JMIR Aging

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

Contents

Original Papers

Reviews

Perceived Use of Web-Based Videoconferencing for Social Connectedness Among Older Adults Living in Long-Term Care: Qualitative Study

Anna Garnett[1], RN, PhD; Halyna Yurkiv[1], BScN, MN; Denise M Connelly[2], PhD; Richard Booth[1], RN, PhD; Lorie Donelle[3], RN, PhD

[1]Arthur Labatt Family School of Nursing, Western University, 1151 Richmond Street, London, ON, Canada

[2]School of Physiotherapy, Western University, London, ON, Canada

[3]Biobehavioral Health & Nursing Science, College of Nursing, University of South Carolina, South Carolina, SC, United States

Corresponding Author:
Anna Garnett, RN, PhD
Arthur Labatt Family School of Nursing, Western University, 1151 Richmond Street, London, ON, Canada

Abstract

Background: The COVID-19 pandemic highlighted how restrictions on in-person interactions within long-term care homes (LTCHs) severely compromised social connectedness among older adults and their families. Post pandemic, despite policy changes supporting greater in-person family engagement, frequent outbreaks continue to disrupt face-to-face interactions, and factors such as geography, life circumstances, and health can constrain family members' ability to make regular in-person visits. Research suggests that web-based videoconferencing technology (WVT) may be a practical solution to help older adults within LTCHs to maintain social connection in the absence of physical gathering. However, increased understanding of end user experience is lacking, and more information on LTCHs' readiness to support and sustain WVT will be needed if this modality is to be successfully and widely implemented.

Objective: This study aimed to understand how older adults living in LTCHs, their families, and LTCH staff members perceived the use and ease of use of WVT devices for facilitating social connectedness.

Methods: Using a qualitative description approach, in-depth semistructured interviews were conducted with 7 older adults, 22 family members, and 10 staff across 3 LTCHs via Zoom (Zoom Communications, Inc), Microsoft Teams, or phone calls. Data were analyzed using a directed content analysis informed by the technology acceptance model.

Results: Findings were structured into 3 main themes: actual system use, perceived usefulness of WVT, and perceived ease of use of WVT. Participants described using a range of WVT hardware and software to promote social connection between older adults and family members. Videoconferencing had a crucial role in supporting older adults and their family members' positive emotional state while also enabling them to maintain life and social roles such as participating in family functions. Despite the perceived use of these tools, participants were concerned about the decline in offering videoconferencing services across LTCHs post pandemic. Some participants noted shifting funding priorities toward supporting in-person recreational activities rather than diversifying web-based social connection options. In addition, factors pertaining to WVT ease of use and integration included limited staff to support older adults with different physical and cognitive needs, variability in digital literacy including knowledge about accessibility features to enhance the ease of use, and families' lack of awareness about the availability of WVT for social connectedness.

Conclusions: Web-based videoconferencing technology has the potential to be a meaningful tool to reduce social isolation and promote a sense of social connectedness among older adults and their families and friends. Future research should explore how WVT could be integrated into care planning for this population, particularly in situations where older adults may be at heightened risk for social isolation. Resource allocation toward equipment, infrastructure, and family and staff training would be well-placed to increase engagement with WVT within LTCHs.

(JMIR Aging 2026;9:e73213) doi:10.2196/73213

KEYWORDS

web-based; videoconferencing; virtual technology; long-term care; social connectedness; older adults

Introduction

Background

Globally, the COVID-19 pandemic impacts were profound, causing widespread loss of life and morbidity [1,2]. During the earliest phases of the pandemic, significant restrictions were placed on many social activities, including normal activities of daily living (eg, shopping and leisure activities) [3-8]. One specific aspect of daily living that was significantly affected included the notion of *social connectedness* [9,10]. Social connectedness can be described as the subjective sense of being in close-knit relationships with others [11,12]. Having interpersonal relationships and being able to gather with others are integral to individuals' health and well-being [13,14]. These avenues for social connectedness can prevent people from feeling lonely and reduce the risk of impaired mental health [14-21]. Canadian older adults residing in long-term care homes (LTCHs) were disproportionally affected by the lack of in-person interactions or access to social spaces over the course of the pandemic [22], resulting in their compromised social and emotional well-being [22-24].

During the height of the COVID-19 pandemic, various web-based videoconferencing technologies (WVTs) were introduced by LTCH staff and family members to help maintain social connectivity between older adults residing in LTCHs and their social circles [25-29]. Leveraging WVTs to facilitate social connectivity among older adults was particularly beneficial as these individuals were at greater risk of isolation and loneliness due to factors such as pre-existing isolation (eg, living alone), being reliant on family for support, and impaired physical health such as frailty [30]. Web-based videoconferencing, the use of networked digital telecommunications technology, is a form of telepresence that "simulates the experience of being physically present in a remote environment" [31]. FaceTime (Apple, Inc), Skype (Microsoft Corp), or Zoom (Zoom Communications, Inc) are among the most popular videoconferencing applications used to continue communication among older adults and family or friends when in-person gatherings are restricted [26,32]. Using WVTs enabled people to visually interact with their family members and, in addition to hearing their voice, offered a greater reassurance of their well-being through a visual confirmation [29]. In addition, seeing older adults within LTCHs using WVTs can help foster family members' belief that their family member was well-cared for [33].

Research suggests that web-based videoconferencing is a viable form of connection for people living in long-term care who may be physically separated from their families and familiar social settings (eg, religious mass services and health care consultations) [34-39]. However, there remains a lack of insight pertaining to the experiences of the individuals directly using this technology. Moreover, additional research is required to inform sustainable and consistent use of WVTs as a reliable option to help older adults maintain social connectedness [40]. Although WVTs offer clear benefits, some literature suggests that this technology may cater more to family members than older adults living in LTCHs, especially if the older adult lives with some degree of cognitive challenge (eg, dementia) [28,33,41]. The types of benefits offered by WVTs across a range of users for the purposes of fostering social connection necessitate further exploration, particularly among older adults with varied cognitive abilities residing in LTCHs. As such, greater knowledge is required to understand how older adults with both physical and cognitive impairments may benefit from or experience challenges in using WVTs [33]. Furthermore, there is a need to understand if and how LTCHs are positioned in terms of infrastructure and staffing to support successful widespread WVT use. For example, arthritis, vision impairments, or neurodegenerative conditions such as Parkinson disease could necessitate the need for adaptive technology or additional assistance as well as greater staff requirements and ultimately cost for the LTCH [42,43]. The potential benefits of WVTs as a tool to support social connection have been shown, but within LTCHs, there is a need to better understand end user experiences if it is to be successfully integrated and sustained as a widespread tool to support social connectedness.

Objective

The purpose of this study was to understand how older adults living in LTCHs, their families, and LTCH staff members perceived the use and ease of use of WVTs as viable modalities for facilitating social connectedness.

Methods

Project Registration

This study presents the findings of the first stage of a multistage qualitative research study described in the published protocol [44].

Study Design

Using a qualitative description methodology [45], this study sought to explore how older adults in LTCHs, their family members, and LTCH staff experience and understand the use of WVTs to support the social connectedness of older adults in LTCHs. Qualitative description methodology enables researchers to capture and describe a wide range of participants' experiences and perceptions in a comprehensive, yet clear and practical way using in-depth interviews with study participants [45,46]. In addition, qualitative description is a low-inference approach used in social sciences and health care to provide researchers with a practical way to derive rich, close to data descriptions of participant experiences without the need to interpret the meaning of participant experiences [46]. In addition to in-depth interviews, observational field notes were kept throughout this study, thereby enhancing the dataset and the overall study rigor [47]. Studies using qualitative description methodology seek to find answers that are of high relevance to practitioners or policymakers to improve the quality of a delivered service, including participants' thoughts or attitudes toward a service; why do participants use certain services more so than others; and how participants use a service in specific instances [45,48].

Theoretical Framework

The technology acceptance model (TAM) [49] was used to shape interview questions and to organize and report qualitative findings in this study. TAM is a widely used model that helps researchers uncover how end users adopt the use of technological systems and how this adoption is influenced by constructs such as usefulness and perceived ease of use [50]. According to TAM, the actual system use construct concerns the observable use of technology [49] and, in the context of this study, refers to the use of WVTs by the study participants (older adults, family members, and staff in LTCHs) for the purpose of social connection. The perceived usefulness construct is the extent to which technology end users perceive technology as enhancing their tasks or helping them achieve a particular goal [49]. In this study, perceived usefulness is the degree to which participants believe that WVT (devices and software) facilitates social connectedness or makes it more meaningful. The perceived ease of use construct pertains to the amount of effort required to operate the technology by end users [49] and, in this study, reflects participants' perceptions of how simple it is to operate or navigate WVTs to engage in video calls with friends and family members.

A range of open-ended interview questions was developed using the TAM constructs to guide the in-depth interviews with study participants. For a full list of interview questions, please see Multimedia Appendix 1 in our protocol [44]. In line with qualitative descriptive methodology supporting the use of deductive approaches (ie, TAM) to guide data analysis [51,52], this study presents descriptions of participants' experiences organized under the following TAM constructs: the (1) *actual system use*, (2) *perceived usefulness*, and (3) *perceived ease of use*.

The Use of TAM With Qualitative Description Approach

Findings grounded in qualitative description methodology result from staying close to the words and events in the research data capturing straightforward descriptive accounts of participant experiences [45,46]. In some qualitative descriptive studies, findings are categorized under constructs of a chosen theoretical framework [51]. This study incorporated TAM as a framework for organizing data-driven coding and practical and identifiable constructs of the TAM that may be of use to policy makers. The process of incorporating TAM into the data analysis phase ensured that participant data were organized and described in a structured way without imposing subjective interpretations or other theoretical assumptions on the data. Such an approach aligns with previously published research that successfully integrated TAM with qualitative description methodology to describe older adults' adoption of digital modalities [51,53,54].

Sample

The desired sample size for qualitative description studies is influenced by factors such as the subject or phenomenon of interest, its representation within the broader population, and the potential variability of experiences with the phenomenon [55]. As such, the recruitment goal of this study was 45 individuals comprised of 3 groups: older adults residing in LTCHs, their family members, and associated staff from 3 participating LTCHs located in southwestern Ontario, Canada. Eligible participants were English-speaking individuals with mild to moderate cognitive disability and able to participate in an interview using the Zoom or Microsoft Teams videoconferencing platforms or a phone. Potential older adult participants with mild cognitive decline living in LTCHs were assessed [56] for suitability to participate in this study by a member of their circle of care within the LTCHs before any study information was shared with them.

Ethical Considerations

This study was reviewed and approved by the Western University Research Ethics Board (2024-121993-91059). Participants were informed about the research study, and their questions were answered by the study research assistant before they consented to participate. Collected data was anonymized and participants who completed an interview were given a $10 (CAD or $7.20 USD) gift card as an honorarium.

Recruitment and Data Collection

Participant recruitment and data collection took place between July 2023 and January 2024 using purposeful and maximum variation sampling to help ascertain a sample with a variety of experiences with WVTs, especially among the older adult and family member participants. Research information was distributed to prospective participants via on-premises physical posters, emails, LTCH newsletters, and word of mouth. Individual in-depth interviews with older adults, their family members, and LTCH staff were offered in person, via phone, or by Zoom videoconferencing to accommodate individual preferences and took place at an agreed-upon date and time. Recreational therapy department staff or a family care partner assisted older adult participants in joining the Zoom or interview if necessary. Interviews with participants were audio recorded and transcribed verbatim prior to analysis.

Data Analysis

The collected data were anonymized and analyzed using NVivo (version 13; Lumivero) qualitative analysis software [57] and directed content analysis, a qualitative analytic technique characterized by a structured stepwise process of coding the data and organizing it into categories and themes [58,59]. The first step of the data analysis entailed deidentification of the cases. Interviews collected from family members were assigned identification such as FM followed by the number of the participant (FM01, FM02, and so on). Similarly, older adults were assigned identification of OA (OA01, OA02, and so on), and staff participants were assigned identification S (S01, S02, and so on). The next step involved reading the interview transcripts, developing memos, and noting down compelling quotes. A formative coding framework was then developed based on the TAM constructs with consensus reached through discussion of the research team. The key constructs of the TAM, perceived usefulness and ease of use of technology, served as initial codes or units of analysis. Data that did not lie within one of the key codes were labeled with a new code and organized into new categories or as subcategories under the initial codes [58]. The preliminary coding framework was tested against the same 5 interview transcripts. Researchers then met on a biweekly basis to refine the coding framework by establishing links between generic and main conceptual categories. The newly identified codes were compared and discussed for similarities and differences until consensus was reached regarding the emergent patterns as the concepts became more densely packed with meaning and evidence. Study rigor was ensured through conducting member checking [60], acknowledging biases researchers may introduce to the research and practicing reflexivity throughout the research processes, and ensuring confirmability of data through triangulation between interviews and field notes [61].

Results

Demographic Characteristics Overview

The total sample comprised 39 participants, which is congruent with sample sizes in other studies using qualitative descriptive methods [51]. The sample included older adults (n=7, 18%), family members (n=22, 56%), and staff (n=10, 26%) from 3 participating LTCH facilities, offering diverse perspectives on WVT use in these settings. The 3 participating homes were licensed, not-for-profit LTCHs spanning 120 kilometers and providing comprehensive care to between 160 to 394 individuals, including an emphasis on person-centered care.

All 7 (18%) older participants were female and aged between 64 and 95 years. Older participants were widowed (28/39, 71%) or divorced (11/39, 29%). All 39 older adults either had a college diploma (n=22, 57%) or a university degree (n=16, 42%), providing some context that may be related to the ease of using WVTs. The majority of older adults (n=28, 71%) reported living with cognitive or physical challenges (eg, poststroke impairment, dementia, and arthritis), which provides essential context for understanding the impact of cognitive or physical limitations on WVTs' usefulness, ease of use, and intention to continue using WVTs. Moreover, 31 out of 39 (nearly 80%) participants in this study identified that physical (eg, arthritis) or mental disability (eg, dementia) affected individuals' ability to use WVTs.

The 22 family members (FMs) consisted of 18 female and 4 male participants, and were aged between 40 and 69 years. Many family members held a university bachelor's degree (16/39, 41%), with a notable portion being retired (14/39, 36%).

All 10 staff participants were female, with just under half (10/39, 40%) working in their profession between 6 and 10 years. Staff participants included a range of health professionals such as recreational therapists (12/39, 30%), social workers (8/39, 20%), personal support workers (8/39, 20%), a registered practical nurse (4/39, 10%), and a spiritual care practitioner (4/39, 10%).

Thematic Presentation of the Findings

Overview

Study findings on participants' roles (older adults, FMs, and staff) and experiences in using WVTs for the purpose of social connectedness are presented in themes informed by the TAM and include the actual system use, with 2 subthemes—(1) WVTs' hardware and software and (2) trajectory of using WVTs from pre- to postpandemic; perceived usefulness, with 4 subthemes—(1) enabling remote connection, (2) providing emotional and psychological benefits, (3) fostering continuation of life and social roles, and (4) the enriching effect of video presence; and perceived ease of use, with 3 subthemes–(1) design and practicality of the devices and platforms used, (2) usability and accessibility of the devices, and (3) impact on the workload (Textbox 1).

Textbox 1. Summary of technology acceptance model–informed themes and subthemes.

Theme 1: actual system use

- Web-based videoconferencing technologies' hardware and software
- Trajectory of using web-based videoconferencing technologies from pre- to postpandemic

Theme 2: perceived usefulness of web-based videoconferencing technologies

- Enabling remote connection
- Providing emotional and psychological benefits
- Fostering continuation of life and social roles
- Enriching effect of video presence

Theme 3: perceived ease of use of web-based videoconferencing technologies

- Design and practicality of the devices and platforms used
- Usability and accessibility of the devices
- Impact on the workload

Theme 1: Actual System Use

Overview

The actual system (WVTs) use theme consisted of 2 subthemes: (1) the hardware and software used by older adults, FMs, and staff, and (2) WVTs use trajectory from pre- to postpandemic timeframes.

WVTs' Hardware and Software

During the pandemic, LTCH staff introduced iPads and tablets as a virtual means to facilitate social connectedness between families. Older adults and families also used a range of personal devices with videoconferencing capabilities such as iPads or tablets, laptops, computers, smartboards, and smartphones. Moreover, some families independently researched and purchased nontraditional devices that FMs suggest have more merit for people with physical or cognitive conditions (eg, dementia, arthritis, and paralysis) to access the WVTs they were using, namely, Portal, Alexa Echo Show, and ViewClix hardware. Participants also used various software platforms for WVTs, including Zoom, Teams, WhatsApp, Skype, Facebook Messenger, Signal, and FaceTime (Table 1).

Table . Web-based videoconferencing technologies used across study participants (N=39).

Web-based videoconferencing technology	Older adults, n	Family members, n	Staff, n	Total, n
Hardware types				
iPad or tablet	6	15	10	31
Computer	3	6	1	10
Laptop	0	4	2	6
Smartboard	0	0	2	2
Smartphone	2	8	3	13
Portal	0	2	0	2
Alexa Echo Show	0	1	0	1
ViewClix	0	1	0	1
Software types				
Zoom	0	10	7	17
Teams	6	1	1	8
WhatsApp	0	1	0	1
Skype	0	3	2	5
Facebook messenger	1	2	2	5
Signal	0	2	0	2
FaceTime	1	12	3	16

Trajectory of WVT Use From Pre- to Postpandemic

The iPads or tablets were predominantly leveraged by LTCHs' recreational department teams during the pandemic-related lockdowns and resultant restrictions on in-person visitations as an alternative to socially connect FMs and older adults residing in LTCHs. However, post pandemic, this service was not unanimously integrated into LTCHs' care delivery. Some LTCHs have embraced the technology, integrating it into daily routines or care planning, while others have shifted funding priorities toward in-person recreational activities, which some FMs shared were not always as meaningful as the ability to see and interact with their loved ones on the screen.

Participants reported varied knowledge of acceptance and normalization of WVT use across different LTCHs from pre- to postpandemic. Most participants who used WVTs to socially connect with people outside LTCHs prepandemic continued to do so during and post pandemic. For these participants, it was the only way to connect with family who lived abroad or far away from LTCHs.

Among the study participants, 14 (36%) participants stated they began using WVTs to support social connectedness due to the pandemic (7 FMs, 5 staff, and 1 older adult). Eleven (28%) participants shared that WVTs continue to be offered by the LTCHs or that they still use it post pandemic (6 FMs, 3 staff, and 1 older adult), while 10 (26%) participants believe the service has stopped being offered post pandemic by the LTCHs (5 FMs, 3 staff, and 1 older adult). A few reasons were provided by the participants to explain the decline in use post pandemic. First, a shift of priorities to do more in-person activities and visitations resulted in a lack of recreational department staff support to help with videoconferencing service. As 1 FM whose mother was admitted to LTCH in March 2023 shared:

> *I've never had it offered to me with my mom. I don't think the emphasis is on it and part of it is because the reason they had time to do more of that [videoconferencing] was because so many other programs were cancelled, right?* [FM21]

A second reason provided was a lack of awareness among older adults and their family members that connecting socially using WVTs was still being offered in the LTCHs:

> *If they have anything to offer, they're certainly not advertising it, I think they don't have it. I think that it was simply cut.* [FM19]

A similar concern was echoed by a staff member participant:

> *I don't even know how much recreation [department] even does help with Zoom calls anymore because all family – like anybody can come in now, right? Or they just call too, right? I don't really see much Zoom,*

> *FaceTime, or anything anymore with residents really.* [S07]

The third identified contributor to the decline of WVT use post pandemic shared by the participants was related to the workload challenges imposed by supporting the use of WVTs to socially connect people:

> *We just simply don't have that manpower to be engaging in this environment as well as supporting you to have a conversation with someone. Even though we understand the value in that, and how important that is for their emotional health, you know, when they are feeling positive and good and they know that their family is safe, they don't have responsive behaviours.* [S02]

In summary, the use of WVTs in LTCHs demonstrated benefits for maintaining social connectedness during the pandemic. However, the transition to postpandemic society revealed a decline in its continued use and only partial normalization of videoconferencing across the 3 LTCHs in this study. The trajectory of WVT use from pre- to postpandemic provides insights into the fluid role that technology can have in supporting social connectedness of older adults in long-term care settings.

Theme 2: Perceived Usefulness of WVTs

Overview

The perceived usefulness of WVTs for social connectedness is represented by 4 subthemes: enabling remote connection, providing emotional and psychological benefits, fostering continuation of life and social roles, and the enriching effects of video presence.

Enabling Remote Connection

WVTs enabled residents to connect remotely with family and friends, thereby bridging physical distances that in some cases were difficult for FMs to overcome. This opportunity was particularly important for those families who lived at a great geographical distance from the LTCHs or were unable to visit frequently due to various constraints:

> *We'll FaceTime with family members that are far away. I mean, last year, we were basically in four different time zones. My sister was Central, I'm Eastern, my daughter was in the U.K. and one of my nephews was in Japan.* [FM09]

Conversely, another participant noted a robust example of how WVTs could connect people in the same building:

> *At one point [my parents] had to be separated and put on separate floors, so we used a lot of FaceTime and – is it Facebook Messenger? Yes, we were using that between the two of them. So, two of us would have to go in every day and one to my father, one to my mother, so that the two of them could talk.* [FM03]

In addition, videoconferencing was a useful tool that enabled family to note health care–related details in older adult's presentation, which they later communicated to the staff, as shared by this FM:

> *I recognized when she had an eye infection, or I could see that her hands were really dry. I couldn't have done that if I was on the phone. We were able to see things and it informed us more about how my mother really was, and it improved her care.* [FM18]

WVTs were also used for accessing care as an accessible alternative in place of the traditional in-person meetings:

> *For the annual team meetings, we had the option to do that virtually or in person. And my sister doesn't drive, and the care home is a long way from her place, so she chose to meet with the care staff virtually, which she really appreciated.* [FM02]

Those who have not used WVTs for health care access expressed receptiveness to this possibility (FM12). However, those FMs who lived within proximity to the LTCH preferred in-person meetings for such matters. For instance, 1 FM shared that*:*

> *there is an annual health care review, and they might have offered that as an option, but I'm just ten minutes from where my mum is, so you know, I would just come in for that meeting.* [FM16]

Another FM raised an interesting point regarding the nonverbal communication that takes place during the in-person meeting that can be missed or hard to sense over the video call:

> *In-person you sense if there's anything going on. You sense it more because when you are videoconferencing, you may see just one person at a time. But when you're sitting in the room, we're all sitting around, and I see everybody, and you feel something if there's something wrong, if they're not telling you everything. On videoconferencing you can't feel it. You just see it, but you don't feel like there's any issues.* [FM22]

These quotes exemplify how WVTs ensured continuous social engagement, helping many residents feel less isolated and more connected to the outside world during lockdown or when they were in isolation.

Providing Emotional and Psychological Benefits

Regular videoconferencing interactions helped support older adults' mental health by reducing feelings of loneliness and isolation. The web-based presence connection with family and friends was a vital source of emotional support. Older adults and their FMs experienced substantial emotional benefits from connecting through the video call via WVTs. It provided them with comfort, joy, and a sense of connectedness, which heightened their positive emotional and psychological well-being:

> *The biggest thing it helps with is their mental health, especially if they're just making that change into long-term care. Or if they've been in there, to help lift them out of some bad times or bad moments, because I know that depression in long-term care is huge. There's 24 hours in a day, and if they're sleeping 10 hours and they only get four hours devoted time from someone, there's a lot of time in between there. So, if there could be some way that*

they could speak to friends or family or anyone else more, then I think it would decrease depression in long-term care. And that would be big. [FM07]

Other examples of the ability of WVTs to provide emotional and psychological benefits include its potential to anecdotally delay the decline associated with a chronic disease such as dementia:

I think, for the dementia and Alzheimer's, for her [mom] to visually see the relatives on whatever we use, whether it's Zoom or FaceTime or Portal, and connect the face with the voice has sort of delayed her [mom's] decline, because you can actually see them. [FM09]

In addition, WVTs were used for prevention of responsive behaviors associated with dementia:

They [staff] proactively call, because they figure it'll help settle her [mom] down. So now they do FaceTime because they obviously realize that if she [mom] can see somebody then it calms her down even more than just a telephone call. [FM09]

Fostering Continuation of Life and Social Roles

Fostering continuation of life and social roles was another subtheme associated with WVT use. For example, WVTs provided an avenue to maintain a sense of normalcy by allowing older adults to continue participating in daily life activities and social roles beyond LTCH walls. For instance, older adults kept in touch with FMs and things that were meaningful to them (eg, their pets), remained updated on family events, and engaged in routine conversations that they would typically have in-person prior to being admitted to LTCHs:

I was trying to keep [mom] updated, how her pet was doing... show her visually–I would take her for a walk, through the gardens. [FM07]

Participants found WVTs useful because they allowed them to fulfill their roles as parents, grandparents, and friends. Even if through the screen, they could still offer advice, participate in family decisions, and stay involved in the lives of their loved ones.

In addition to continuing to fulfill their social roles within the family and household, WVTs allowed older adults to participate in other social events. For example, 1 staff member described how WVTs were used for facilitating music concerts for older adults*:*

We would get music entertainers at home through video chat and each home area would have on their smart board to get the video chat going. [S03]

Another example was using videoconferencing to stream mass in churches facilitated by a spiritual care practitioner (1 staff). As a result, 1 older adult requested to attend a virtual mass daily, as it allowed this resident to feel connected to a church community they belonged to prior to the pandemic. These examples demonstrate the expanded role of WVTs in fostering a sense of connectedness with other community members.

Enriching Effects of Video Presence

Compared to other means of communication (eg, landline phone), WVTs were described as having an enriching effect on social connectedness. Participants shared that being able to see faces, share visual experiences, and express emotions more vividly via the screen contributed to a deeper sense of connection:

How many times you read an email where you take it wrong. How many times you get on a phone, and you don't understand a person may be sounding angry, but they're visually upset. I think the video adds a lot in those situations, particularly since in a nursing home it is so emotional. [FM01]

In addition to the interactions between family and older adults, participants described how videoconferencing allowed them to record interactions and watch them later or send them to others (FM01) or observe the resident-staff interactions:

A side benefit is, every so often I'll catch one of the staff members in the room with him [my husband], so I have a nice chat with the nurse. I live next door, so I'm there all the time anyway, but it gives me a chance to see the interactions between him and the staff. [FM17]

In summary, these findings underscore the multifaceted usefulness of WVTs in maintaining social connectedness and well-being not just for older adults, but also for FMs. Usefulness of WVTs may be hindered or enhanced based on how intuitive and user-friendly these technologies are for older adults, FMs, and LTCH staff.

Theme 3: Perceived Ease of Use of Web-Based Videoconferencing Technologies

Overview

The perceived ease of use was informed by 3 subthemes: design and practicality of the hardware and software, usability and accessibility of the WVTs, and ease of use in relation to the impact on the workload for staff members supporting WVT use.

Design and Practicality

Factors impacting the external design of the devices, such as size of the screen, weight, and the availability of supportive accessories like stands, iPad covers, stylus, external speakers, and headphones influenced the practicality and ease of use of WVTs for older adults.

Hardware devices with larger screens, clear audio, and simplified navigation were reported as easier to use. Accessories like headphones or speakers also improved the ease of use. Participants also benefited from using devices with software interfaces requiring minimum to no effort to operate:

With our mom, she didn't have to know anything, she just would see faces pop up. [FM04]

Portable hardware devices were preferred by some, as they could be conveniently moved and adjusted according to the residents' needs and preferences and provided more privacy (4 staff). For instance, 1 FM described how they enjoyed going for a walk into the garden and video calling their family from that location. At the same time, portability potentially breached privacy when residents were situated in an open public space like the dining area due to a lack of private space or convenience for staff (1 FM and 2 staff). Eight (21%) participants also expressed their concern for device misplacement, theft, or a risk of accessing passwords and private information through older adults' personal devices (1 older adult, 5 FMs, and 2 staff), which could be amplified by the devices' portability.

Usability and Accessibility

The usability and accessibility of the devices were key factors impacting participants' use of WVTs. In this study, 31 out of 39 (nearly 80%) participants identified physical (eg, arthritis) or mental disability (eg, dementia) as the primary reason impeding their ease of using WVTs for social connectedness. Both staff and FMs described older adults' difficulties in comprehending the concept of videoconferencing, which often resulted in disengagement during the video call. Despite this, FMs still appreciated the ability to see their loved ones on screen, even if for a brief period. However, this stimulated FMs to seek other easier-to-use WVTs requiring only the FM to start and end a video call.

In addition to physical or cognitive abilities, digital literacy also varied among participants. Staff noted that the pandemic catalyzed WVT use, and as a result, they had to quickly adapt and engage in learning to use the technologies introduced by their workplaces. Families also engaged in mobile learning either individually or with the help of their technologically literate friends or family. These FMs often assumed responsibility for teaching older adults living in LTCHs how to use WVTs. For those older adults who were particularly isolated or for FMs who lacked support, staff members became responsible for teaching, setting up, or troubleshooting the WVTs for them. One staff member shared a story about teaching a husband outside the LTCH how to download and set up Skype so that he could speak to his isolated wife in the LTCH during the pandemic. Some reported that the assessment for the family and older adult's capacity to use technology and planning to support the ease of use would usually take place during the resident admission into LTCH (2 staff), although this was not always the case. Findings also suggest that some families were unaware they could set up video calls upon admission (3 family members).

Surprisingly, only 1 FM explicitly received special training on accessibility features embedded into the device through their workplace to enhance its ease of use for people with varied cognitive and physical abilities. Moreover, only a few (3 FMs and 1 staff) were aware of the accessibility features and how to enable them for easier device use.

Impact on the Workload

The ease of WVT use impacted the workload among staff within LTCHs. Setting up and supporting older adults with videoconferencing was seen as an additional task to do and required coordination between more than 1 health care team, which led to some resistance to use WVTs among staff. One staff member deployed to assist residents with setting up video calls shared:

> *I wouldn't get the residents ready myself; I would ask a personal support worker to get the resident ready. A lot of times I would go to the floor and some of the residents would still be in bed, so I'd be like: "Someone needs to get this person ready for me because it's out of my scope.* [S05]

Another frontline staff member shared that not all staff were willing to facilitate the video calls*:*

> *There's one resident who has her own tablet, who will ask us: "Can you call my daughter?" And we will go and do it for her. Now, when I say we, it's mostly me and one other person. A lot of people don't want to stop and facilitate that because a lot of times there's troubleshooting with a tablet. And people aren't always comfortable doing that kind of thing. Or they don't feel like it's their job to do.* [S09]

Interestingly, staff reticence to assist with web-based videoconferencing technologies use was received with understanding among family and older adults (1 older adult and 1 FM) with some noting that staff already had many work demands and were too overworked to be helping them with videoconferencing. However, 1 FM also added that contextualizing the reason for a video call (eg, social connection) could overcome the staff's resistance to help with setting it up and staying with the person for the duration of a video call if needed:

> *[…] if you had a palliative resident in the middle of the night that wanted to say last good-byes to somebody, wouldn't it be wonderful to be able to connect them virtually with somebody that couldn't be there. I think if it were put in the context like that, it could be woven into some of the other training that you give your nursing staff.* [FM21]

The perceived ease of use of WVTs in this study was influenced by a range of factors, from choosing the ergonomically suitable device to calibrating it based on unique needs and integrating its use into a care plan given the realities of staff workload.

Discussion

Principal Findings

This qualitative study examined perceptions of usefulness and ease of using web-based videoconferencing hardware and software to facilitate social connectedness between older adults living in LTCHs and FMs. In-depth interviews conducted with 3 participant groups—older adults living in LTCHs, FMs, and LTCH staff—suggest that many older adults and FMs benefited from WVTs, and families were creative in finding workable solutions to facilitate older adults' use of technology despite their potential physical and cognitive limitations. Findings informed by the TAM model were collated in 3 overarching themes: actual system use, perceived usefulness, and perceived ease of use to help inform understanding of WVT use in LTCHs. Key findings in this study include introduction of the WVTs during the COVID-19 pandemic positively impacted emotional well-being of socially isolated older adults in LTCHs and their FMs when in-person visitations were restricted; design and ergonomics greatly influenced the use of WVTs, and staff along with FMs were instrumental in finding ways to make WVTs useful and easy to use, especially for older adults with cognitive or physical impairments; and there was a notable decline in WVT use in LTCHs for social connectedness post COVID-19 pandemic despite heavy reliance on WVT programs to maintain a sense of connectedness during the pandemic.

Actual System Use

Besides commonly used WVTs like iPads or Zoom platform, FMs and staff participants in this study who provided care to the older adults living with dementia used several devices such as Portal, ViewClix, Alexa Echo Show, and Signal, which are understudied in the literature. For instance, a literature search yielded only 1 study exploring Alexa Echo Show to maintain social connection among older adults without impaired cognitive abilities during the COVID-19 pandemic [62]. Smartphones, iPads, or laptop devices, as well as FaceTime, Skype, or Zoom digital software are the most commonly used WVTs reported in the literature to facilitate social engagement and residents' emotional well-being [25,29,36,41,63-66]. These hardware and software are useful for many people; however, they might not have embedded accessibility features required for people with special needs, which was noted in this study. In addition, WVT users may require additional assistance, either provided by staff (which may not be realistic to integrate in the current workflow) or due to the variable availability of informal support by friends or family.

Participant experiences highlight the limited number of formal roles (eg, recreational activity specialists), resources (eg, tailored technologies and their maintenance), and programs (eg, digital literacy workshops) to support ongoing WVT use. The lack of WVT use for social connectedness, particularly post pandemic, signals an emphasis on in-person activities and, concomitantly, ongoing staff shortages, which are also reported in other research [67]. The sustained use of WVTs for social connectedness in this study was found among those who were most familiar and comfortable with technology and who were separated by great geographical distances, a finding corroborating other research [66]. However, there were FMs of older adults in this study who were unaware of the ongoing availability of WVTs despite expressed interest in using this modality. This indicates a growing preference for multiple modes of fostering social connection with those residing in LTCHs.

Perceived Usefulness

Study findings demonstrated that LTCH staff valued using WVTs with families as a means to enhance social presence and for its positive effects on older adults' well-being, even if they had advanced dementia and were unable to fully comprehend how videoconferencing functioned, which is echoed in other research [36,63,67,68]. Participants in this study and other research [69] reported increased usefulness of WVTs for social connectedness facilitated by LTCH staff or family members, as it allowed older adults to remain remotely connected with their families despite restrictions on in-person visits. Although some of the older adults with more pervasive cognitive shortcomings found it challenging to engage in conversation, videoconferencing experiences still elicited positive reactions and emotions, in accord with findings reported in the literature [36,63,67]. Despite the potential challenges of WVT use, they were perceived as a valuable tool to support well-being, foster engagement between older adults and their family, and enhance patient-centered care [67].

Notable was the lack of clarity voiced by some participants about the availability of and ability to use WVTs, particularly once the acuity of the pandemic had subsided. Although WVT use declined post pandemic in LTCHs, older adults and their FMs described it as a useful and desired service for ongoing use. To achieve optimum quality of life while residing in LTCHs, older adults and their FMs expect greater opportunities for social interactions and higher quality and quantity of family engagement [70], which can be addressed by using WVTs. For example, participants in this study shared that videoconferencing allowed them to stay connected to their life and maintain habits established before moving to LTCHs (eg, virtually visiting the garden they used to walk through and interacting with their pet). This propensity of WVTs to provide access to other contexts may foster a sense of aging in place [71] among older adults who are living in LTCHs.

In addition to supporting social well-being, WVTs may also have an important clinical role, in the form of telemedicine, as these modalities reduce time needed to travel, allow timely access to specialized health care personnel [37] and are a cost-effective solution that could alleviate the need for emergent visits [72]. The added examples of usefulness make it worthwhile not only to sustain but also expand WVT service [73].

Perceived Ease of Use

Participants in this study independently researched and shared their knowledge of Portal, ViewClix, or Alexa Echo Show. These WVTs were perceived as easier to use for people with cognitive impairment (eg, dementia) than iPads. Although iPads have been commonly leveraged for their convenience and wide societal acceptance, these types of mobile devices often presented challenges for those older adults who had cognitive or physical disabilities or who lacked experience in using WVTs. In line with this finding, higher tablet or iPad use is predicted by factors such as younger age, higher cognitive functioning, and absence of hearing impairment [66]. The use of devices such as Portal, ViewClix, and Alexa Echo Show highlighted in this study warrants further exploration. These devices could be well-situated to address some of the identified challenges of WVT use among older adult populations and in those with increasing health complexity, such as is often found within LTCH settings [74]. This is especially relevant because a lack of social contact was identified as a risk factor for developing dementia, and preventative action is best initiated early [75].

Findings in this study also suggest that staff members noted digital literacy challenges among some FMs as well as older adults living in LTCH, which is echoed across other studies [63,67]. Low digital literacy among older adults may amplify social isolation by excluding these individuals from participating in social processes or spaces [76,77], many of which have shifted their operation to digital spaces in recent years. Moreover, since in many instances the FMs were pivotal for initiating the video calls [36], low digital literacy among them resulted in additional responsibility being assumed by LTCH staff who were required to provide education and manual support [36,67]. These findings highlight the need for broad approaches to foster widespread WVT uses among older adults that consider cost, digital literacy, individual capabilities, usability of these tools, and training and use requirements.

Although WVTs are challenging to use for some older adults [64], looking to the future, more people will likely expect to use and be supported in using WVTs in LTCHs as younger generations of society use these devices regularly and consistently within their daily lives. The importance of digital literacy training and accessibility features to facilitate independent use of WVTs by older adults is documented in the literature [78]. This study also identified limited awareness about, and knowledge of, accessibility features embedded in the WVTs by participants which, in turn, could enhance the ease of their use. Although previous works have identified low digital literacy among older adults [64] and a lack of training available within LTCHs to guide their use of WVTs [36,66], there remains a gap in research regarding implementing digital literacy training programs with older adults in LTCHs and their families. In fact, research suggests that in some cases, FMs have been the ones to take initiative and provide the technology necessary for WVTs within LTCHs [79]. Moving forward, additional funding and education to increase technology availability, enhance digital literacy, and increase staffing support roles within LTCH are warranted.

Staff participants in this study and in others [36,63,67] expressed concern regarding their workload and challenges to integrate videoconferencing with their other roles and responsibilities. Interestingly, a few FM participants in this study were accepting of limited opportunities to connect with their older adult member using WVTs due to staffing shortfalls. It is plausible that widespread awareness of the workload and employment challenges common within LTCHs across Canada contributed to their acceptance [80]. This finding demonstrates families' awareness and compassion for those employed within LTCHs, but it also highlights the ongoing paucity of resources directed toward supporting the psychosocial well-being of older adults within LTCHs. In addition, it highlights the lack of policy to support widespread use of WVTs in LTCHs. For instance, burdened staff being too busy to provide assistance with WVT use is documented in the literature [66,79]. Institutional supports are paramount to sustain the WVT service for social connectedness and include additional staffing and redefining roles and responsibilities among LTCH staff to foster digital literacy among older adults in LTCHs or, alternatively, to assist older adults and their families in using WVTs [81]. Addressing these factors simultaneously may facilitate smoother integration of this service into LTCHs while limiting the burden on older adults, their families, and staff.

Implications

Findings from this study highlight the importance of a breadth and depth of approaches to effectively implement and sustain the integration of WVT programs for social connectedness within LTCHs. Government and the private sectors must allocate funding to allow for technology procurement and infrastructure improvements to support wide-scale WVT use within LTCHs. This could be achieved by targeted funding grants or through capital funding initiatives. To maintain WVT use over time, policies and funding to guide its use are also required. Recognition that personnel is required for successful long-term use is paramount if older adults and families are going to benefit. There could be opportunities for expanded roles or even volunteer positions to help older adults use WVTs. In terms of choosing the WVTs, guidelines that ensure accessibility and user-centeredness for individuals with cognitive or physical impairment are required to ensure equitable use of WVTs. In anticipation of the increasingly digitalized health care and potential future increase in using WVTs by individuals in our society, additional roles for staff should be created to be able to support, sustain, and expand WVT programs for social connectedness and health care in LCTHs.

Interdisciplinary care teams in LTCHs can leverage WVTs in their daily routines or activities to improve older adults' physical and psychosocial well-being and create opportunities for more family engagement. Care teams should create opportunities for older adults and their families to engage in digital literacy workshops to optimize benefits from using WVTs by a greater number of older adults and their families. Staff members should also engage in digital literacy training with a particular focus on troubleshooting WVTs and innovating ways they can incorporate WVTs into their daily activities to promote a sense of meaningful social connectedness among older adults residing in LTCHs.

This study highlights the need to explore long-term impacts on health and well-being of using WVTs with residents and their families. Future studies should focus on exploration and comparison of WVTs to determine the most appropriate WVTs to promote social connectedness, especially for individuals living with cognitive or physical impairments. In addition, metrics on the impact of using WVTs for older adults' well-being should be collected to determine different aspects of WVTs' usefulness for older adults in LTCHs. Future research studies should also explore or develop sustainable models to enhance digital literacy training for older adults and their families.

Strengths and Limitations

A strength of this study is the robust participant sample representing multiple perspectives and experiences in using WVTs for social connectedness. Although staff participants were all female, this parallels research suggesting that 90% of personal support workers employed in Ontario are female and 75% of care workers across Canada are female [82,83]. In addition, collaborating with 3 LTCH sites for participant recruitment supported collection of a diverse range of experiences across the different settings. This mitigated a risk for biased results that would be imposed by the nature and operations of a single facility.

While efforts were made to recruit a range of older participants living in LTCHs, we recognize that the inclusion criteria of mild or no cognitive impairment constrained the participation of older adults. Those who had advanced dementia or other severe neurocognitive deficits were not included, which likely represented a notable portion of those residing in LTCHs who live with dementia [84]. Those older adults who had mild deficits or other physical limitations that impacted their ability to participate were offered assistance in setting up the technology so they could still participate. The findings revealed that frontline staff, particularly nurses and personal support workers, were instrumental in facilitating WVT use among older adults and their families, although this responsibility posed potential burdens on staff workload. Future work could address targeted personnel support for WVT use, including the potential for volunteer technology facilitation or expanding frontline provider roles to include assisting older adults to use technology such as WVTs. Experiences of staff in various LTCH facilities may differ in terms of digital literacy training provided to staff or role expectations. Moving forward, it would be helpful to have clear guidelines to inform staff roles in providing technology support, particularly in the case of supporting older adults' well-being, including social connectedness. In addition, this study did not delve into other factors that may contribute to the decline in using WVTs post pandemic, like for instance, the complex psychosocial factors that influence technology adoption.

Conclusions

The experiences of older adults residing in LTCHs, FMs, and staff demonstrate that using WVTs for social connectedness positively impacts older adults and their FMs' emotional and social well-being. However, WVTs service needs to be tailored to the needs of the families, including the choice of the device, digital literacy training, and provision of human resources to support connections. Moving forward, LTCHs should develop formal programs that allow for integration of WVTs service to expand the opportunities for older adults in the LTCHs to connect socially with their families or use WVTs in other ways that create a sense of social connectedness. To achieve this, funding initiatives such as capital improvement funds and clearer policies on the roles of personnel support will be required. Prospective studies should implement WVTs in collaboration with other actors participating in the process of socially connecting people with WVTs, such as technology industry partners, companies allowing people to participate in or attend leisure activities virtually, and public health organizations to explore additional impacts of WVTs on social and clinical well-being of older adults in LTCHs and their families.

Acknowledgments

The authors acknowledge Jessica Burford and Christina Oleynikov for their assistance with this project.

Funding

This study was supported by a Social Sciences and Humanities Research Council Insight Development Grant from the Government of Canada. The funders had no role in the study's design; collection, analyses, or interpretation of data; writing of the manuscript; or decision to publish the results.

Authors' Contributions

Conceptualization: AG
Data curation: AG (lead), HY (supporting)
Formal analysis: AG (equal), HY (equal), RB (supporting), LD (supporting), DMC (supporting)
Funding acquisition: AG (lead), RB (supporting), LD (supporting), DMC (supporting)
Investigation: AG (lead), HY (supporting)
Methodology: AG
Project administration: AG (lead), HY (supporting), RB (supporting), LD (supporting), DMC (supporting)
Resources: AG
Supervision: AG
Validation: AG
Visualization: AG (lead), HY (supporting)
Writing–original draft: AG (equal), HY (equal)
Writing–review & editing: AG (lead), HY (supporting), RB (supporting), LD (supporting), DMC (supporting)

Conflicts of Interest

None declared.

References

1. Aburto JM, Schöley J, Kashnitsky I, et al. Quantifying impacts of the COVID-19 pandemic through life-expectancy losses: a population-level study of 29 countries. Int J Epidemiol 2022 Feb 18;51(1):63-74. [doi: 10.1093/ije/dyab207] [Medline: 34564730]

2. Wang H, Paulson KR, Pease SA. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020-21. Lancet 2022 Apr 16;399(10334):1513-1536. [doi: 10.1016/S0140-6736(21)02796-3] [Medline: 35279232]
3. Felipe SGB, Parreira Batista P, da Silva CCR, de Melo RC, de Assumpção D, Perracini MR. Impact of COVID-19 pandemic on mobility of older adults: a scoping review. Int J Older People Nurs 2023 Jan;18(1):e12496. [doi: 10.1111/opn.12496] [Medline: 35941278]
4. Trevissón-Redondo B, López-López D, Pérez-Boal E, et al. Use of the Barthel index to assess activities of daily living before and after SARS-COVID 19 Infection of institutionalized nursing home patients. Int J Environ Res Public Health 2021 Jul 7;18(14):7258. [doi: 10.3390/ijerph18147258] [Medline: 34299709]
5. Şahiner TA, Şahiner M, Cinar N, Beynex Research Group. In Covid‐19 pandemic, quarantine measures affect daily living activities in normal individuals, but do not change the cognitive state. Alzheimer’s Dementia 2021 Dec;17(S10):1-4. [doi: 10.1002/alz.055639] [Medline: 34971285]
6. Champlin C, Sirenko M, Comes T. Measuring social resilience in cities: an exploratory spatio-temporal analysis of activity routines in urban spaces during Covid-19. Cities 2023 Apr;135:104220. [doi: 10.1016/j.cities.2023.104220] [Medline: 36743889]
7. Erfani G, Bahrami B. COVID and the home: the emergence of new urban home life practised under pandemic-imposed restrictions. Cities & Health 2023 Jul 4;7(4):548-555. [doi: 10.1080/23748834.2022.2029241]
8. Breaux R, Cash AR, Lewis J, Garcia KM, Dvorsky MR, Becker SP. Impacts of COVID-19 quarantine and isolation on adolescent social functioning. Curr Opin Psychol 2023 Aug;52:101613. [doi: 10.1016/j.copsyc.2023.101613] [Medline: 37364468]
9. Hugelius K, Harada N, Marutani M. Consequences of visiting restrictions during the COVID-19 pandemic: an integrative review. Int J Nurs Stud 2021 Sep;121:104000. [doi: 10.1016/j.ijnurstu.2021.104000] [Medline: 34242976]
10. Killgore WDS, Cloonan SA, Taylor EC, Lucas DA, Dailey NS. Loneliness during the first half-year of COVID-19 lockdowns. Psychiatry Res 2020 Dec;294:113551. [doi: 10.1016/j.psychres.2020.113551] [Medline: 33171416]
11. Seppala E, Rossomando T, Doty JR. Social connection and compassion: important predictors of health and well-being. Soc Res 2013 Jun;80(2):411-430. [doi: 10.1353/sor.2013.0027]
12. Lee RM, Robbins SB. Measuring belongingness: the social connectedness and the social assurance scales. J Couns Psychol 1995;42(2):232-241. [doi: 10.1037/0022-0167.42.2.232]
13. Townsend KC, McWhirter BT. Connectedness: a review of the literature with implications for counseling, assessment, and research. J Couns Dev 2005 Apr;83(2):191-201. [doi: 10.1002/j.1556-6678.2005.tb00596.x]
14. Parsons MA. “I miss my friends, but I also miss strangers”: pandemic loneliness and the importance of place and practice. SSM Ment Health 2022 Dec;2:100121. [doi: 10.1016/j.ssmmh.2022.100121]
15. Wickramaratne PJ, Yangchen T, Lepow L, et al. Social connectedness as a determinant of mental health: a scoping review. PLoS ONE 2022;17(10):e0275004. [doi: 10.1371/journal.pone.0275004] [Medline: 36228007]
16. Holt-Lunstad J. Social connection as a critical factor for mental and physical health: evidence, trends, challenges, and future implications. World Psychiatry 2024 Oct;23(3):312-332. [doi: 10.1002/wps.21224] [Medline: 39279411]
17. Neves BB, Colón Cabrera D, Sanders A, Warren N. Pandemic diaries: lived experiences of loneliness, loss, and hope among older adults during COVID-19. Gerontologist 2023 Jan 24;63(1):120-130. [doi: 10.1093/geront/gnac104] [Medline: 35868621]
18. Lorber M, Černe Kolarič J, Kmetec S, Kegl B. Association between loneliness, well-Being, and life satisfaction before and during the COVID-19 pandemic: a cross-sectional study. Sustainability 2023;15(3):2825. [doi: 10.3390/su15032825]
19. Smith CB, Wong KLY, To-Miles F, et al. Exploring experiences of loneliness among Canadian long-term care residents during the COVID-19 pandemic: a qualitative study. Int J Older People Nurs 2023 Jan;18(1):e12509. [doi: 10.1111/opn.12509] [Medline: 36347829]
20. Kadowaki L, Wister A. Older adults and social isolation and loneliness during the COVID-19 pandemic: an integrated review of patterns, effects, and interventions. Can J Aging 2023 Jun;42(2):199-216. [doi: 10.1017/S0714980822000459] [Medline: 36345649]
21. Prommas P, Lwin KS, Chen YC, et al. The impact of social isolation from COVID-19-related public health measures on cognitive function and mental health among older adults: a systematic review and meta-analysis. Ageing Res Rev 2023 Mar;85:101839. [doi: 10.1016/j.arr.2022.101839] [Medline: 36596396]
22. The impact of COVID-19 on long-term care in canada: focus on the first 6 months. Canadian Institute for Health Information. URL: https://www.cihi.ca/en/long-term-care-and-covid-19-the-first-6-months [accessed 2025-12-01]
23. Paananen J, Rannikko J, Harju M, Pirhonen J. The impact of Covid-19-related distancing on the well-being of nursing home residents and their family members: a qualitative study. Int J Nurs Stud Adv 2021 Nov;3(May):100031. [doi: 10.1016/j.ijnsa.2021.100031] [Medline: 34095858]
24. Stall NM, Brown KA, Maltsev A, et al. COVID‐19 and Ontario’s long‐term care homes. J Elder Policy 2021 Sep;1(3):65-110. [doi: 10.18278/jep.1.3.3]

25. Gallistl V, Seifert A, Kolland F. COVID-19 as a “digital push?” research experiences from long-term care and recommendations for the post-pandemic era. Front Public Health 2021;9:660064. [doi: 10.3389/fpubh.2021.660064] [Medline: 34041216]
26. Banskota S, Healy M, Goldberg EM. 15 smartphone apps for older adults to use while in isolation during the COVID-19 pandemic. West J Emerg Med 2020 Apr 14;21(3):514-525. [doi: 10.5811/westjem.2020.4.47372] [Medline: 32302279]
27. Lemaire C, Humbert C, Sueur C, Racin C. Use of digital technologies to maintain older adults’ social ties during visitation restrictions in long-term care facilities: scoping review. JMIR Aging 2023 Feb 10;6:e38593. [doi: 10.2196/38593] [Medline: 36599164]
28. Mitchell LL, Albers EA, Birkeland RW, et al. Caring for a relative with dementia in long-term care during COVID-19. J Am Med Dir Assoc 2022 Mar;23(3):428-433. [doi: 10.1016/j.jamda.2021.11.026] [Medline: 34929196]
29. Kelly RM, Xing Y, Baker S, Waycott J. Video calls as a replacement for family visits during lockdowns in aged care: interview study with family members. JMIR Aging 2023 Jun 12;6:e40953. [doi: 10.2196/40953] [Medline: 37191951]
30. Hwang TJ, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. Int Psychogeriatr 2020 Oct;32(10):1217-1220. [doi: 10.1017/S1041610220000988] [Medline: 32450943]
31. Kuksa I, Childs M. Remediating technology, translating experience, immersing in spaces. In: Kuksa I, Childs M, editors. Making Sense of Space: Chandos Publishing; 2014:3-22. [doi: 10.1533/9781780634067.1.3]
32. Khosravi P, Rezvani A, Wiewiora A. The impact of technology on older adults’ social isolation. Comput Human Behav 2016 Oct;63:594-603. [doi: 10.1016/j.chb.2016.05.092]
33. Garnett A, Pollock H, Floriancic N, et al. Social connectedness between family caregivers and older adults living in long-term care homes in the context of COVID-19. Can J Aging 2024 Mar;43(1):33-44. [doi: 10.1017/S0714980823000351] [Medline: 37727879]
34. Hart JL, Turnbull AE, Oppenheim IM, Courtright KR. Family-centered care during the COVID-19 era. J Pain Symptom Manage 2020 Aug;60(2):e93-e97. [doi: 10.1016/j.jpainsymman.2020.04.017] [Medline: 32333961]
35. Hardy MS, Fanaki C, Savoie C, et al. Acceptability of videoconferencing to preserve the contact between cognitively impaired long-term care residents and their family caregivers: a mixed-methods study. Geriatr Nurs 2022;48:65-73. [doi: 10.1016/j.gerinurse.2022.09.006] [Medline: 36155311]
36. Naudé B, Rigaud AS, Pino M. Video calls for older adults: a narrative review of experiments involving older adults in elderly care institutions. Front Public Health 2021;9:751150. [doi: 10.3389/fpubh.2021.751150] [Medline: 35096731]
37. Tan AJQ, Rusli KDB, McKenna L, Tan LLC, Liaw SY. Telemedicine experiences and perspectives of healthcare providers in long-term care: a scoping review. J Telemed Telecare 2024 Feb;30(2):230-249. [doi: 10.1177/1357633X211049206] [Medline: 34666535]
38. Sixsmith A, Horst BR, Simeonov D, Mihailidis A. Older people’s use of digital technology during the COVID-19 pandemic. Bull Sci Technol Soc 2022 Jun;42(1-2):19-24. [doi: 10.1177/02704676221094731] [Medline: 38603230]
39. Ang S, Lim E, Malhotra R. Health-related difficulty in internet use among older adults: correlates and mediation of its association with quality of life through social support networks. Gerontologist 2021 Jul 13;61(5):693-702. [doi: 10.1093/geront/gnaa096] [Medline: 32744310]
40. Hacker J, vom Brocke J, Handali J, Otto M, Schneider J. Virtually in this together—how web-conferencing systems enabled a new virtual togetherness during the COVID-19 crisis. Eur J Inf Syst 2020 Sep 2;29(5):563-584. [doi: 10.1080/0960085X.2020.1814680]
41. Hung L, Chow B, Shadarevian J, et al. Using touchscreen tablets to support social connections and reduce responsive behaviours among people with dementia in care settings: a scoping review. Dementia (London) 2021 Apr;20(3):1124-1143. [doi: 10.1177/1471301220922745] [Medline: 32380856]
42. Vaportzis E, Clausen MG, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. Front Psychol 2017 Oct 4;8:1687. [doi: 10.3389/fpsyg.2017.01687] [Medline: 29071004]
43. Gell NM, Rosenberg DE, Demiris G, LaCroix AZ, Patel KV. Patterns of technology use among older adults with and without disabilities. Gerontologist 2015 Jun;55(3):412-421. [doi: 10.1093/geront/gnt166] [Medline: 24379019]
44. Garnett A, Yurkiv H, Booth R, Connelly D, Donelle L. Web-based presence for social connectedness in long-term care: protocol for a qualitative multimethods study. JMIR Res Protoc 2023 Oct 27;12(1):e50137. [doi: 10.2196/50137] [Medline: 37889518]
45. Sandelowski M. Whatever happened to qualitative description? Res Nurs Heal 2000;23(4):334-340. [doi: 10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G] [Medline: 10940958]
46. Sandelowski M. What’s in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
47. Phillippi J, Lauderdale J. A guide to field notes for qualitative research: context and conversation. Qual Health Res 2018 Feb;28(3):381-388. [doi: 10.1177/1049732317697102] [Medline: 29298584]
48. Colorafi KJ, Evans B. Qualitative descriptive methods in health science research. HERD 2016 Jul;9(4):16-25. [doi: 10.1177/1937586715614171] [Medline: 26791375]
49. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]

50. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A systematic review of the technology acceptance model in health informatics. Appl Clin Inform 2018 Jul;9(3):604-634. [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
51. Kim H, Sefcik JS, Bradway C. Characteristics of qualitative descriptive studies: a systematic review. Res Nurs Health 2017 Feb;40(1):23-42. [doi: 10.1002/nur.21768] [Medline: 27686751]
52. Doyle L, McCabe C, Keogh B, Brady A, McCann M. An overview of the qualitative descriptive design within nursing research. J Res Nurs 2020 Aug;25(5):443-455. [doi: 10.1177/1744987119880234] [Medline: 34394658]
53. Rourke S. How does virtual reality simulation compare to simulated practice in the acquisition of clinical psychomotor skills for pre-registration student nurses? A systematic review. Int J Nurs Stud 2020 Feb;102:103466. [doi: 10.1016/j.ijnurstu.2019.103466] [Medline: 31783192]
54. Perdomo Delgato C, Ripat J, Mallory-Hill S, Bohunicky S. Using the technology acceptance model to explore the user experience of smart kitchen use among older adult women. In: Marcus A, Rosenzweig E, Soares MM, editors. Design, User Experience, and Usability: 13th International Conference, DUXU 2024, Held as Part of the 26th HCI International Conference, HCII 2024, Washington, DC, USA, June 29 – July 4, 2024, Proceedings, Part III: Springer:3-12. [doi: 10.1007/978-3-031-61356-2_1]
55. Bradshaw C, Atkinson S, Doody O. Employing a qualitative description approach in health care research. Glob Qual Nurs Res 2017;4:2333393617742282. [doi: 10.1177/2333393617742282] [Medline: 29204457]
56. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
57. NVivo 13. Lumivero. 2023. URL: https://lumivero.com/products/nvivo/ [accessed 2023-06-01]
58. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
59. Assarroudi A, Heshmati Nabavi F, Armat MR, Ebadi A, Vaismoradi M. Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. J Res Nurs 2018 Feb;23(1):42-55. [doi: 10.1177/1744987117741667] [Medline: 34394406]
60. Johnson JL, Adkins D, Chauvin S. A review of the quality indicators of rigor in qualitative research. Am J Pharm Educ 2020 Jan;84(1):7120. [doi: 10.5688/ajpe7120] [Medline: 32292186]
61. Creswell JW, Miller DL. Determining validity in qualitative inquiry. Theory Pract 2000 Aug;39(3):124-130. [doi: 10.1207/s15430421tip3903_2]
62. Soubutts E, Ayobi A, Eardley R, McNaney R, Cater K, O’Kane AA. Amazon Echo Show as a multimodal human-to-human care support tool within self-isolating older UK households. Proc ACM Hum-Comput Interact 2022 Nov 7;6(CSCW2):1-31. [doi: 10.1145/3555193]
63. Zamir S, Hennessy C, Taylor A, Jones R. Intergroup “Skype” quiz sessions in care homes to reduce loneliness and social isolation in older people. Geriatrics (Basel) 2020 Nov 11;5(4):90. [doi: 10.3390/geriatrics5040090] [Medline: 33187242]
64. Moyle W, Jones C, Murfield J, Liu F. ‘For me at 90, it’s going to be difficult’: feasibility of using iPad video-conferencing with older adults in long-term aged care. Aging Ment Health 2020 Feb 1;24(2):349-352. [doi: 10.1080/13607863.2018.1525605] [Medline: 30621436]
65. Tsai HH, Cheng CY, Shieh WY. Effectiveness of laptop-based versus smartphone-based videoconferencing interaction on loneliness, depression and social support in nursing home residents: a secondary data analysis. J Telemed Telecare 2023 Apr;29(3):177-186. [doi: 10.1177/1357633X20972004] [Medline: 33197365]
66. Astell A, Dosanjh S, D’Elia T, et al. Personalized tablets for residents in long-term care to support recreation and mitigate isolation. J Am Med Dir Assoc 2024 Jul;25(7):105022. [doi: 10.1016/j.jamda.2024.105022] [Medline: 38763162]
67. Hardy MS, Fanaki C, Savoie C, et al. Long-term care staffs’ experience in facilitating the use of videoconferencing by cognitively impaired long-term care residents during the COVID-19 pandemic: a mixed-methods study. BMC Health Serv Res 2024 May 21;24(1):646. [doi: 10.1186/s12913-024-11095-9] [Medline: 38769512]
68. Schmidt SB, Isbel S, John B, Subramanian R, D’Cunha NM. Examining technology perspectives of older adults with mild cognitive impairment: scoping review. JMIR Aging 2025 Oct 30;8(1):e78229. [doi: 10.2196/78229] [Medline: 41166548]
69. Nguyen M, Fujioka J, Wentlandt K, et al. Using the technology acceptance model to explore health provider and administrator perceptions of the usefulness and ease of using technology in palliative care. BMC Palliat Care 2020 Sep 7;19(1):138. [doi: 10.1186/s12904-020-00644-8] [Medline: 32895060]
70. Liougas MP, Sommerlad A, O’Rourke HM, et al. Assessing social connection for long‐term care home residents: a scoping review of measure content. Alzheimers Dement (N Y) 2024;10(3):e12488. [doi: 10.1002/trc2.12488] [Medline: 39315314]
71. Ratnayake M, Brathwaite S, Lukas S, et al. Aging in place: are we prepared? Dela J Public Health 2022 Aug;8(3):28-31. [doi: 10.32481/djph.2022.08.007] [Medline: 36177171]
72. De Guzman KR, Snoswell CL, Caffery LJ, Smith AC. Economic evaluations of videoconference and telephone consultations in primary care: a systematic review. J Telemed Telecare 2024 Jan;30(1):3-17. [doi: 10.1177/1357633X211043380] [Medline: 34617819]
73. Grant KL, Lee DD, Cheng I, Baker GR. Reducing preventable patient transfers from long-term care facilities to emergency departments: a scoping review. CJEM 2020 Nov;22(6):844-856. [doi: 10.1017/cem.2020.416] [Medline: 32741417]

74. Ng R, Lane N, Tanuseputro P, et al. Increasing complexity of new nursing home residents in Ontario, Canada: a serial cross-sectional study. J Am Geriatr Soc 2020 Jun;68(6):1293-1300. [doi: 10.1111/jgs.16394] [Medline: 32119121]
75. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet Standing Commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
76. Zapletal A, Wells T, Russell E, Skinner MW. On the triple exclusion of older adults during COVID-19: technology, digital literacy and social isolation. Soc Sci Humanit Open 2023;8(1):100511. [doi: 10.1016/j.ssaho.2023.100511] [Medline: 37021073]
77. Reneland-Forsman L. 'Borrowed access'—the struggle of older persons for digital participation. In J Lifelong Educ 2018 May 4;37(3):333-344. [doi: 10.1080/02601370.2018.1473516]
78. Kelly RM, Xing Y, Baker S, Waycott J. "More than just holding the iPad": family Members' perspectives on the work of video calling in aged care. Proc ACM Hum-Comput Interact 2024 Nov 7;8(CSCW2):1-26. [doi: 10.1145/3687007]
79. Chu CH, Yee A, Stamatopoulos V. Poor and lost connections: essential family caregivers' experiences using technology with family living in long-term care homes during COVID-19. J Appl Gerontol 2022 Jun;41(6):1547-1556. [doi: 10.1177/07334648221081850] [Medline: 35416076]
80. Long-term care and COVID-19: report of a special task force prepared for the Chief Science Advisor of Canada. : Government of Canada; 2020 URL: https://science.gc.ca/site/science/en/office-chief-science-advisor/initiatives-covid-19/long-term-care-and-covid-19 [accessed 2020-10-29]
81. Newbould L, Ariss S, Mountain G, Hawley MS. Exploring factors that affect the uptake and sustainability of videoconferencing for healthcare provision for older adults in care homes: a realist evaluation. BMC Med Inform Decis Mak 2021 Jan 6;21(1):13. [doi: 10.1186/s12911-020-01372-y] [Medline: 33407437]
82. Long-term care staffing study. : Ontario Ministry of Long-Term Care; 2020 URL: https://files.ontario.ca/mltc-long-term-care-staffing-study-en-2020-07-31.pdf [accessed 2025-11-15]
83. Khanam F, Langevin M, Savage K, Uppal S. Women working in paid care occupations. : Statistics Canada; 2022 URL: https://www150.statcan.gc.ca/n1/pub/75-006-x/2022001/article/00001-eng.pdf [accessed 2025-11-15]
84. Dementia in long-term care. Canadian Institute for Health Information. 2025. URL: https://www.cihi.ca/en/dementia-in-canada/dementia-care-across-the-health-system/dementia-in-long-term-care [accessed 2025-11-15]

Abbreviations

FM: family member
LTCH: long-term care home
TAM: technology acceptance model
WVT: web-based videoconferencing technology

Edited by Y Luo; submitted 27.Feb.2025; peer-reviewed by CK Tan, R Kelly; revised version received 04.Nov.2025; accepted 19.Nov.2025; published 09.Jan.2026.

Please cite as:
Garnett A, Yurkiv H, Connelly DM, Booth R, Donelle L
Perceived Use of Web-Based Videoconferencing for Social Connectedness Among Older Adults Living in Long-Term Care: Qualitative Study
JMIR Aging 2026;9:e73213
URL: https://aging.jmir.org/2026/1/e73213
doi:10.2196/73213

Predictive Model of Acupuncture Adherence in Alzheimer Disease: Secondary Analysis of Randomized Controlled Trials

Ze-Hao Chen[1*], MM; Ran Li[2*], MD; Yu-Hang Jiang[1*], MM; Jia-Kai He[3], MD; Shan-Shan Yan[1], MM; Guan-Hua Zong[1], BM; Zong-Xi Yi[4], BM; Xin-Yu Ren[4], BM; Bao-Hui Jia[2*], MD, PhD

[1]Department of Acupuncture and Moxibustion, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, 5 Beixiange, Xicheng District, Beijing, China

[2]Department of Rehabilitation Medicine, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, 5 Beixiange, Xicheng District, Beijing, China

[3]Department of Traditional Chinese Medicine, Peking University People's Hospital, Beijing, China

[4]Department of Acupuncture and Moxibustion, Guang'anmen Hospital, Beijing University of Chinese Medicine, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Bao-Hui Jia, MD, PhD
Department of Rehabilitation Medicine, Guang'anmen Hospital, China Academy of Chinese Medical Sciences, 5 Beixiange, Xicheng District, Beijing, China

Abstract

Background: The therapeutic efficacy of acupuncture in treating Alzheimer disease (AD) largely depends on consistent treatment adherence. Therefore, identifying key factors influencing adherence and developing targeted interventions are crucial for enhancing clinical outcomes.

Objective: This study aims to develop and validate a predictive model for identifying patients with AD who are likely to maintain good adherence to acupuncture treatment.

Methods: This secondary analysis included 108 patients with probable AD, aged 50 to 85 years, from 2 independent randomized controlled trials conducted at Guang'anmen Hospital, China Academy of Chinese Medical Sciences. Of all, 66 patients were assigned to the development cohort and 42 to the external validation cohort. Acupuncture adherence was defined as the proportion of completed sessions relative to scheduled sessions, with good adherence defined as ≥80% completion. Baseline data included demographic, clinical, cognitive, functional, psychological, and caregiving variables. Multivariable logistic regression with backward stepwise selection was used to identify significant predictors, and a nomogram was constructed based on the final model. Model performance was assessed using receiver operating characteristic curves, calibration plots, and decision curve analysis, with external validation performed by receiver operating characteristic analysis. Sensitivity analysis was performed using alternative adherence thresholds of 70% and 90%.

Results: A higher number of treatments during the first month was associated with a significant increase in the odds of good adherence (odds ratio [OR] 3.06, 95% CI 1.68 - 7.01; *P*=.002), while longer disease duration (OR 0.97, 95% CI 0.94 - 1.00; *P*=.049) and receiving care from a part-time caregiver (OR 0.19, 95% CI 0.04 - 0.72; *P*=.022) were associated with lower odds of adherence. Sensitivity analyses further supported the stability and reliability of the model.

Conclusions: This study is the first to develop and validate a predictive model for acupuncture adherence in patients with AD. In clinical research, it can facilitate participant stratification and help identify individuals who may need additional adherence support, thereby reducing bias and enhancing trial quality. In clinical practice, the nomogram enables proactive adherence management by prospectively identifying high-risk patients and guiding targeted strategies to improve adherence and optimize therapeutic outcomes.

(JMIR Aging 2026;9:e82787) doi:10.2196/82787

KEYWORDS

alzheimer disease; acupuncture; adherence; nomogram; predictive model

Introduction

Alzheimer disease (AD) is a progressive neurodegenerative disorder primarily characterized by cognitive decline, functional impairment in activities of daily living, and neuropsychiatric symptoms [1]. Acupuncture, given its favorable safety profile and potential for symptomatic improvement [2-4], has been recommended as a promising nonpharmacological long-term therapy in clinical practice guidelines for the management of AD [5]. However, the therapeutic efficacy of acupuncture for AD remains a subject of ongoing debate [6,7]. This inconsistency in findings may be partially explained by a critical, yet often underinvestigated factor: treatment adherence.

Consistent adherence is crucial to therapeutic efficacy, particularly in chronic disease management and nonpharmacological interventions. At present, adherence has been studied primarily in the context of pharmacotherapy, with reported rates of long-term pharmacological adherence in patients with AD ranging from 16.5% to 51% [8,9]. Some studies have developed models to predict patients' medication adherence [10,11]. In contrast, adherence to nonpharmacological therapies such as acupuncture remains insufficiently explored. This knowledge gap is compounded by the lack of a standardized definition for what constitutes good or poor adherence. This variability in definitions across studies makes it difficult to compare findings and may be a key confounding factor obscuring the true dose-response relationship of acupuncture [12].

Clinical studies have shown that factors influencing acupuncture adherence are multifaceted, encompassing patient subjective norms, perceived behavioral control, and treatment commitment [13]. Specific factors include perceived effectiveness, family support, emotional status, patient recognition and acceptance of acupuncture, and the availability of medical subsidies [14]. Nonetheless, most of these studies have provided only descriptive insights or general recommendations based on literature reviews [15,16], while few have offered quantitative evidence derived from clinical data. Notably, research specifically focusing on acupuncture adherence in individuals with AD is exceedingly scarce. Given the substantial cognitive decline, behavioral symptoms, and caregiver dependence associated with AD, patients in this population may face unique barriers to maintaining adherence. Thus, findings from studies in other populations may not be directly generalizable to patients with AD. Understanding the specific determinants of adherence in this unique population is crucial for optimizing treatment delivery and improving outcomes.

Given the ongoing debate over acupuncture's efficacy and the critical yet poorly understood role of adherence in treatment outcomes, identifying predictors for treatment engagement in a methodologically robust setting is paramount. In this study, we conducted a secondary analysis of data from 2 independent randomized controlled trials (RCTs) previously conducted by our team, with the aim of identifying key predictors of acupuncture adherence among patients with AD and developing a clinically applicable predictive model. By identifying individuals at high risk of poor adherence and enabling the development of targeted intervention strategies, this study seeks to enhance the adherence and efficacy of acupuncture treatment and provide evidence-based support for adherence management in nonpharmacological interventions for AD.

Methods

Study Design and Data Source

This study was a secondary analysis based on data collected from 2 independent RCTs conducted at Guang'anmen Hospital, China Academy of Chinese Medical Sciences (CACMS). Participants were primarily recruited between December 2021 and June 2024. The first RCT, conducted from December 2021 to June 2024, enrolled 66 patients and served as the model development cohort. The second RCT, conducted from June 2022 to November 2022, included 42 patients and served as an external validation cohort for the predictive model.

Eligibility screening and clinical assessments were conducted by licensed physicians from the departments of encephalopathy and neurology. Inclusion criteria and exclusion criteria are present in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

- **Inclusion criteria:**
  - A diagnosis of probable Alzheimer disease according to the National Institute on Aging–Alzheimer's Association criteria [17]
  - Age between 50 and 85 years
  - A Clinical Dementia Rating score ≥0.5
  - A Mini-Mental State Examination score ≤26
  - A Hachinski Ischemic Scale score ≤4
- **Exclusion criteria:**
  - Other neurological or systemic disorders known to cause progressive cognitive impairment
  - Recent use of medications or exposure to substances known to impair cognition
  - A history of trypanophobia or active skin infections
  - Acupuncture or electroacupuncture treatment within the past 2 weeks
  - Participation in other clinical trials

Ethical Considerations

Both studies were conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. This study is a secondary analysis of data derived from 2 RCTs previously conducted at Guang'anmen Hospital, CACMS. Both original trials received independent ethical approval from the ethics committee of Guang'anmen Hospital, CACMS (approval: 2021-056-KY-01 and 2022-087-KY). Written informed consent was obtained from all participants or their legally authorized representatives prior to data collection. Participants did not receive compensation for participation in the original trials. All treatments and assessments were provided free of charge. To protect participant privacy, all data used in this secondary analysis were anonymized and deidentified prior to analysis.

Treatment and Adherence Assessment

Participants received 20-minute acupuncture sessions 3 times per week (on nonconsecutive days) for a total of 12 weeks. Adherence was assessed by calculating the proportion of completed treatment sessions relative to the total number of scheduled sessions during the intervention period. The proportion of days covered (PDC) was used as the adherence metric. In the absence of an established gold standard for adherence in nonpharmacological trials like acupuncture, we adopted the widely accepted threshold of PDC ≥80% from pharmacotherapy research [9]. This threshold is a well-validated proxy for sufficient exposure to treatment in chronic disease management. Participants with a PDC ≥80% were classified as having good adherence, while those with a PDC <80% were categorized as having poor adherence.

Data Collection

The following clinical data were collected: sex, age, disease duration, disease severity, educational level, occupation, history of acupuncture treatment, Mini-Mental State Examination score, Alzheimer's Disease Assessment Scale–Cognitive Subscale score, basic activities of daily living, instrumental activities of daily living, presence of behavioral and psychological symptoms of dementia, Patient Health Questionnaire-9 score for depressive symptoms, caregiving role, number of treatments in the first month, travel time to the hospital, recruitment method, and adherence outcomes.

Screening of Influencing Factors

A linear regression model was used to assess multicollinearity among independent variables. Multicollinearity was quantified by calculating the variance inflation factor and tolerance values. Variables with a variance inflation factor <10 and tolerance >0.1 were considered to have acceptable levels of multicollinearity [18]. Variables meeting these criteria were subsequently included in a multivariable logistic regression model. Variable selection was performed using a backward stepwise regression approach based on the likelihood ratio test, with a significance threshold of *P*<.05 for retention in the model. This data-driven approach was chosen to build a parsimonious model and reduce the risk of overfitting. We also conducted exploratory analyses by forcing clinically relevant but nonsignificant variables into the model, but this did not lead to a significant improvement in model performance and increased model complexity. Therefore, the final model retained only the statistically significant predictors. Final variables were required to meet criteria for statistical significance, low multicollinearity, and satisfactory predictive performance. To improve model interpretability, Shapley Additive Explanations (SHAP) values were calculated to quantify the relative contribution of each predictor to the outcome [19]. SHAP values were visualized using beeswarm plots.

Construction and Evaluation of the Nomogram

A nomogram was developed based on the final multivariable logistic regression model to predict adherence. The nomogram serves as a graphical tool to visualize the relationships between multiple predictors and the outcome, facilitating individualized risk assessment and clinical decision-making. In the nomogram, each predictor is aligned with its corresponding axis; by drawing a vertical line from the predictor's value to the point scale, a score can be assigned. The total score, obtained by summing the individual scores, corresponds to a predicted probability of adherence on the nomogram's outcome axis.

Model performance was assessed through receiver operating characteristic (ROC) curve analysis, calibration plots, the Hosmer-Lemeshow goodness-of-fit test, and decision curve analysis using the development dataset. Internal validation was conducted using the bootstrap method with 1000 resampling iterations. Agreement between predicted and observed outcomes was evaluated using the κ statistic. An area under the ROC curve (AUC) between 0.5 and 0.7 was interpreted as indicating low discrimination, 0.7 to 0.9 as moderate, and >0.9 as high discrimination [20]. The nomogram's robustness was further evaluated by performing ROC analysis in the external validation cohort.

Statistical Methods

Data completeness for the variables included in the final model was assessed prior to analysis. There were no missing values for the variables included in the final model in either the development or validation cohorts, as complete data collection was a requirement for the per-protocol analysis in the 2 RCTs. All statistical analyses were conducted using R software (version 4.4.2; R Foundation for Statistical Computing). Continuous variables with a normal distribution were presented as mean (SD), while nonnormally distributed variables were expressed as median (IQR), with distribution assessed using the Shapiro-Wilk test. Categorical variables were reported as frequencies and percentages. Group comparisons were performed using independent-samples *t* tests for normally distributed continuous variables and the Mann-Whitney *U* test for nonnormally distributed continuous variables. Categorical variables were compared using the chi-square test or the Fisher exact test, as appropriate. All statistical tests were 2-tailed, and a *P* value <.05 was considered statistically significant. To assess the robustness of our model to the primary adherence definition (PDC ≥80%), a sensitivity analysis was performed by repeating the multivariable logistic regression using alternative thresholds of 70% and 90%.

Results

Baseline Characteristics of the Training Set

The predictive model was developed using data from the development cohort (cohort 1), which comprised 66 patients with AD (Table 1). Participants were stratified by treatment adherence into a good adherence group (n=43) and a poor adherence group (n=23). Of the total participants, 34 (51.5%) were female participants and 32 (48.5%) were male participants. The mean age was 71.8 (SD 7.9) years, and the mean disease duration was 50.0 (SD 26.0) months. Univariate analysis revealed significant differences between the good and poor adherence groups in caregiving status (*P*=.025) and the number of treatment sessions during the first month (*P*=.001).

The *P* values for testing differences between patients with good and poor adherence to acupuncture treatment were derived from independent samples *t* tests, Mann-Whitney *U* tests and chi-square tests or the Fisher exact tests.

Table . Comparison of baseline characteristics between patients with good and poor adherence (n=66).

Variable	Total	Good adherence (n=43)	Poor adherence (n=23)	*t* test (*df*)[a]	Wilcoxon rank-sum test[a]	Chi-square (*df*)	*P* value
Sex, n (%)				—	—[b]	1.88 (1)	.171
Female	34 (52)	19 (44)	15 (65)				
Male	32 (48)	24 (56)	8 (35)				
Age (y), mean (SD)	71.8 (7.9)	70.9 (7.7)	73.5 (8.3)	−1.28 (64)	—	—	.204
Disease duration (mon), mean (SD)	50.0 (26.0)	48.1 (26.0)	53.5 (26.2)	−0.80 (64)	—	—	.427
Disease severity,[c] n (%)				—	—	—	.490
Mild	29 (44)	21 (49)	8 (35)				
Moderate	28 (42)	16 (37)	12 (52)				
Severe	9 (14)	6 (14)	3 (13)				
Education level, n (%)				—	—	0.05 (1)	.816
No higher education	40 (61)	27 (63)	13 (57)				
Higher education	26 (39)	16 (37)	10 (43)				
Occupation, n (%)				—	—	0.00 (1)	.99
Manual work	21 (32)	14 (33)	7 (30)				
Nonmanual work	45 (68)	29 (67)	16 (70)				
MMSE[d], median (IQR)	16.5 (11.0 - 21.0)	18.0 (12.0 - 21.5)	16.0 (9.5 - 20.0)	—	582	—	.241
ADAS-Cog[e], median (IQR)	22.0 (15.0 - 38.8)	22.0 (14.5 - 39.0)	24.0 (16.0 - 38.5)	—	454	—	.590
BADL[f], median (IQR)	10.0 (9.0 - 13.0)	10.0 (9.0 - 12.5)	11.0 (9.5 - 13.5)	—	444.5	—	.500
IADL[g], median (IQR)	26.0 (18.2 - 33.0)	27.0 (16.0 - 32.5)	25.0 (19.5 - 33.5)	—	472.5	—	.772
BPSD[h],[c] n (%)				—	—	—	.99
Present	53 (80)	34 (79)	19 (83)				
Absent	13 (20)	9 (21)	4 (17)				
PHQ-9[i],[c] n (%)				—	—	—	.99
Depressive symptoms	3 (5)	2 (5)	1 (4)				
Normal	63 (95)	41 (95)	22 (96)				
Travel time to the hospital (min), median (IQR)	60.0 (40.0 - 100.0)	50.0 (35.0 - 95.0)	60.0 (50.0 - 135.0)	—	413	—	.272
First-month treatment sessions, median (IQR)	11.0 (11.0 - 12.0)	12.0 (11.0 - 12.0)	10.0 (8.0 - 12.0)	—	736	—	.001
Caregiving role, n (%)				—	—	5.05 (1)	.025

Variable	Total	Good adherence (n=43)	Poor adherence (n=23)	*t* test (*df*)[a]	Wilcoxon rank-sum test[a]	Chi-square (*df*)	*P* value
Full-time caregiver	34 (52)	27 (63)	7 (30)				
Part-time caregiver	32 (48)	16 (37)	16 (70)				
History of acupuncture, n (%)				—	—	0.00 (1)	.99
No	30 (45)	20 (47)	10 (43)				
Yes	36 (55)	23 (53)	13 (57)				
Recruitment method[c], n (%)				—	—	—	.844
Nursing home	4 (6)	3 (7)	1 (4)				
Multimedia	30 (45)	19 (44)	11 (48)				
Study team referral	15 (23)	11 (26)	4 (18)				
Outpatient clinic	17 (26)	10 (23)	7 (30)				

[a]Continuous variables were compared using independent-samples *t* tests when normally distributed and Wilcoxon rank-sum tests otherwise. Categorical variables were compared using the chi-square test or Fisher exact test, as appropriate.

[b]—: not applicable.

[c]Evaluated using the Fisher exact test.

[d]MMSE: Mini-Mental State Examination.

[e]ADAS-Cog: Alzheimer's Disease Assessment Scale–Cognitive Subscale.

[f]BADL: basic activities of daily living.

[g]IADL: instrumental activities of daily living.

[h]BPSD: behavioral and psychological symptoms of dementia.

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

Independent Predictors of Adherence to Acupuncture Treatment Among Patients with AD

Initially, 17 potential predictors of adherence to acupuncture treatment in patients with AD were considered. Following multicollinearity analysis and assessment of clinical relevance, 16 variables were ultimately selected for model construction (Table S1 in Multimedia Appendix 1). Multivariable logistic regression analysis was performed to develop a predictive model of adherence to acupuncture treatment in this population. The final model is represented by the following equation (Figure 1; Table S2 in Multimedia Appendix 1):

$$\log(p/(1-p)) = -2.44 - 0.03 \times DiseaseDuration + 1.12 \times NumberofTreatmentsFirstMonth - 1.68 \times CaregivingRole$$

Disease duration, the number of treatments in the first month, and caregiving role were independent predictors of adherence to acupuncture treatment among patients with AD. A higher number of treatments during the first month was associated with a significant increase in the odds of good adherence (odds ratio [OR] 3.06, 95% CI 1.68 - 7.01; *P*=.002), while longer disease duration (OR 0.97, 95% CI 0.94 - 1.00; *P*=.049) and receiving care from a part-time caregiver (OR 0.19, 95% CI 0.04 - 0.72; *P*=.022) were associated with lower odds of adherence. SHAP analysis quantified the contributions of these predictors, confirming their importance in the model (Figure 2). The SHAP results were consistent with the logistic regression findings, enhancing the interpretability of the model.

The results indicated that the number of treatment sessions during the first month, caregiving role, and disease duration were the 3 most important factors influencing adherence. This ranking was highly consistent with the findings of the logistic regression model, demonstrating the substantial contribution of these variables to the model's predictive performance.

The SHAP summary plot visualizes the impact of each variable on the predicted probability of adherence. Each dot represents the SHAP value of an individual observation, indicating the degree and direction of that variable's influence on the model output. Higher SHAP values correspond to a stronger positive contribution to adherence probability. The color gradient from blue to red indicates the relative value of each variable for that observation (blue =lower value; red =higher value). For example, a higher number of treatments in the first month (red) is associated with higher SHAP values, reflecting its positive impact on adherence.

Figure 1. Forest plot of factors influencing adherence to acupuncture treatment among patients with Alzheimer disease based on multivariable logistic regression analysis. Dots represent odds ratios, and horizontal lines indicate 95% CIs. Longer disease duration and care provided by part-time caregivers were associated with lower adherence (odds ratio<1), while a higher number of treatment sessions during the first month significantly increased adherence (odds ratio >1). All variables included in the model were statistically significant (*P*<.05).

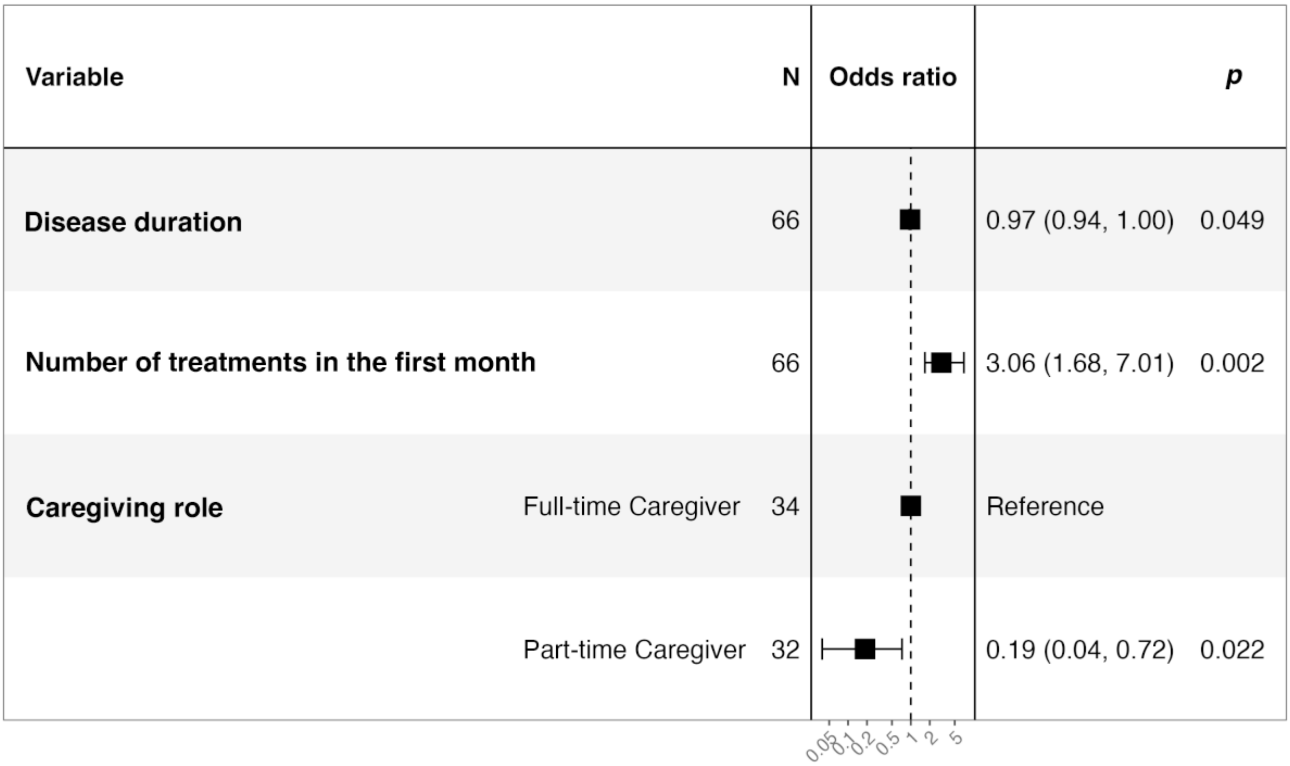
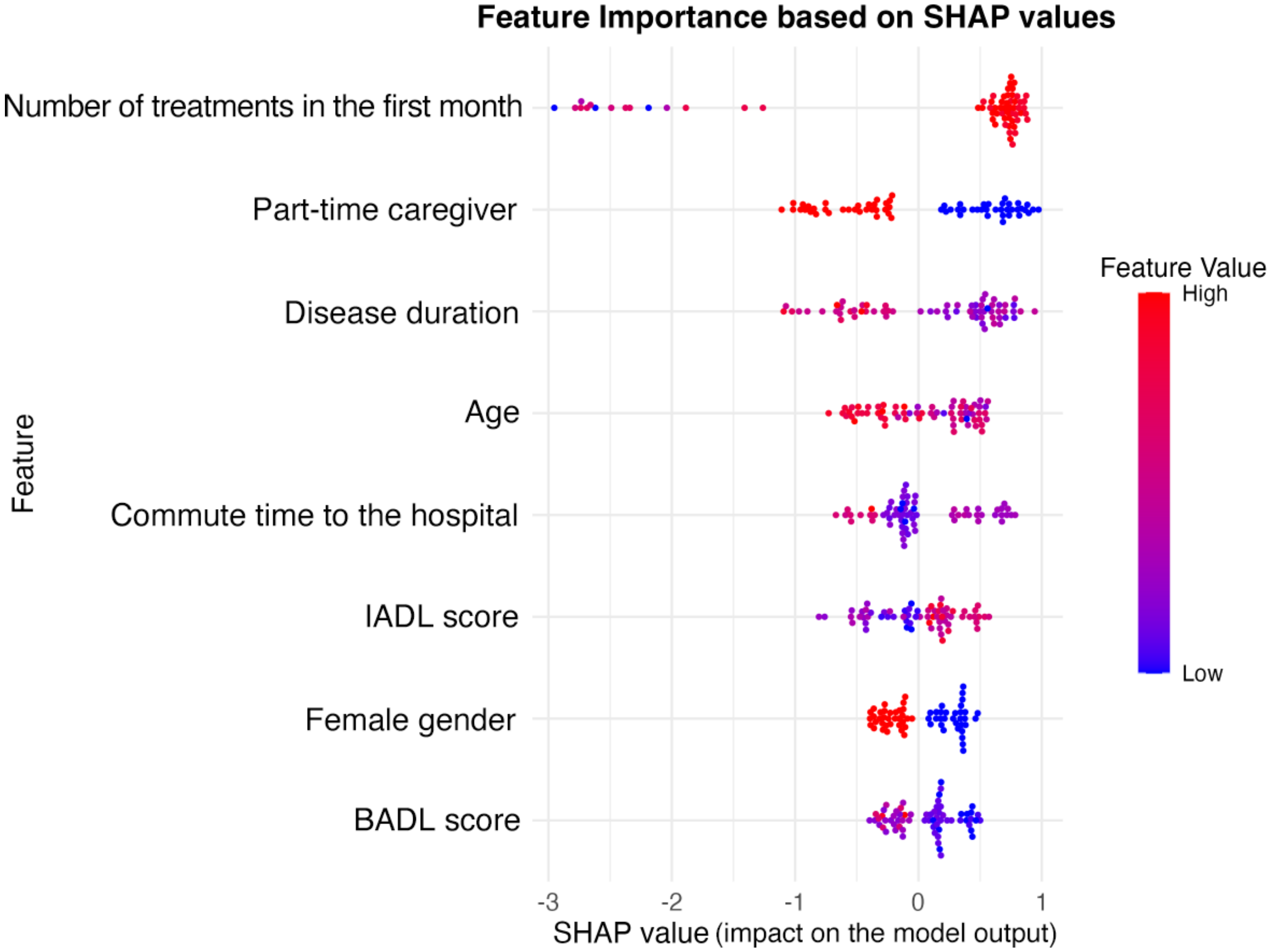


Figure 2. Shapely Additive Explanations (SHAP) summary plot of variable importance for predicting adherence to acupuncture treatment in patients with Alzheimer disease based on multivariable logistic regression analysis. BADL: basic activities of daily living; IADL: instrumental activities of daily living.



Development and Evaluation of the Nomogram

A nomogram was developed based on the final predictive model, incorporating the number of treatments in the first month, disease duration, and caregiving role (Figure 3). The model demonstrated excellent discrimination, with an area under the ROC curve (AUC) of 0.914 (Figure 4A). Validation was performed using cohort 2 (test set), which included 42 patients with AD (Table S3 in Multimedia Appendix 1). The calibration curve from internal validation indicated good agreement between predicted and observed probabilities (mean absolute error =0.04; mean squared error =0.003; 90th quantile absolute error =0.078) (Figure 4B). The Hosmer-Lemeshow goodness-of-fit test showed that the model predictions were well calibrated (χ^2_8=10.9; *P*=0.21). The decision curve analysis showed that the nomogram provided a net clinical benefit across a wide range of threshold probabilities (Figure 4D). Internal validation yielded an overall predictive accuracy of 89.4%, with a κ statistic of 0.759, indicating good consistency and potential clinical utility. Furthermore, the model performed well in the external validation cohort, achieving an AUC of 0.833 (Figure 4C).

For example, a patient with AD with a disease duration of 60 months, part-time caregiving, and 11 acupuncture sessions during the first month would obtain corresponding point values on each variable axis, indicated by red dots on the upper horizontal scales. The points for the 3 variables are summed to yield a total score of approximately 234. On the bottom axis, a total score of 234 corresponds to a predicted adherence probability of approximately 0.30, as marked by the red arrow. This relatively low predicted probability highlights the need for proactive adherence management and targeted interventions in clinical practice.

Figure 3. Nomogram for predicting adherence to acupuncture treatment among patients with Alzheimer disease.

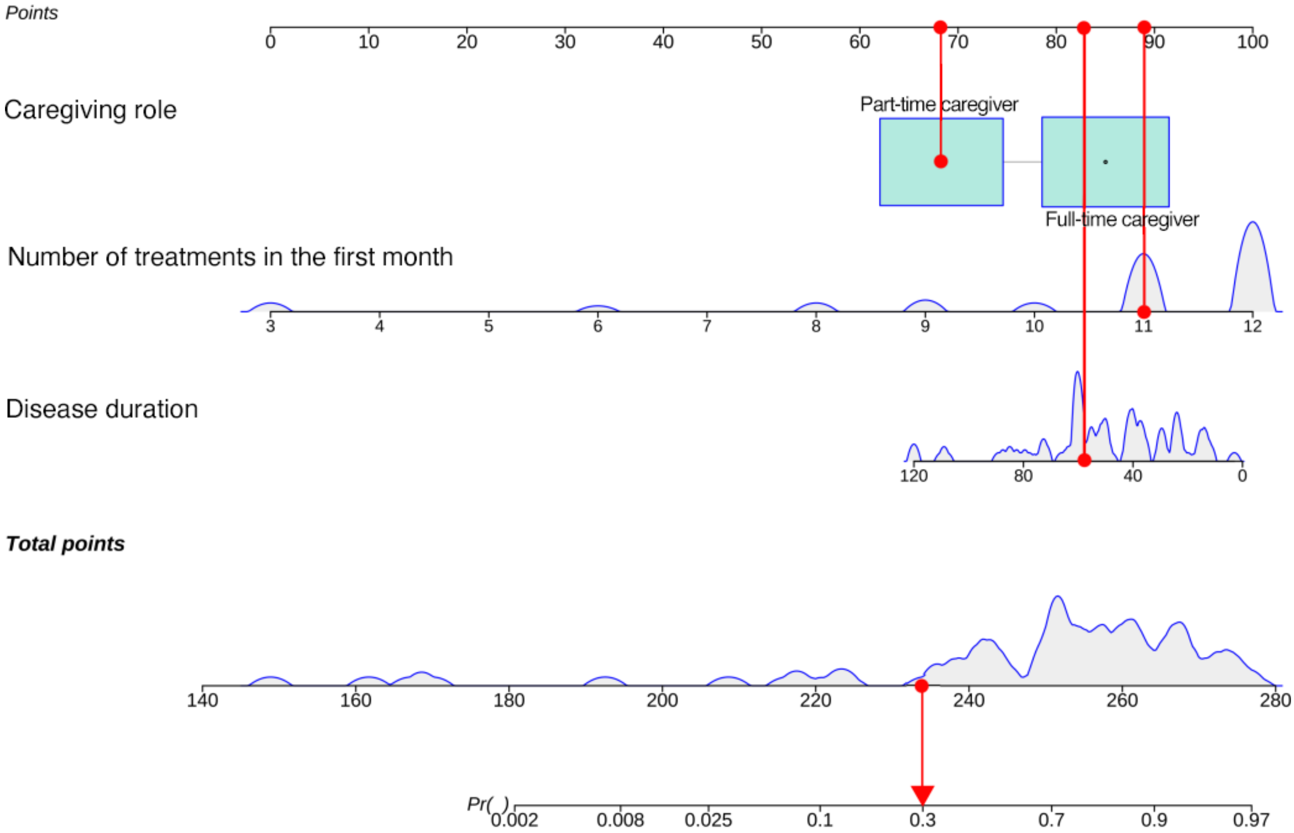
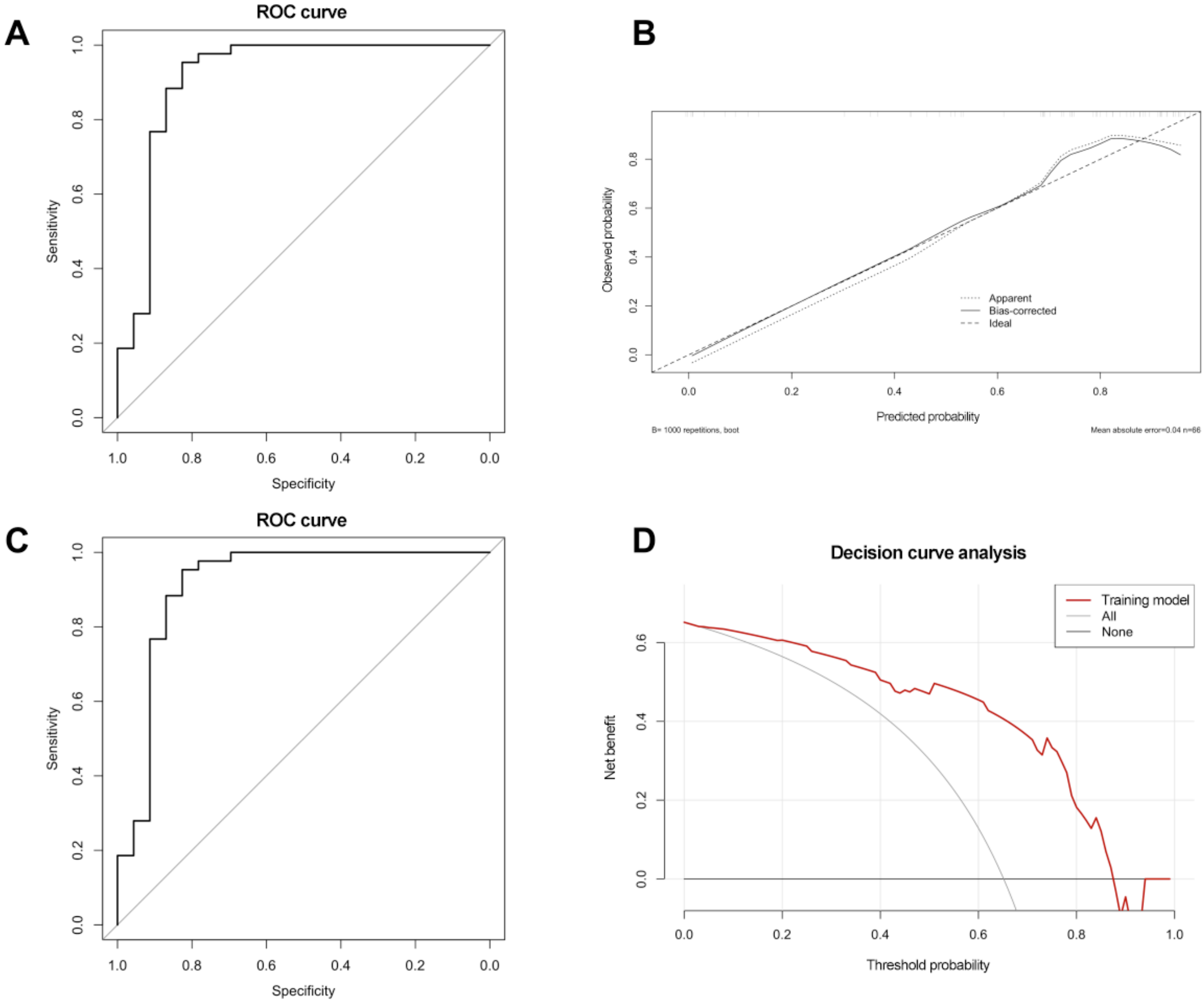


Figure 4. Evaluation and validation of the nomogram model for predicting adherence to acupuncture treatment in patients with Alzheimer disease. (A) Receiver operating characteristic (ROC) curve assessing the discriminative ability of the model, with an area under the ROC curve of 0.914, indicating excellent predictive performance. (B) Calibration curve illustrating agreement between predicted and observed adherence probabilities. The x-axis represents predicted probability, while the y-axis shows the observed probability of adherence. The Ideal line indicates perfect concordance between predicted and actual outcomes. The apparent line reflects model performance in the original sample, and the bias-corrected line shows model performance adjusted for overfitting using the bootstrap method (1000 resamples). The close alignment of the calibration curves with the ideal line demonstrates the high accuracy and reliability of the model. (C) ROC curve of the nomogram model in the validation cohort, with an area under the ROC curve of 0.838. (D) Decision curve analysis of the nomogram model in the training cohort. The x-axis represents the threshold probability, and the y-axis indicates the net clinical benefit across different thresholds.



Sensitivity Analysis

To evaluate the robustness of our model, a sensitivity analysis was performed using alternative adherence thresholds of 70% and 90%. The model's predictive performance remained strong across all definitions, with an area under the ROC curve (AUC) of 0.892 for the 70% threshold and 0.860 for the 90% threshold. The statistical significance of the number of treatments in the first month (*P*=.01 at 70%; *P*=.05 at 90%), caregiving role (*P*=.03 at 90%), and disease duration (*P*=.08 at 70%; *P*=.09 at 90%) fluctuated near the *P*=.05 cutoff; this could be caused by the dataset size. Overall, these findings support the stability of our model. Detailed results of the sensitivity analysis are provided in the supplementary materials (Figure S1 in Multimedia Appendix 1).

Adverse Events

Safety monitoring was conducted throughout both RCTs. No serious adverse events related to acupuncture treatment were reported in either the development cohort or the validation cohort. Adverse events unrelated to the intervention are summarized in Table S4 in Multimedia Appendix 1.

Discussion

Predictors of Treatment Adherence in Patients with AD

This study identified the number of treatments in the first month, caregiving role, and disease duration as significant predictors of adherence to acupuncture treatment among patients with AD. The predictive nomogram constructed with these 3 variables demonstrated excellent discrimination in the development cohort (AUC=0.914) and acceptable performance in the external validation cohort (AUC=0.838). With its simplicity, interpretability, and ease of application, the nomogram may serve as a practical tool to assist clinicians in identifying patients at risk of poor adherence and tailor intervention strategies accordingly.

To understand the clinical implications of these predictors, it is useful to interpret them through the established framework of intentional versus unintentional nonadherence. Unintentional nonadherence typically arises from practical barriers, which can be internal to the patient's condition (eg, forgetfulness and functional limitations) or external and situational (eg, transportation difficulties, inclement weather, and systemic disruptions like a pandemic). In contrast, intentional nonadherence involves a deliberate decision to cancel the treatment plan, often driven by subjective factors such as a perceived lack of efficacy, treatment fatigue, or shifting personal priorities. Our findings suggest that the identified predictors likely influence adherence through mechanisms related to both categories.

Increasing the Number of Treatments in the First Month May Improve Adherence

A key finding of this study is that a higher number of treatments during the first month was a powerful predictor of adherence. This suggests that early and intensive engagement is critical for establishing sustained treatment behaviors. From the perspective of unintentional nonadherence, a structured, frequent schedule in the initial phase may help patients and caregivers with cognitive and organizational deficits to more quickly accept the therapeutic routine, making it a habitual part of their lives [21].

This intensive approach can also mitigate intentional nonadherence through two possible ways. First, given that the therapeutic effect of acupuncture is often cumulative [22], increasing treatment frequency may accelerate the perception of clinical benefits [23,24]. When patients and caregivers observe an improvement early on, their motivation and belief in the treatment's value are naturally reinforced [25]. Second, frequent sessions offer more opportunities for communication among patients, caregivers, and clinicians, enabling the early detection and management of emerging issues. However, this approach requires careful consideration. It is important to acknowledge that a higher treatment frequency does not universally guarantee improved clinical outcomes [26] and may increase the treatment burden on families. The optimal number and timing of sessions may vary depending on disease severity, stage, and individual patient needs [27]. In conclusion, tailoring the intensity of acupuncture interventions to individual profiles remains a key consideration for future research and clinical practice for patients with AD, balancing clinical benefit and treatment burden.

Caregiver Capacity Is a Critical Determinant of Adherence

Our finding that patients supported by part-time caregivers (defined as providing fewer than 41 h of care per week [28]) were significantly less likely to adhere underscores the critical role of caregiver capacity in treatment engagement. This highlights a powerful driver of unintentional nonadherence, as dementia caregiving is a long-term, high-burden undertaking [29-31]. In advanced stages, care demands can often exceed 100 hours per week [32]. This presents formidable logistical barriers for part-time caregivers, who are typically adult children trying to balance employment, their own household duties, and the challenge of not living with the person they care for. These challenges frequently limit their ability to schedule and accompany patients to appointments, provide consistent emotional support, or respond to emergent care needs, thereby compromising adherence despite their best intentions [29].

This relentless demand also takes a significant physical and psychological toll, leading to caregiver burden and fatigue that can further diminish the capacity to support treatment [33-37]. These inherent difficulties are often compounded by unpredictable external factors, such as sudden illness or bad weather, which can disproportionately disrupt the routines of caregivers with less flexibility. A more subtle yet powerful factor is a form of intentional nonadherence driven by altruism, where patients may forgo appointments out of a desire not to burden their children, ultimately leading to treatment discontinuation.

In contrast, spouses often serve as full-time caregivers (≥41 h/wk) [38,39] and are typically more emotionally invested and committed to maintaining treatment routines. Some may even retire early to provide round-the-clock care [40]. Our study also found that professional caregiving within institutional settings can offer stable, structured, and high-quality care, which may facilitate better adherence. To enhance adherence in clinical practice, health care providers should actively involve caregivers in treatment planning and offer tailored education to improve their understanding of disease progression, treatment goals, and the importance of adherence [41,42]. Such interventions can mitigate both unintentional nonadherence (by improving scheduling and problem-solving skills) and intentional nonadherence (by reinforcing the perceived value of the treatment). Strengthening caregivers' motivation and capacity to support treatment is essential. When informal caregiving resources are insufficient, incorporating professional home care services or transitioning to institutional care may help ensure treatment continuity and effectiveness. Notably, countries such as Denmark and the Netherlands have developed comprehensive formal care systems for dementia that integrate medical and social support services [30,43]. These models may provide valuable reference points for improving dementia care infrastructure in China.

Early Detection, Prevention, and Intervention Still Key to Treating AD

Our study found that longer disease duration was associated with poorer adherence to acupuncture treatment. Primarily, this is a form of unintentional nonadherence driven by the patient's own progressive cognitive and functional decline, which impairs their capacity to independently manage appointments [44]. A higher prevalence of neuropsychiatric symptoms can reduce patient cooperation, while accumulating physical comorbidities and mobility limitations create new logistical hurdles. As the disease advances, a cascade of factors converges to further undermine adherence. This decline simultaneously increases the burden on caregivers, who may experience emotional exhaustion and a deterioration in their own health, diminishing their ability to provide consistent support [45]. The combination of increasing patient dependency, rising caregiver exhaustion, and mounting logistical obstacles creates a formidable barrier to sustained treatment in the later stages of the disease.

This underscores the critical importance of a proactive and early approach to management. Our findings emphasize that the timing of intervention is paramount. Initiating treatment when patients retain greater cognitive and functional capacity offers a crucial window of opportunity. Early initiation of acupuncture may help establish regular treatment routines, foster therapeutic rapport, and enhance patient motivation. Moreover, health care providers should prioritize early education and ongoing support for both patients and caregivers. This includes training in caregiving skills, psychological counseling, and practical strategies to reduce caregiver burden and enhance quality of life [46]. These factors can contribute not only to improved adherence but also to more favorable long-term outcomes by potentially slowing the trajectory of disease progression.

Limitations

This study has limitations. First, although this study suggests that early, intensive treatment may enhance adherence, the optimal frequency and total number of acupuncture sessions for patients with AD remain undetermined. In real-world settings, increasing treatment frequency may impose greater transportation, time, and financial burdens on both patients and caregivers, potentially reducing their motivation and adherence. Second, our analysis was based on data from previous RCTs that were provided free of charge. Consequently, we could not assess the influence of crucial socioeconomic factors, such as treatment costs, or household income, which are known to be factors of health care engagement. Their influence on adherence may have been underestimated. Third, our sample size (N=108) was adequate for the primary analysis; it may be underpowered to detect predictors with more subtle effects. Finally, although we performed external validation, both cohorts were recruited from a single hospital. This shared clinical and demographic context limits the generalizability of our nomogram. Studies are needed to validate our model in multicenter or community-based cohorts to confirm its broader applicability.

Conclusions

In conclusion, this study is the first to develop and validate a predictive model for acupuncture treatment adherence in patients with AD, offering a novel, evidence-based tool for both clinical research and practice. For clinical research, this model provides a method to stratify enrollment or identify participants who may require enhanced adherence support, thereby reducing bias and improving the integrity of future trials. In clinical practice, the nomogram enables a shift from reactive problem-solving to proactive adherence management. By prospectively identifying patients at high risk, clinicians can address specific barriers and implement targeted strategies to improve adherence and, ultimately, enhance therapeutic outcomes.

Acknowledgments

The authors extend their heartfelt gratitude to all patients with Alzheimer disease and their family members for their participation in this study, without whom this research would not have been possible. During the preparation of this work, the authors used ChatGPT to improve the language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the final content of the publication.

Funding

This study was supported by the Clinical Research and Achievement Transformation Capacity Improvement Project of High-Level Traditional Chinese Medicine Hospitals (HLCMHPP2023069); the Scientific and Technological Innovation Project—Innovation Team of the China Academy of Chinese Medical Sciences (CI2021B012); and the Surplus Fund of the Guang'anmen Hospital Scientific Research Project of the China Academy of Chinese Medical Sciences (2023013).

Data Availability

The ethical approval document specifies that access to the data collected in this project is limited to core members of the project team to safeguard patient privacy. Data may, however, be available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: ZHC, BHJ
Data curation: ZHC, RL, YHJ, SSY, GHZ, ZXY, XYR, JKH
Formal analysis: ZHC
Methodology: ZHC, RL, JKH
Supervision: BHJ
Writing – original draft: ZHC
Writing – review & editing: ZHC, RL, YHJ, JKH, BHJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
.Supplementary analyses supporting the development and validation of the adherence prediction model.
[DOCX File, 256 KB - aging_v9i1e82787_app1.docx]

References

1. Tatulian SA. Challenges and hopes for Alzheimer's disease. Drug Discov Today 2022 Apr;27(4):1027-1043. [doi: 10.1016/j.drudis.2022.01.016] [Medline: 35121174]
2. Dong XQ, Li XY, Kong XH, et al. Analysis of clinical application patterns in acupuncture-moxibustion treatment of Alzheimer disease. J Acupunct Tuina Sci 2020 Jun;18(3):238-246. [doi: 10.1007/s11726-020-1183-y]
3. Xia KP, Pang J, Li SL, Zhang M, Li HL, Wang YJ. Effect of electroacupuncture at governor vessel on learning-memory ability and serum level of APP, Aβ1-42 in patients with Alzheimer's disease. Zhongguo Zhen Jiu 2020 Apr 12;40(4):375-378. [doi: 10.13703/j.0255-2930.20190728-0003] [Medline: 32275365]
4. Ouyang Q, Mu YY. Clinical observation of electroacupuncture combined with perphenazine in treating psychiatric symptoms of Alzheimer's disease. Shanghai J Acupunct Moxibustion 2000;06:16-17. [doi: 10.13460/j.issn.1005-0957.2000.06.009]
5. Lin L, Ma X, Wang G, Wang HZ, Wang ZQ, Wang ZW, et al. Chinese guidelines for early prevention of Alzheimer's disease (2024). Chin J Alzheimer Dis Relat Disord 2024;7(3):168-175. [doi: 10.3969/j.issn.2096-5516.2024.03.002]
6. Wang YY, Yu SF, Xue HY, Li Y, Zhao C, Jin YH. Effectiveness and safety of acupuncture for the treatment of Alzheimer's disease: a systematic review and meta-analysis. Front Aging Neurosci 2020;12:98. [doi: 10.3389/fnagi.2020.00098] [Medline: 32435187]
7. Ke C, Shan S, Yu J, Wei X, Pan J, Zhang W. Acupuncture for patients with Alzheimer's disease: an evidence map of randomized controlled trials, systematic reviews, and meta-analysis. J Alzheimers Dis 2024 Dec;102(4):924-942. [doi: 10.1177/13872877241295400] [Medline: 39544007]
8. Olchanski N, Daly AT, Zhu Y, et al. Alzheimer's disease medication use and adherence patterns by race and ethnicity. Alzheimers Dement 2023 Apr;19(4):1184-1193. [doi: 10.1002/alz.12753] [Medline: 35939325]
9. Xiong S, Wu J, Li M, Lu K. PDG38 association between satisfaction with quality of care and ANTI-dementia medication adherence among elderly adults with Alzheimer's disease and related dementias. Value Health 2021 Jun;24:S93. [doi: 10.1016/j.jval.2021.04.486]
10. Wu XW, Yang HB, Yuan R, Long EW, Tong RS. Predictive models of medication non-adherence risks of patients with T2D based on multiple machine learning algorithms. BMJ Open Diabetes Res Care 2020 Mar;8(1):e001055. [doi: 10.1136/bmjdrc-2019-001055] [Medline: 32156739]
11. Koesmahargyo V, Abbas A, Zhang L, et al. Accuracy of machine learning-based prediction of medication adherence in clinical research. Psychiatry Res 2020 Dec;294:113558. [doi: 10.1016/j.psychres.2020.113558] [Medline: 33242836]
12. Nagpal TS, Mottola MF, Barakat R, Prapavessis H. Adherence is a key factor for interpreting the results of exercise interventions. Physiotherapy 2021 Dec;113:8-11. [doi: 10.1016/j.physio.2021.05.010] [Medline: 34555674]
13. Wang XL, Hu HQ, Wen Q, Li N. Considerations and suggestions on influencing factors of compliance in clinical trial of acupuncture and moxibustion: experience in case study of knee osteoarthritis treated with acupuncture. Zhongguo Zhen Jiu 2021 Sep 12;41(9):1045-1048. [doi: 10.13703/j.0255-2930.20200910-k0002] [Medline: 34491656]
14. Duan HL, Du ZJ, Zhang WR, Dong XF. Influencing factors of acupuncture adherence based on literature review. World's Latest Medical Information Digest 2018;18(98):146-148. [doi: 10.19613/j.cnki.1671-3141.2018.98.063]
15. Yu MK, Hu RX, Wen LZ, Li X, Zhao ZY, et al. Characteristics of the methodological studies on patient compliance in clinical trials in China. Chin J EvidBased Med 2019;19(6):708-714. [doi: 10.7507/1672-2531.201812053]
16. Wei J, Liu HL, Wang LC, Wang LP, Liu ZS, Yi Y. Discussion on improving compliance of participants in acupuncture-moxibustion clinical trial. Beijing J Tradit Chin Med 2014;33(2):97-99. [doi: 10.16025/j.1674-1307.2014.02.005]
17. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):263-269. [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
18. Cui Y, Zhang J, Wang Y, et al. Multivariate predictive model of the therapeutic effects of metoprolol in paediatric vasovagal syncope: a multi-centre study. EBioMedicine 2025 Mar;113:105595. [doi: 10.1016/j.ebiom.2025.105595] [Medline: 39946834]
19. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. 2017 Presented at: 31st Conference on Neural Information Processing Systems (NIPS 2017); Dec 4-9, 2017 URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf [accessed 2025-12-30]
20. David W, Lemeshow S, Sturdivant RX. Applied Logistic Regression: John Wiley & Sons; 2013. [doi: 10.1002/9781118548387]
21. Wang YY, Liu Z, Wu Y, et al. Acupuncture for smoking cessation in Hong Kong: a prospective multicenter observational study. Evid Based Complement Alternat Med 2016;2016:2865831. [doi: 10.1155/2016/2865831] [Medline: 28003848]
22. White A, Cummings M, Barlas P, et al. Defining an adequate dose of acupuncture using a neurophysiological approach--a narrative review of the literature. Acupunct Med 2008 Jun;26(2):111-120. [doi: 10.1136/aim.26.2.111] [Medline: 18591910]

23. Xu G, Lei H, Huang L, et al. The dose-effect association between acupuncture sessions and its effects on major depressive disorder: a meta-regression of randomized controlled trials. J Affect Disord 2022 Aug 1;310:318-327. [doi: 10.1016/j.jad.2022.04.155] [Medline: 35504399]
24. Zhang Y, Yang R, Zhang C, Han L. Dose-effect relationship between the number of acupuncture sessions and efficacy for cervical vertigo: a meta-regression analysis based on randomized controlled trials. Zhongguo Zhen Jiu 2025 Aug 12;45(8):1180-1186. [doi: 10.13703/j.0255-2930.20240711-0002] [Medline: 40825705]
25. Li J, Aulakh N, Culum I, Roberts AC. Adherence to non-pharmacological interventions in Parkinson's disease: a rapid evidence assessment of the literature. J Parkinsons Dis 2024;14(s1):S35-S52. [doi: 10.3233/JPD-230266] [Medline: 38640167]
26. Wang TS, Bai P. Clinical study on 37 cases of periarthritis treated with different acupuncture frequencies. Jiangsu J Tradit Chin Med 2019;51(7):62-65 [FREE Full text] [doi: 10.3969/j.issn.1672-397X.2019.07.021]
27. Li Y. Study on the time-effect relationship of acupuncture between needle retaining time and the interval time of acupuncture. Liaoning J Tradit Chin Med 2013;40(3):538-540. [doi: 10.13192/j.ljtcm.2013.03.160.liy.029]
28. Caregiving in the United States 2020. : Health Resources and Services Administration; 2020 URL: https://www.hrsa.gov/sites/default/files/hrsa/advisory-committees/nursing/reports/report-caregiving-us-2020.pdf [accessed 2025-12-30]
29. Committee on Family Caregiving for Older Adults, Board on Health Care Services, Health and Medicine Division, National Academies of Sciences, Engineering, and Medicine. In: Schulz R, Eden J, editors. Families Caring for an Aging America: National Academy of Sciences; 2016. [doi: 10.17226/23606]
30. Chen S, Cao Z, Nandi A, et al. The global macroeconomic burden of Alzheimer's disease and other dementias: estimates and projections for 152 countries or territories. Lancet Glob Health 2024 Sep;12(9):e1534-e1543. [doi: 10.1016/S2214-109X(24)00264-X] [Medline: 39151988]
31. Xu P, Zhong QL. Research progress on long-term care services for disabled elderly in community-based home settings. Chin J Gerontol 2016;36(12):3076-3078 [FREE Full text] [doi: 10.3969/j.issn.1005-9202.2016.12.116]
32. Ma WJ, Wang YH. Influencing factors of informal home care time. Chin J Gerontol 2018;38(14):3560-3563 [FREE Full text] [doi: 10.3969/j.issn.1005-9202.2018.14.088]
33. Berge LI, Angeles RC, Gedde MH, et al. Burden and care time for dementia caregivers in the LIVE@Home.Path trial. Alzheimers Dement 2025 Mar;21(3):e14622. [doi: 10.1002/alz.14622] [Medline: 40042468]
34. Golics CJ, Basra MKA, Finlay AY, Salek S. The impact of disease on family members: a critical aspect of medical care. J R Soc Med 2013 Oct;106(10):399-407. [doi: 10.1177/0141076812472616] [Medline: 23759884]
35. Feng Y, Wu QH, Lin SB, Lin H, Huang XH, Liu BB. Study on the correlation between caregiver burden, coping style, and acceptance behavior in primary caregivers of Alzheimer's disease patients. Chin Gen Pract Nurs 2022;20(32):4593-4596 [FREE Full text] [doi: 10.12104/j.issn.1674-4748.2022.32.035]
36. Du X, Shen J, Wu YF. Correlation between caregiving experience and fatigue among dementia caregivers in old age. Chin J Gerontol 2015;35(9):2504-2506 [FREE Full text] [doi: 10.3969/j.issn.1005-9202.2015.09.093]
37. Pinquart M, Sörensen S. Correlates of physical health of informal caregivers: a meta-analysis. J Gerontol B Psychol Sci Soc Sci 2007 Mar;62(2):126-137. [doi: 10.1093/geronb/62.2.p126] [Medline: 17379673]
38. Wenborn J, O'Keeffe AG, Mountain G, et al. Community occupational therapy for people with dementia and family carers (COTiD-UK) versus treatment as usual (Valuing Active Life in Dementia [VALID]) study: a single-blind, randomised controlled trial. PLoS Med 2021 Jan;18(1):e1003433. [doi: 10.1371/journal.pmed.1003433] [Medline: 33395437]
39. Lamb SE, Sheehan B, Atherton N, et al. Dementia And Physical Activity (DAPA) trial of moderate to high intensity exercise training for people with dementia: randomised controlled trial. BMJ 2018 May 16;361:k1675. [doi: 10.1136/bmj.k1675] [Medline: 29769247]
40. Lilly MB, Laporte A, Coyte PC. Labor market work and home care's unpaid caregivers: a systematic review of labor force participation rates, predictors of labor market withdrawal, and hours of work. Milbank Q 2007 Dec;85(4):641-690. [doi: 10.1111/j.1468-0009.2007.00504.x] [Medline: 18070333]
41. Liu JH, Wang XL, Li WL, Yang L, Wang JG. Effect of synchronous education for children caregivers on hip function recovery after intertrochanteric fracture surgery in elderly patients. Clin Res 2021;29(7):153-154 [FREE Full text]
42. Liu YY, Zhang L. Application of children caregivers' synchronous education in postoperative rehabilitation of elderly patients undergoing heart valve replacement. Contemp Nurse 2022;29(4):150-152. [doi: 10.19791/j.cnki.1006-6411.2022.10.044]
43. Coley N, Gallini A, Garès V, Gardette V, Andrieu S, ICTUS/DSA group. A longitudinal study of transitions between informal and formal care in Alzheimer disease using multistate models in the European ICTUS cohort. J Am Med Dir Assoc 2015 Dec;16(12):1104. [doi: 10.1016/j.jamda.2015.09.010] [Medline: 26593306]
44. MacNeil-Vroomen JL, Thompson M, Leo-Summers L, Marottoli RA, Tai-Seale M, Allore HG. Health-care use and cost for multimorbid persons with dementia in the National Health and Aging Trends Study. Alzheimers Dement 2020 Sep;16(9):1224-1233. [doi: 10.1002/alz.12094] [Medline: 32729984]
45. Balkrishnan R, Housman TS, Carroll C, Feldman SR, Fleischer AB. Disease severity and associated family impact in childhood atopic dermatitis. Arch Dis Child 2003 May;88(5):423-427. [doi: 10.1136/adc.88.5.423] [Medline: 12716715]

46. Zhang DY, Yang JM, Liu H. Correlation between psychological resilience and caregiving burden among caregivers of patients with senile dementia. Chin Gen Pract Nurs 2018;16(32):3969-3973 [FREE Full text] [doi: 10.12104/j.issn.1674-4748.2018.32.001]

Abbreviations

AD: Alzheimer disease
AUC: area under the receiver operating characteristic curve
CACMS: China Academy of Chinese Medical Sciences
OR: odds ratio
PDC: proportion of days covered
RCT: randomized controlled trial
ROC: receiver operating characteristic
SHAP: Shapley Additive Explanations

Edited by Q Fan; submitted 21.Aug.2025; peer-reviewed by H Wu, WX Li, X Wang; accepted 17.Dec.2025; published 21.Jan.2026.

Please cite as:
Chen ZH, Li R, Jiang YH, He JK, Yan SS, Zong GH, Yi ZX, Ren XY, Jia BH
Predictive Model of Acupuncture Adherence in Alzheimer Disease: Secondary Analysis of Randomized Controlled Trials
JMIR Aging 2026;9:e82787
URL: https://aging.jmir.org/2026/1/e82787
doi:10.2196/82787

Original Paper

The Virtual Kitchen Challenge—Version 2: Validation of a Digital Assessment of Everyday Function in Older Adults

Marina Kaplan[1], BA; Moira McKniff[1], MA; Stephanie M Simone[1,2], PhD; Molly B Tassoni[1], MA; Katherine Hackett[1,3]; Sophia Holmqvist[1], MA; Rachel E Mis[1,4], PhD; Kimberly Halberstadter[1], MA; Riya Chaturvedi[1], BS; Melissa Rosahl[1], BA; Giuliana Vallecorsa[1,5], MS; Mijiail D Serruya[6], MD, PhD; Deborah A G Drabick[1], PhD; Takehiko Yamaguchi[7], PhD; Tania Giovannetti[1], PhD

[1]Department of Psychology and Neuroscience, Temple University, Philadelphia, PA, United States
[2]Department of Psychiatry, Neuropsychology Section, Michigan Medicine, Ann Arbor, MI, United States
[3]Division of General Internal Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, United States
[4]Department of Neurology, Dell Medical School, University of Texas at Austin, Austin, TX, United States
[5]Department of Psychology, LaSalle University, Philadelphia, PA, United States
[6]Raphael Center for Neruorestoration, Thomas Jefferson University, Philadelphia, PA, United States
[7]Department of Applied Information Engineering, Suwa University of Science, Nagano, Japan

Corresponding Author:
Tania Giovannetti, PhD
Department of Psychology and Neuroscience
Temple University
Weiss Hall
1701 N 13th Street
Philadelphia, PA, 19122
United States
Phone: 1 484 843 1321
Email: tgio@temple.edu

Abstract

Background: Conventional methods of functional assessment include subjective self- or informant report, which may be biased by personal characteristics, cognitive abilities, and lack of standardization (eg, influence of idiosyncratic task demands). Traditional performance-based assessments offer some advantages over self- or informant reports but are time-consuming to administer and score.

Objective: This study aims to evaluate the validity and reliability of the Virtual Kitchen Challenge—Version 2 (VKC-2), an objective, standardized, and highly efficient alternative to current functional assessments for older adults across the spectrum of cognitive aging, from preclinical to mild dementia.

Methods: A total of 236 community-dwelling, diverse older adults completed a comprehensive neuropsychological evaluation to classify cognitive status as healthy, mild cognitive impairment, or mild dementia, after adjustment for demographic variables (age, education, sex, and estimated IQ). Participants completed 2 everyday tasks (breakfast and lunch) in a virtual kitchen (VKC-2) using a touchscreen interface to select objects and sequence steps. Automated scoring reflected completion time and performance efficiency (eg, number of screen interactions, percentage of time spent off-screen, interactions with distractor objects). Participants also completed the VKC-2 tasks using real objects (Real Kitchen). All participants and informants for 219 participants completed questionnaires regarding everyday function. A subsample of participants (n=143) performed the VKC-2 again in a second session, 4-6 weeks after the baseline, for retest analyses. Analyses evaluated construct and convergent validity, as well as retest and internal reliability, of VKC-2 automated scores.

Results: A principal component analysis showed that the primary VKC-2 automated scores captured a single dimension and could be combined into a composite score reflecting task efficiency. Construct validity was supported by analyses of covariance results showing that participants with healthy cognition obtained significantly better VKC-2 scores than participants with cognitive impairment (all *Ps*<.001), even after controlling for demographics and general computer visuomotor dexterity. Convergent validity was supported by significant correlations between VKC-2 scores and performance on the Real Kitchen (r=−0.58 to 0.64, *Ps*<.001), conventional cognitive test scores (r=−0.50 to −0.22, *Ps*<.001), and self- and informant report questionnaires evaluating everyday function (r=0.25 to 0.43, *Ps*<.001). Intraclass correlation coefficients (ICCs) indicated moderate to excellent retest reliability (ICC=0.70-0.90) for VKC-2 scores after 4-6 weeks. Reliability improved in analyses including only participants who reported no change in cognitive status between time 1 and time 2 (n=123). Spearman-Brown correlations showed acceptable to good internal consistency between the VKC-2 tasks (breakfast and lunch) for all scores (0.77-0.84), supporting the use of total scores.

Conclusions: The VKC-2 is an efficient, valid, and sensitive measure of everyday function for diverse older adults and holds promise to improve the status quo of functional assessment in aging, particularly when informants are unavailable or unreliable.

(JMIR Aging 2026;9:e82092) doi:10.2196/82092

KEYWORDS

everyday function; activities of daily living; assessment; dementia; Alzheimer disease; neuropsychology; cognition; mild cognitive impairment; virtual reality; digital assessment

Introduction

As the US population ages and interventions for Alzheimer disease and Alzheimer disease–related dementias become available [1], highly sensitive, objective, and efficient measures of functional abilities are needed for multiple purposes. Mild functional difficulties are among the strongest predictors of future cognitive decline and dementia [2-5]; thus, accurate measurement of functional ability will improve prognostic prediction and help identify the need for early intervention. Given that functional ability level is often the criterion that distinguishes mild cognitive impairment (MCI) from mild dementia, accurate assessment is critical for diagnostic decision-making [6,7]. According to the Food and Drug Administration, the approval of pharmacological treatments for dementia, even at the very early, presymptomatic stage, is contingent on demonstrating gains on meaningful measures of functioning [8]. Recently approved treatments have relied on composite measures such as the Clinical Dementia Rating-Sum of Boxes and the integrated Alzheimer’s Disease Rating Scale, but these measures require specialized training, are not readily deployable in typical clinical settings, and lack sensitivity to the earliest functional changes [9,10]. There exists a critical need for sensitive and efficient functional assessment tools that are clinically meaningful, psychometrically sound, and practically implementable across diverse health care settings [11,12]. We developed a nonimmersive virtual reality (VR) measure, the Virtual Kitchen Challenge—Version 2 (VKC-2), an objective, sensitive, efficient, and theoretically based tool for assessing everyday function in older adults to address the gaps in current functional assessments. Here we report results on VKC-2 validity and reliability in racially diverse, community-dwelling older adults with healthy cognition, MCI, or mild dementia.

Self/informant reports of everyday function, which are easy to administer and score, are the current standard method for functional assessment. When used with reliable, observant, and knowledgeable reporters, they generate useful information about how a person is functioning in everyday life [13-15]. In many circumstances, however, the accuracy of self and informant reports is uncertain. Their subjective nature makes them prone to over- or underreporting due to faulty cognitive abilities, psychological factors (eg, denial, depression, burden), or cultural beliefs [16]. Informant reports are often unavailable, as many older adults do not have a living spouse, nearby family members, or close friends. Even when available and willing, informants may have limited opportunities to observe daily functioning and may lack knowledge, particularly when functional difficulties are mild and may be masked by compensatory behaviors [17,18].

Another limitation of questionnaires is that older adults vary widely in the activities they perform and the contexts in which they perform them. For example, informant-reported difficulties with medication management may be profoundly different for an older adult managing a single prescription while residing in a small, highly organized home with her spouse versus an older adult taking dozens of medications while living alone in a large, cluttered house [19]. However, given identical clinical presentations and cognitive test scores suggesting mild cognitive decline, the latter patient would likely be diagnosed with clinical dementia if she were unable to independently manage her medications. Failure to account for context and task complexity confounds the informant report of everyday function and precludes clear comparisons of functional abilities across individuals.

Further, many questionnaires do not distinguish difficulties due to physical versus cognitive limitations [14], and if they do, it may be difficult for an informant to fully understand the nature of the functional difficulties, particularly because physical and cognitive limitations often co-occur [20-22]. Informant and self-reports also do not offer a detailed characterization of types of functional difficulties arising from different underlying cognitive problems (eg, slowing, disorganized actions vs omission of crucial task steps), which could offer insights into interventions for improving function and reducing the risk of future functional disability [23,24].

Performance-based measures of function address many of the limitations of questionnaires; they are objective, standardize task complexity and context, and allow for detailed analysis of behavior and systematic comparison across individuals. The Naturalistic Action Test (NAT), for example, is a performance-based test of everyday function with strong psychometric properties, normative data, and suggested cut scores for healthy cognition versus MCI versus mild dementia [23,25-35]. Scoring NAT performance for subtle inefficient errors, called micro-errors, has increased the sensitivity of NAT tasks for detecting mild difficulties with everyday tasks [35-38]. Results from performance-based tests, such as the NAT, with added sensitive scoring procedures, have demonstrated that (1) healthy older adults make more errors and require more time to complete everyday tasks than younger adults [36,37,39-43]; (2) people with MCI make more errors than healthy controls but fewer errors than individuals with dementia [32,35,44-46]; and (3) the ability to accurately and efficiently perform everyday tasks is moderately correlated with performance on cognitive tests [27,28,35,37,47] and informant report of everyday function [27,29,31,35,48]. Together, these findings and others [37,49-52] suggest that standardized performance-based assessment of function is valid and reliable.

Despite their objectivity, validity, and potential for rich characterizations of function, current performance-based tests require extraordinary effort, limiting their implementation and scalability. Scoring, particularly scoring for subtle errors and inefficiencies, is time-intensive and requires video recording, detailed scoring instructions, and trained coders. Although some performance-based tests may be scored quickly as pass/fail without video recording [40,53], such gross measures are less sensitive to mild difficulties (ie, MCI) [54], do not advance our understanding of the nature of functional problems [5,55,56], or still require considerable effort to administer. To streamline administration and scoring, a nonimmersive VR task called the Virtual Kitchen, modeled after the NAT, was developed. The original version of the Virtual Kitchen [57] required a mouse to move objects on a computer screen to complete a coffee-making task. Results showed that people with dementia accomplished fewer steps and made more errors than healthy controls on the Virtual Kitchen. Validity was also supported by significant correlations between Virtual Kitchen scores and performance of real tasks, cognitive tests, and informant reports of functioning [57].

Our team revised the original Virtual Kitchen [57] by implementing the following updates: (1) expanding the coffee task to include a more extensive breakfast; (2) adding a lunch task; (3) updating the graphics; and (4) transitioning from a mouse to a computer touchscreen to make interactions more natural [39]. We also added a brief training task to familiarize participants with the touchscreen interface. Automated scores were expanded to include measures computed based on interactions with the touchscreen to increase sensitivity (ie, number of screen interactions). Preliminary results from the revised task, which we called the Virtual Kitchen Challenge (VKC), demonstrated validity and good internal consistency [39]. The VKC automated scores have been validated against conventional cognitive tests in young adults [52] and against neuroimaging markers of cerebral vascular disease (white matter hyperintensities) in a small sample of community-dwelling older adults [48].

In this paper, we present the psychometric properties of the automated scores from the most recent revision of the Virtual Kitchen, the VKC-2. This version includes enhanced graphics and a more extensive basic familiarization task for practice and to obtain a score of participants' digital visuomotor dexterity that may be used as a control measure. We evaluated construct and convergent validity as well as retest and internal reliability in a large, community-based sample of racially diverse older adults with healthy cognition, MCI, or mild dementia. Construct validity of the VKC-2 automated scores was evaluated in a known-group comparison (healthy cognition vs MCI vs mild dementia). Convergent validity was evaluated with correlations between automated VKC-2 measures and performance on the real versions of the VKC-2 tasks (Real Kitchen), demographically adjusted cognitive test scores, and conventional self/informant questionnaires of everyday function. Retest reliability was evaluated over a period of 4-6 weeks. Internal consistency was evaluated for the 2 VKC-2 tasks (breakfast and lunch).

Methods

Participants

Participants were recruited for an observational, longitudinal psychometric study designed to evaluate the psychometric properties of the VKC-2 (n=217; grant R01AG062503) or for a separate, smaller study on activity tracking (n=20; grant F31AG089944). Procedures for the baseline visit of both studies were the same, designed and conducted in accordance with the Helsinki Declaration, and approved by the Institutional Review Board at Temple University (institutional review board protocols 23116 and 29712). All participants and a knowledgeable informant signed informed consent forms, were compensated for their participation (US $50 for participants per session and US $25 for informants per session), and were assigned study numbers to protect their privacy when storing research records. At the end of the study, participants were also offered a research report with their cognitive test scores, if interested.

All participants were recruited from community outreach events, fliers, and referrals from neurology departments in Philadelphia, Pennsylvania, from September 2020 to June 2025. Inclusion and exclusion criteria were screened by phone, with only minor differences between the 2 studies. In both studies, participants were excluded for the following reasons: lifetime history of severe psychiatric disorder (eg, schizophrenia, bipolar disorder); nervous system infections or disorders (eg, epilepsy, brain tumor); current metabolic or systemic disorders (eg, B_{12} deficiency, renal failure, cancer); current moderate-severe depression; current moderate-severe anxiety symptoms; severe sensory deficits that would preclude visual detection or identification of common everyday objects used in the study or the ability to hear task directions (eg, blindness, total hearing loss); severe motor weakness that would preclude the use of everyday objects (eg, severe deformities or paralysis of both upper extremities); intellectual disability; and not being a fluent English speaker. The inclusion criteria for the larger study required participants to be at least 65 years old and have an available informant who could serve as a study partner. Informants were screened by phone for the following eligibility criteria: 18 years of age or older; fluent English speaker; available and willing to complete study questionnaires in person, by phone, or online; has daily contact with the participant; and reports knowledge of the participant's daily functioning. Inclusion criteria for the second, smaller study required participants to be at least 55 years old and did not require a study informant/partner.

Procedures

At the baseline visit (session 1), participants (N=237) completed informed consent, cognitive testing, the VKC-2, the real version of the VKC-2 tasks (ie, Real Kitchen), and questionnaires regarding demographic information, familiarity with the tasks used in the VKC-2, and their ability to perform activities in everyday life. The order of the Real Kitchen and VKC-2 was counterbalanced across participants to control for order effects. At session 1, informants completed questionnaires in person, online, or at home by mail. After reaching our target sample size (n=140) for retest reliability analyses (June 2024), participants were no longer requested to return for a second session 4-6 weeks after session 1 [58]. A total of 143 participants completed session 2, which included a brief interview (for both the participant and informant) regarding changes in cognition or health status (eg, medication changes, falls, illnesses, hospitalizations) since session 1, as well as repeat administration of the VKC-2 and Real Kitchen.

Measures

Conventional Cognitive Tests

Cognitive tests were administered to characterize the sample, classify participants according to their cognitive status, and evaluate the convergent validity of the VKC-2. The cognitive testing protocol is described in Table 1. The protocol included 2 tests from 4 different cognitive domains to classify participants according to Jak/Bondi actuarial criteria [59,60] and clinical criteria originally proposed by Petersen [6] and McKhann [61]. Normative data from the Calibrated Neuropsychological Normative System [62] were used to enable raw score adjustments for sex, age, education, and IQ estimated by a test of reading/vocabulary. Such demographic adjustments are critical for confirming group membership in a diverse sample of older adults [63,64]. Further details on how tests were used for classifying cognitive abilities are provided in Multimedia Appendix 1.

Table 1. Cognitive tests administered at session 1.

Cognitive domain and test	Score(s)	Reference
Premorbid intellectual functioning (IQ)		
Hopkins Reading Test	Estimated IQ	Schretlen et al [65]
Global cognitive status		
Mini-Mental State Examination	Total correct	Folstein et al [66]
Episodic memory		
Hopkins Verbal Learning Test—Revised[a,b]	Delayed free recall total correct	Brandt and Benedict [67]
Brief Visual Memory Test—Revised[a]	Delayed free recall total correct	Benedict et al [68]
Language		
Category (Animal) Fluency[a,b]	Total correct	Schretlen et al [62]
Boston Naming Test—30 item[a]	Total correct	Goodglass and Kaplan [69]
Executive function		
Trail Making Test—Part B[a,b]	Completion time	Reitan [70]
Digit Span Backward[a]	Longest span	Wechsler [71]
Processing speed		
Salthouse Letter Comparison[a,b]	Total correct	Salthouse [72]
Salthouse Pattern Comparison[a,b]	Total correct	Salthouse [72]
Attention		
Digit Span Forward[a]	Longest span	Wechsler [71]
Trail Making Test A[a]	Completion time	Reitan [70]

[a] *t* scores from these tests were used for healthy versus mild cognitive impairment versus dementia classification.

[b] *t* scores from these tests were averaged to compute the modified Knight-Preclinical Alzheimer Cognitive Composite.

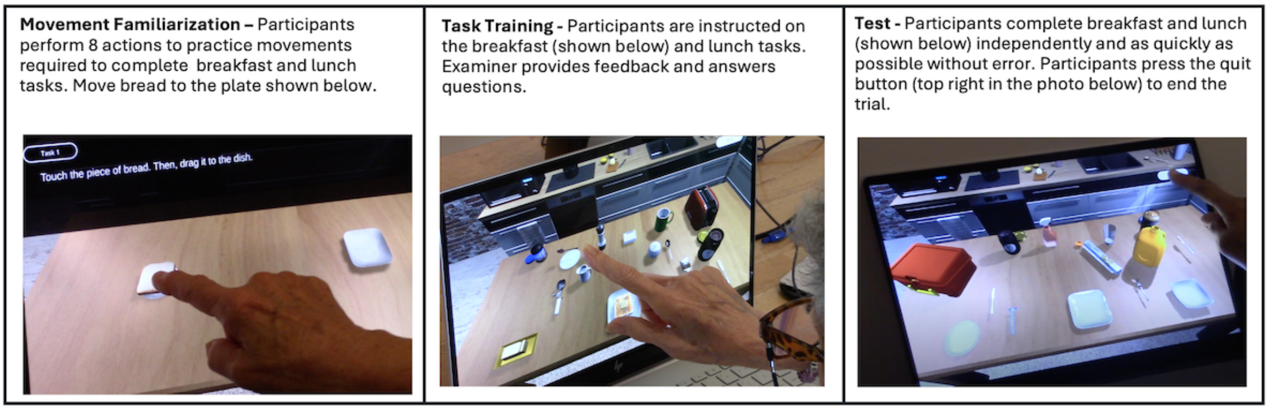
For the analysis of VKC-2 convergent validity, composite scores were computed by averaging demographically adjusted *t* scores from tests within each domain (eg, Episodic Memory, Language). A global cognitive composite was modeled after the Knight-Preclinical Alzheimer Cognitive Composite (modified [m]Knight-PACC) [73], which has been validated as a sensitive measure of early cognitive change due to neurodegenerative disease.

Virtual Kitchen Challenge-Version 2

The VKC-2 is a nonimmersive VR test of everyday function that requires participants to complete 2 everyday tasks (breakfast and lunch) by moving virtual objects using a touchscreen [39,57]. The VKC-2 tasks and objects were modeled after the NAT [25], an extensively studied and theoretically based performance-based test of everyday function that involves completion of familiar everyday tasks using real objects. The VKC-2 breakfast and lunch tasks were designed to be of comparable complexity and difficulty, with each task including 13 target objects and 4 distractor objects. For this study, the VKC-2 was administered on an MSI Creator Z16-A12UET laptop (12th Gen Intel Core i9 Processor) with a 16″ QHD+ (Quad High Definition Plus) (2560 × 1600), 120 Hz, IPS (In-Plane Switching)-level touchscreen display to maximize visibility and portability. Participants were instructed to use the index finger of their dominant hand to move and manipulate objects on the touchscreen.

The VKC-2 included 3 phases: Movement Familiarization, Task Training, and Test. See Figure 1 and the text below for more details.

Figure 1. Photos of participants completing each phase of the VKC-2.



VKC-2 Movement Familiarization

Participants were directed to perform 8 basic touchscreen actions (eg, tap, drag) to complete the following task steps: (1) move bread to dish, (2) stir mug with spoon, (3) pour juice, (4) place thermos in lunch box, (5) spread jelly on bread, (6) wrap cookies in foil, (7) place bread in toaster, and (8) add sugar to mug. Participants first performed all basic touchscreen actions with guidance from the examiner and had the opportunity to ask questions and repeat each action as needed. Next, participants were asked to complete all 8 trials independently as quickly and efficiently as possible. Completion time of the second, independent trial was computed as a measure of basic digital visuomotor dexterity (Digital Dexterity Score).

VKC-2 Task Training

The examiner reviewed the written instructions presented on the computer screen for each task. Participants were asked to point to each of the target objects needed for the task. For example, training for the breakfast task included the direction to “point to all of the objects you will need for the toast” while the examiner named each object out loud (eg, “bread,” “toaster”). Participants were also asked to point to each of the distractor objects and were told that they would not need to touch or use those objects. Participants then proceeded to practice trials, making breakfast and lunch with prompting, cues, and error correction from the examiner. The examiner also answered questions to ensure that participants fully understood each task.

VKC-2 Test

Breakfast and lunch tasks were completed independently without feedback. Instructions regarding the task objectives, which were reviewed during the practice trials, were repeated (eg, “pack a lunch for someone who wants a sandwich, snack, and a drink”). Participants were also instructed to complete test trials as quickly as possible, without making errors, and using clear and precise movements. They were told to touch the quit button at the top right of the screen to end the trial (see Figure 1). Participants were asked to verbally repeat the directions before each task to ensure comprehension; instructions were repeated as often as needed before the participant initiated the task.

VKC-2 Test Automated Scores

Performance on the VKC-2 Test tasks (breakfast and lunch) was scored using data from the touchscreen, as described and validated in our pilot work with the original version of the VKC [39,48,52]:

- Completion time (time) was recorded in seconds from the moment the virtual kitchen screen appeared (after instructions) until the participant pressed the quit button. Results from prior studies of the original VKC indicate that completion time differed significantly between older and younger participants and correlated with completion time on the Real Kitchen, cognitive tests of executive function and episodic memory [39], and neuroimaging markers of cerebrovascular disease [48].
- The number of screen interactions (touches) included the number of discrete instances the participant made contact with the computer touchscreen. Touches were collected as a measure of performance efficiency, with fewer screen interactions reflecting more precise and deliberate actions. Results from the original VKC showed that older adults made significantly more touches than younger adults, with additional touches by older adults including both inefficient correct actions and errors. A higher number of touches was significantly associated with more total errors scored by trained coders who watched video recordings of VKC

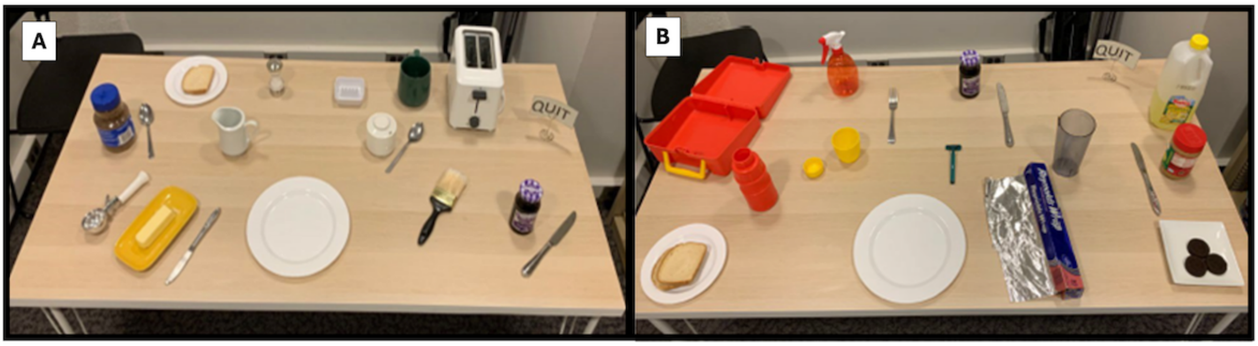
performance. Additionally, screen touches were significantly associated with performance on the Real Kitchen and cognitive tests of executive function and episodic memory [39].

- The percentage of time off-screen (%off-screen) was the percentage of time spent working on the VKC-2 when the participant was not touching the screen. It was computed by subtracting the time spent touching the screen from the completion time, dividing by the completion time, and multiplying by 100. The percentage of time off-screen also reflects performance efficiency. Pilot data from the original VKC [39] indicated that older adults spent a significantly higher percentage of their total time off-screen than younger adults. Correlations between %off-screen and human codes of VKC performance suggested that higher %off-screen times were due to multiple factors, including slower planning, difficulties locating target objects, difficulty resolving competition for object selection, and misreaching toward the computer screen (ie, micro-errors). Higher %off-screen times were significantly associated with more errors on the Real Kitchen, poorer scores on tests of executive function [52] and episodic memory [39], and neuroimaging markers of cerebrovascular disease [48].
- The number of distractor object interactions (distractor interactions) included instances when a distractor object was touched or moved. Our pilot work in a sample of healthy older and younger adults indicated that distractor interactions occurred too infrequently for analysis [39], but they have not been studied in participants with cognitive impairment.

Real Kitchen

The Real Kitchen required participants to complete the breakfast and lunch tasks using real objects placed on a table (Figure 2). Instructions for the Real Kitchen were identical to those for the VKC-2, including the instruction to "press the quit button when finished." In the Real Kitchen, the Quit Button was a piece of paper on the right side of the table labeled "QUIT." Real task objects were similar in appearance (color and shape) to the simulated objects in the VKC-2. Participants repeated the directions before each test trial to ensure comprehension; instructions were repeated as often as needed. Participants were video recorded, and recordings were labeled using a code so that human coders were unaware of participant classification and study session.

Figure 2. Real Kitchen breakfast (A) and lunch (B) tasks.



Real Kitchen performance was scored according to detailed instructions using validated scores and procedures. Real Kitchen scores from a subset of the current sample have been published and show strong interrater reliability, significant differences between participants with healthy cognition versus cognitive impairment, and correlations with cognitive tests and self/informant reports of everyday function [35]. For our current study, the following Real Kitchen scores were used to validate (convergent validity) the VKC-2 automated measures:

- Real Kitchen completion time was recorded in seconds and reliably coded by starting the timer when the first step was initiated and ending when the participant touched the quit button. Prior work shows that participants with greater cognitive impairment demonstrate longer completion times than participants with healthy cognition [35].
- Accomplishment was coded for each completed step and scored from 0 to 13 for the breakfast task and 0 to 20 for the lunch task. A total accomplishment score was computed (0-33), with higher scores reflecting a greater number of task steps accomplished.
- Total errors were coded according to a taxonomy studied in a range of clinical populations [25,74], showing validity and strong interrater reliability in people with stroke [75,76], dementia [27,28,30,47], MCI [26,32,33], and healthy controls [37,38,49], as well as a subset of participants from this sample [35]. The error taxonomy (see Multimedia Appendix 1) includes overt errors (eg, performing task steps in the wrong sequence) and micro-errors (eg, reaching toward a distractor object). In studies of participants with dementia, total overt errors correlate with cognitive tests and informant reports of function. The micro-error category was added to improve detection of subtle, inefficient behaviors in healthy and MCI participants [35,37,38,49]. As overt errors occur with relatively low frequency, they were combined with micro-errors to compute a total error score [35].
- Motor errors were tracked separately from total errors. Motor errors involved instances in which a corrective action was performed with motor or spatial imprecision (eg, spilling coffee grounds, dropping a knife).

Participant Questionnaires

Participants completed a demographic form assessing age, sex, race, ethnicity, income, and education level, as well as the following questionnaires.

The Past Experience Scale [45,77] assessed familiarity with the breakfast and lunch subtasks that comprise the VKC/Real Kitchen. The scale included 4 items (toast, coffee, sandwich, and thermos), each rated from 0 (not at all familiar) to 4 (very familiar). The total familiarity score ranged from 0 to 16, with higher scores reflecting greater familiarity. Participants also rated the frequency with which they had completed each subtask in their day-to-day life over the past 5-10 years, using a scale from 0 (never) to 4 (just about every day), with total scores ranging from 0 (never performed any of the tasks) to 16 (performed each task just about every day).

Functional Activity Questionnaire (FAQ) [14] instructions were modified to reflect only difficulties due to cognitive problems (not physical problems, fatigue, etc) for 10 activities (eg, preparing a balanced meal). Each activity is rated on a scale from 0 (performs normally) to 3 (dependent). Total FAQ scores range from 0 to 30, with higher scores reflecting greater dependence on others in everyday tasks due to cognitive difficulties.

The 12-item Everyday Cognition Scale (ECog-12) [13,78] measures decline over the past 10 years in 12 everyday cognitive abilities (eg, remembering where you have placed objects) on a scale from 1 (better or no change) to 4 (much worse all the time). Total scores reflect an average across all completed items and range from 1 to 4, with higher scores indicating greater decline in everyday cognition.

The Instrumental Activities of Daily Living—Compensation (IADL-C) [15] scale measures the need for assistance and compensatory strategies when performing 27 daily activities (eg, preparing one's own meals). Each activity is rated on a scale from 1 (independent, no aid) to 8 (not able to complete the activity anymore). The total score is the sum of all item responses, with a possible range from 27 (completely independent, no aid needed for any tasks) to 216 (no longer able to perform any task).

Informant Questionnaires

Informants completed questionnaires regarding their demographic information (eg, age, education), their relationship with the participant (eg, cohabitation, years known, hours in contact with the participant), and the participants' everyday function, including the ECog-12 [13,78], FAQ [14], and IADL-C [15]. Instructions and scoring for each questionnaire were the same as those for the participant versions described above.

Analysis Plan

Preliminary Analysis

Analyses were conducted using SPSS version 29.0 (IBM Corp) [79]. VKC-2 automated scores were examined for outliers and Winsorized at the first and ninety-ninth percentiles. The VKC-2 distractor interaction score was dichotomized because a few participants interacted with distractor objects (0=no interactions; 1=at least one interaction with a distractor object during completion of the VKC-2). Relations among VKC-2 scores were evaluated using bivariate correlations. Additionally, a principal component analysis (PCA) was conducted, including the 3 primary VKC-2 variables (time, touches, and %off-screen), to determine whether the dimensional VKC-2 automated scores could be combined into a single composite score. The Digital Dexterity score was not included in the PCA because it is derived from a separate condition intended to be used as a control for basic visuomotor skills. The distractor interaction score was not included because dichotomous variables are not appropriate for PCA. The suitability of the data for PCA was evaluated using the Kaiser-Meyer-Olkin measure of sampling adequacy and the Bartlett test of sphericity.

Construct Validity

VKC-2 automated scores were compared across groups known to differ in functional ability level: healthy cognition, MCI, and mild dementia. As the size of the dementia subgroup was relatively small (n=16), statistical analyses focused on differences between participants with healthy cognition and those with MCI. Participants with dementia were included for descriptive comparisons. One-way analyses of covariance (ANCOVA) were used to test group differences for each VKC-2 automated score (digital dexterity, time, touches, %off-screen, and VKC-2 composite) after controlling for demographics. Group differences were also evaluated in ANCOVA models that controlled for the digital dexterity score to determine whether significant group differences were explained by differences in basic visuomotor or computer abilities. Group differences on the dichotomized VKC-2 distractor interaction score were evaluated using chi-square tests. Significant between-group differences with at least small effect sizes (ie, partial η^2>.01; phi [ϕ] coefficient>.30) were interpreted as supporting the construct validity of the VKC-2 automated scores.

Receiver operating characteristic analyses comparing participant groups (healthy cognition vs impaired cognition [MCI + dementia]; healthy cognition vs MCI) were performed to identify cutoff values for each of the VKC-2 automated scores. Youden indices were used to identify cutoff scores that optimized sensitivity and specificity [80].

Convergent Validity

Correlations between the VKC-2 automated measures and the ability to perform tasks with real objects (Real Kitchen), demographically adjusted cognitive test scores of overall cognition and specific cognitive abilities, and self/informant reports of everyday functioning were performed to evaluate convergent validity. Pearson correlation coefficients were computed using the full sample. Spearman rank-order correlations were also performed and are included in Tables S3-S5 in Multimedia Appendix 1. Significant and moderate-level relationships were interpreted as supporting the convergent validity of the VKC-2 automated scores.

Reliability

Retest reliability was assessed using intraclass correlation coefficients (ICCs), calculated with a 2-way mixed-effects model based on absolute agreement and average measures [81].

ICC values range from 0 to 1, with values above 0.75 generally indicating good reliability and values above 0.90 considered excellent [82]. 95% CIs were computed for each ICC, and significance was determined using *F* tests. Retest reliability for the distractor interaction score (dichotomous variable) was examined using Cohen κ [83]. Retest reliability was evaluated for the full sample who completed session 2 (n=143) and for a subsample that reported no change in cognitive abilities since session 1 (123/143). Internal consistency between the 2 VKC-2 tasks (breakfast and lunch) was tested using the Spearman-Brown formula (*r*), with coefficients>0.70 interpreted as evidence of strong internal consistency [84].

Results

Participant Characteristics

A total of 237 participants were recruited from June 2021 to June 2025 for studies on everyday function. One participant with mild dementia refused to complete the study tasks; thus, the final analytic sample included 236 participants, of whom 172 were classified as having healthy cognition, 48 as having MCI, and 16 as having mild dementia. On average, participants were 72 years old and had completed 15 years of education; of the 236 participants, 156 (66.1%) were women, and nearly equal numbers identified as Black (n=106, 44.9%) and White (n=113, 47.9%). Demographic characteristics of the groups are reported in Table 2. The groups differed in age and education, but post hoc comparisons did not reach statistical significance (*P*>.051 for all). There were no group differences in estimated IQ or in the distributions of sex, Black/African American versus White race, or ethnicity.

Table 2. Demographic and descriptive characteristics by group.

Variable	Healthy (n=172)	Mild cognitive impairment (n=48)	Mild dementia (n=16)	*F* test (*df*) or chi-square (*df*)	*P* value
Age, mean (SD); range	71.95 (6.56); 58-94	74.54 (7.27); 61-98	74.50 (8.70); 55-91	3.30 (2, 235)	.04
Education (years), mean (SD); range	16.06 (2.51); 10-20	15.40 (3.25); 10-20	14.06 (3.04); 10-20	4.64 (2, 235)	.01
Estimated IQ, mean (SD); range	112.44 (11.73); 87-139	112.06 (13.30); 88-138	108.25 (13.19); 88-139	0.87 (2, 235)	.42
Sex: women, n (%)	114 (66.3)	33 (68.8)	9 (56.3)	0.85 (2)	.67
Race				3.53 (2)	.17
Black	72 (41.9)	26 (54.2)	8 (50.0)		
White	89 (51.7)	17 (35.4)	7 (43.8)		
Asian	5 (2.9)	3 (6.3)	1 (6.3)		
Pacific Islander/Hawaiian	2 (1.2)	0 (0)	0 (0)		
American Indian	0 (0)	1 (2.1)	0 (0)		
Multiracial	3 (1.7)	1 (2.1)	0 (0)		
Not reported	1 (0.6)	0 (0)	0 (0)		
Latino/Hispanic Ethnicity	2 (1.2)	0 (0)	0 (0)	0.75 (2)	.69
Past Experience Scale					
Familiarity rating, mean (SD); range	14.20 (2.44); 6-16	13.45 (3.27); 3-16	10.80 (4.44); 4-16	10.69 (2, 235)	<.001
Frequency rating, mean (SD); range	7.87 (2.82); 0-16	7.89 (3.27); 2-13	8.27 (1.62); 5-11	1.11 (2, 235)	.87

Results from the Past Experience Scale showed that task familiarity ratings were generally high, indicating that, on average, the breakfast and lunch tasks were “pretty” to “very” familiar. The groups differed on the familiarity rating, with post hoc tests indicating that the dementia group reported significantly lower task familiarity than the healthy cognition group (*P*<.001) and the MCI group (*P*=.005); however, the healthy cognition group and the MCI group did not differ (*P* =.32). According to the frequency ratings, participants reported that, on average, they had performed the VKC-2 tasks about once per month over the past 5-10 years. Frequency ratings did not differ across groups.

Demographic characteristics of participants who returned for session 2 and were included in the retest reliability analysis are reported in Table S1 in Multimedia Appendix 1. Compared with participants who did not return, the returning participants had completed significantly more years of education, obtained higher estimated IQ scores, and included a greater proportion of White participants.

Informant Characteristics

A total of 219 informants participated in the study. On average, informants were 63.97 years old (SD 13.99 years; range 20-90 years) and had completed 15.73 years of education (SD 2.43 years; range 10-21 years). Informants included spouses (95/219, 43.4%), children (60/219, 27.4%), friends (41/219, 18.7%), and other family members (23/219, 10.5%).

Correlations Among VKC-2 Scores and Principal Component Analysis

Average VKC-2 scores and their bivariate correlations indicate significant, moderate associations among all scores (Table 3). The relationship between VKC-2 time and touches was particularly strong, reflecting nearly overlapping scores, with more touches associated with longer completion times.

Table 3. VKC-2[a] scores and correlation coefficients in the full sample (N=236).

VKC-2 score	Digital dexterity	Time	Touches	%Off-screen
Time	0.67[b]	N/A[c]	N/A	N/A
Touches	0.51[b]	0.83[b]	N/A	N/A
%Off-screen	0.49[b]	0.42[b]	0.34[b]	N/A
Distractor interactions	0.34[b]	0.39[b]	0.42[b]	0.26[b]
Mean	86.76	197.47	67.82	0.48
SD	27.47	111.57	50.59	0.09

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]*P*<.001 (2-tailed). A total of 21 (8.9%) participants interacted with distractor objects; the mean and SD for the distractor interactions score are not reported because it was dichotomized.

[c]N/A: not applicable.

According to the Kaiser-Meyer-Olkin measure of sampling adequacy (0.579) and the significant Bartlett test of sphericity (χ^2_3=311.45, *P*<.001), there was a modest but acceptable level of shared variance among variables and a suitable correlation matrix for factor analysis. PCA results showed that only 1 component was extracted (eigenvalue=2.09), accounting for 69.74% of the total variance. All variables loaded positively on this component (time=0.930, touches=0.905, and %off-screen=0.639), suggesting a single underlying factor representing a common dimension. Thus, a VKC-2 composite score was computed by averaging sample-based *z* scores for time, touches, and %off-screen, with higher scores reflecting worse (ie, more inefficient) performance.

Construct Validity

The construct validity of the VKC-2 automated scores was evaluated by assessing differences among groups known to differ in functional abilities: healthy, MCI, and mild dementia. As shown in Figure 3, average scores on each VKC-2 measure were consistently worse for the dementia group. The same pattern was observed in the VKC-2 composite score (healthy: mean –0.22, SD 0.49; MCI: mean 0.43, SD 1.07; and dementia: mean 1.27, SD 1.36). Statistical analyses focused on differences between the healthy and MCI groups due to the relatively small number of participants with dementia. ANCOVA results comparing healthy versus participants with MCI are reported in Tables 4 and 5 and showed significant group differences (*P*<.001) in all measures after controlling for age. After controlling for the digital dexterity score and age (see Tables 4 and 5), the difference in the time score was no longer significant (*P*=.06), suggesting that the difference in completion time could be explained by low-level visuomotor skill differences between the MCI and healthy groups. By contrast, after controlling for digital dexterity and age, the differences in touches (*P*=.004), %off-screen (*P*=.01), and the VKC-2 (*P*<.001) composite score remained statistically significant, indicating that these between-group differences could not be explained by basic visuomotor skills. Thus, aside from time, the VKC-2 scores—particularly the composite score, which showed the strongest effect size after controlling for digital dexterity and age—likely reflect more than simple visuomotor abilities and capture the cognitive processes required to perform everyday tasks (ie, goal maintenance and control over task goals for the efficient execution of multistep everyday tasks).

Figure 3. Unadjusted VKC-2 mean scores by group.

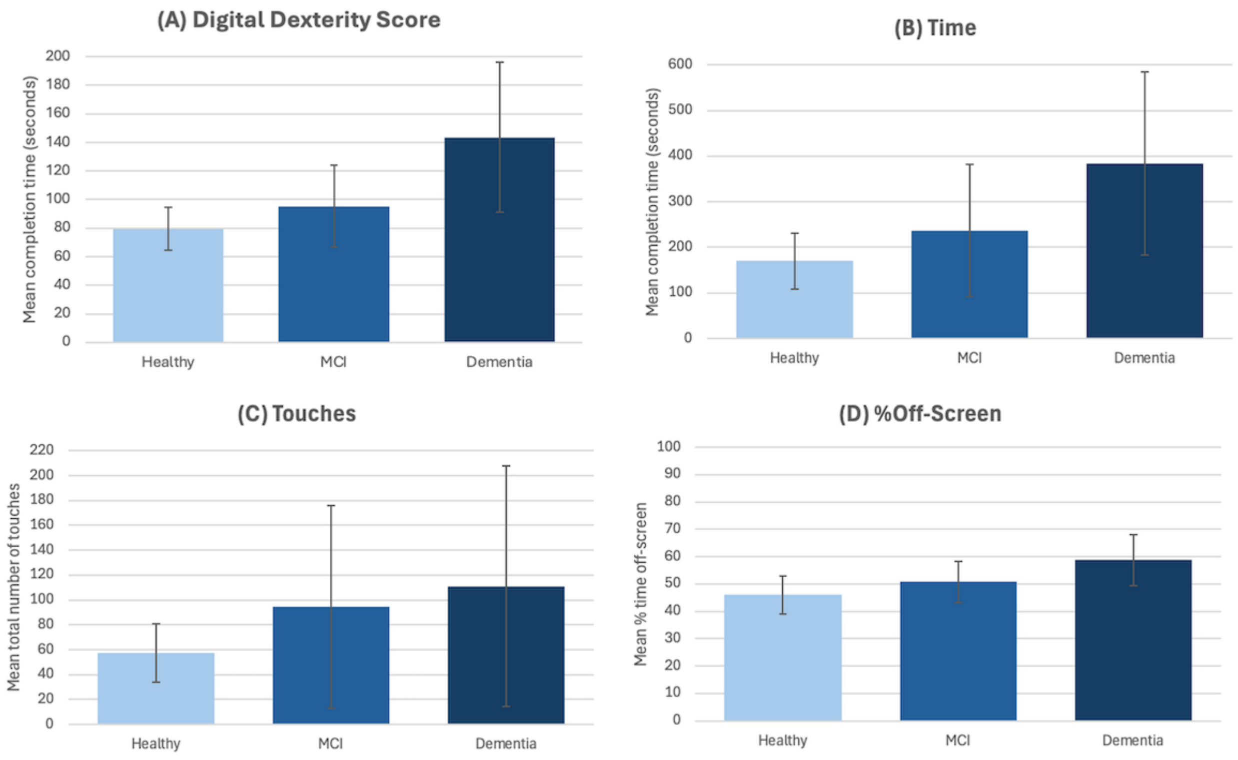


Table 4. Analysis of covariance results comparing participants with healthy cognition (n=172) versus those with mild cognitive impairment (n=48) on all VKC-2[a] automated scores: controlling for age.[b]

VKC-2 score	*F* test (*df*)	*P* value	η^2 (partial eta^2)	Effect size
Digital dexterity	20.68 (2, 219)	<.001	0.087	Medium to large
Time	17.54 (2, 219)	<.001	0.075	Medium
Touches	23.35 (2, 219)	<.001	0.097	Large
%Off-screen	12.71 (2, 219)	<.001	0.055	Medium
VKC-2 composite	29.70 (2, 219)	<.001	0.120	Medium to large

[a]VKC-2: Virtual Kitchen Challenge—Version 2.
[b]Effect sizes (η^2) are interpreted as follows: small=0.01, medium=0.06, and large=0.14.

Table 5. Analysis of covariance results comparing participants with healthy cognition (n=172) versus those with mild cognitive impairment (n=48) on all VKC-2[a] automated scores: controlling for age and digital dexterity.[b]

VKC-2 score	*F* test (*df*)	*P* value	η^2 (partial eta^2)	Effect size
Time	3.47 (3, 219)	.06	0.016	Small
Touches	8.60 (3, 219)	.004	0.038	Small to medium
%Off-screen	6.42 (3, 219)	.01	0.029	Small to medium
VKC-2 composite	11.68 (3, 219)	<.001	0.051	Medium

[a]VKC-2: Virtual Kitchen Challenge—Version 2.
[b]Effect sizes (η^2) are interpreted as follows: small=0.01, medium=0.06, and large=0.14.

The distributions of the distractor interaction score across the 3 groups (not reported in Figure 3) indicated a higher percentage of participants interacting with distractors in the groups with cognitive impairment (dementia: 5/16, 31%; MCI: 8/48, 17%; and healthy: 5/114, 4.3%). The difference in distractor interactions between the MCI and healthy groups was statistically significant (χ^2_1=8.03, *P*=.005; ϕ=0.191, small-to-medium effect size).

Classification Analyses

Classification analyses for distinguishing participants with healthy cognition from those with cognitive impairment (MCI

+ mild dementia combined) are reported in Table 6. All predictors showed statistically significant areas under the curve (AUCs; *P*<.001 for all), indicating that they were better than chance at predicting impaired group status. Time and the VKC-2 composite score were the strongest predictors, as indicated by their high AUCs and sensitivity. Time demonstrated particularly high sensitivity, making it useful for maximizing the identification of participants with impairment for early detection. By contrast, the %off-screen score showed the highest specificity, suggesting it may be more useful for ruling out individuals with healthy cognition during diagnostic confirmation. As expected, scores that increase sensitivity reduce specificity, reflecting the inherent trade-off between identifying true positives and minimizing false positives. Analyses distinguishing participants with healthy cognition versus those with MCI demonstrated similar AUCs (0.68-0.74), cutoff scores, and patterns of sensitivity and specificity, and are reported in Table S2 in Multimedia Appendix 1.

Table 6. Area under the curve values, optimal cutoffs, and specificity/sensitivity for predicting cognitive impairment from VKC-2[a] scores (N=236).

VKC-2 score	AUC[b]	95% CI	SE	*P* value	Optimal cutoff	Sensitivity	Specificity
Digital dexterity	0.73	0.65-0.81	0.41	<.001	87.12	0.67	0.77
Time	0.75	0.68-0.82	0.04	<.001	163.47	0.82	0.58
Touches	0.70	0.62-0.78	0.04	<.001	65.5	0.59	0.78
%Off-screen	0.71	0.64-0.79	0.04	<.001	0.53	0.48	0.87
VKC-2 composite	0.76	0.70-0.84	0.04	<.001	–0.041	0.69	0.72

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]AUC: area under the curve.

Convergent Validity Against Real Kitchen Scores

Bivariate correlations between VKC-2 scores and Real Kitchen scores are reported in Table 7. Correlations with Real Kitchen completion time, accomplishment, and total errors were consistently significant (*P*<.001 for all) and moderate to strong, supporting the convergent validity of the VKC-2 scores against the real versions of the VKC-2 tasks. Relations between VKC-2 measures and motor errors on the Real Kitchen were relatively weaker and not consistently significant (*P* values ranged from <.001 to .08), suggesting that VKC-2 scores correspond more strongly with the cognitive aspects of Real Kitchen performance rather than visuomotor errors made with the real tasks (Table 7). Spearman rank-order correlations showed the same pattern of results and are reported in Table S3 in Multimedia Appendix 1.

Table 7. Correlation coefficients (and P values) between VKC-2[a] scores and Real Kitchen scores (n=201).

VKC-2 score	Completion time	Accomplishment score	Total errors	Motor errors
Digital dexterity	0.59[b]	–0.58[b]	0.50[b]	0.13 (*P*=.08)
Time	0.58[b]	–0.53[b]	0.64[b]	0.22 (*P*=.004)
Touches	0.38[b]	–0.26[b]	0.53[b]	0.30[b]
%Off-screen	0.44[b]	–0.44[b]	0.40[b]	0.15 (*P*=.057)
VKC-2 composite	0.56[b]	–0.48[b]	0.62[b]	0.27[b]
Mean	244.35	32.09	7.42	2.37
SD	93.73	2.35	5.88	2.49

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]*P*<.001 (2-tailed).

Convergent Validity Against Conventional Cognitive Tests

Bivariate correlations between VKC-2 scores and demographically adjusted cognitive test scores are reported in Table 8. The coefficients were statistically significant (*P* values ranged from <.001 to .03), indicating that participants with higher cognitive test scores completed the VKC-2 tasks more quickly and efficiently, supporting the convergent validity of the VKC-2 scores. Spearman rank-order correlations showed the same pattern of results and are reported in Table S4 in Multimedia Appendix 1.

Table 8. Correlation coefficients (and P values) between VKC-2[a] scores and cognitive test scores (N=236).

VKC-2 score	Global Cognition mKnight-PACC[b]	Executive function composite	Episodic memory composite	Processing speed composite	Language composite
Digital dexterity	–0.50[c]	–0.29[c]	–0.42[c]	–0.41[c]	–0.36[c]
Time	–0.42[c]	–0.27[c]	–0.39[c]	–0.30[c]	–0.34[c]
Touches	–0.23[c]	–0.14 (*P*=.03)	–0.29[c]	–0.11 (*P*=.10)	–0.22[c]
%Off-screen	–0.38[c]	–0.27[c]	–0.40[c]	–0.26[c]	–0.27[c]
VKC-2 composite	–0.42[c]	–0.27[c]	–0.42[c]	–0.28[c]	–0.34[c]
Mean	50.50	49.15	45.04	53.08	47.97
SD	7.66	8.77	10.04	9.75	9.55

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]mKnight-PACC: modified Knight-Preclinical Alzheimer Cognitive Composite.

[c]*P*<.001 (2-tailed).

Convergent Validity Against Self/Informant Questionnaires of Everyday Function

Table 9 shows the relationships between VKC-2 scores and questionnaires assessing everyday function completed by participants and informants. Results from participant questionnaires indicated that the associations between VKC-2 scores and the IADL-C and FAQ, which assess current functional abilities, were statistically significant (*P* values ranged from <.001 to .02) and in the expected direction. That is, participants who reported greater current functional difficulties (IADL-C and FAQ) also performed the VKC-2 tasks less quickly and efficiently. The relationship between VKC-2 scores and participants' reports of functional decline (ECog-12) was not significant (*P* values ranged from .21 to .55). By contrast, informant reports of both current functional difficulties (IADL-C and FAQ) and functional decline (ECog-12) were significantly associated with lower VKC-2 scores. Overall, correlations between the VKC-2 and participant/informant questionnaires support the validity of the VKC-2 and are comparable to or stronger than the relationships reported between conventional performance-based tests and questionnaires in the literature [43]. Spearman rank-order correlations showed a similar pattern of results and are reported in Table S5 in Multimedia Appendix 1.

Table 9. Correlation coefficients (and P values) between VKC-2[a] scores and questionnaires.

VKC-2 score	Participant questionnaires (n=236)			Informant questionnaires (n=194)		
	IADL-C[b]	FAQ[c]	ECog-12[d]	IADL-C	FAQ	ECog-12
Digital dexterity	0.26[e]	0.32[e]	0.08 (*P*=.26)	0.43[e]	0.32[e]	0.34[e]
Time	0.28[e]	0.28[e]	0.06 (*P*=.38)	0.41[e]	0.32[e]	0.26[e]
Touches	0.15 (*P*=.03)	0.20 (*P*=.003)	0.04 (*P*=.55)	0.20 (*P*=.005)	0.10 (*P*=.17)	0.07 (*P*=.35)
%Off-screen	0.26[e]	0.25[e]	0.09 (*P*=.21)	0.35[e]	0.29[e]	0.31[e]
VKC-2 composite	0.27[e]	0.30[e]	0.07 (*P*=.30)	0.37[e]	0.28[e]	0.24[e]
Mean	44.76	2.12	1.56	46.75	2.92	1.40
SD	21.41	3.85	.97	30.10	5.98	.54

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]IADL-C: Instrumental Activities of Daily Living—Compensation.

[c]FAQ: Functional Activity Questionnaire.

[d]ECog-12: 12-item Everyday Cognition Scale.

[e]*P*<.001 (2-tailed).

Retest Reliability

ICCs are reported in Table 10 and indicate moderate to excellent reliability for the VKC-2 automated scores. Cohen κ, used to assess agreement between distractor interaction scores at time 1 and time 2, showed only fair agreement (κ=0.27, *P*<.001), indicating limited but statistically significant consistency over time. When ICCs were rerun, including only participants who reported no change in their cognitive status from session 1 (123/143, 86%), results yielded comparable or slightly improved coefficients relative to the full sample (see Table S6 in Multimedia Appendix 1).

Table 10. Intraclass correlation coefficients for VKC-2[a] scores over time (n=143).

VKC-2 score	Intraclass correlation coefficient[b] (average measures)	95% CI	*F* test (*df*)	*P* value
Digital dexterity	0.844	0.766-0.893	6.965 (142, 142)	<.001
Time	0.812	0.736-0.865	5.469 (142, 142)	<.001
Touches	0.849	0.783-0.893	6.943 (140, 140)	<.001
%Off-screen	0.703	0.523-0.806	3.837 (140, 140)	<.001
VKC-2 composite	0.899	0.860-0.923	9.965 (140, 140)	<.001

[a]VKC-2: Virtual Kitchen Challenge—Version 2.

[b]Type A using an absolute agreement definition.

Internal Reliability

Internal consistency between the VKC-2 breakfast and lunch tasks at time 1 was evaluated using Spearman-Brown coefficients in the full sample (N=236). Results indicated acceptable to good internal consistency for all scores (time: 0.81; touches: 0.81; %off-screen: 0.77; and VKC-2 composite score: 0.84).

Discussion

Results of this study support the validity and reliability of the VKC-2 automated scores as measures of everyday function in older adults. As predicted, VKC-2 scores differed significantly between groups known to vary in functional ability (healthy vs MCI vs mild dementia), supporting the construct validity of the VKC-2. Convergent validity was further supported by significant correlations between VKC-2 scores and performance on the real versions of the VKC-2 tasks (Real Kitchen), conventional cognitive test scores, and self/informant questionnaires assessing everyday functioning. Retest reliability analyses showed fair to excellent reliability for the VKC-2 automated scores over 4-6 weeks. Internal consistency between the 2 VKC-2 tasks (breakfast and lunch) was also good. Additionally, participants reported that the tasks included in the VKC-2 were highly familiar (Past Experience Questionnaire). These findings suggest that the VKC-2 automated scores hold strong potential for addressing critical gaps in functional assessment across multiple contexts, including screening older adults at risk for decline in meaningful everyday activities in primary care and serving as a functional end point in clinical trials of Alzheimer disease/Alzheimer disease–related disorder treatments.

To our knowledge, this is the first study to demonstrate significant differences between older adults with healthy cognition and those with MCI on the VKC-2 automated scores. Group differences in all scores except time persisted even after controlling for the digital dexterity score, a novel feature of the updated VKC-2. Thus, differences between MCI and healthy participants on the VKC-2 cannot be attributed solely to differences in digital visuomotor skills or touchscreen accuracy, but rather reflect the additional cognitive demands required to perform everyday tasks accurately and efficiently (eg, accurate object selection, sequencing of task steps, performance monitoring [74,85,86]). This conclusion is further supported by significant correlations with Real Kitchen scores (see also [57,87]) and by the fact that differences between participants with MCI and healthy participants on the VKC-2 mirror those observed on performance-based tasks with real objects in previous studies [26,32,33,35,45]. Significant associations with cognitive tests of episodic memory and language, which do not primarily measure motor skills or processing speed, provide additional evidence that the automated VKC-2 scores reflect cognitive abilities. Collectively, these results strongly support the construct validity of the VKC-2, offering a novel approach to identify everyday task difficulties without the need for video recording or trained coders—a major advantage over traditional performance-based tests—providing a highly efficient, scalable, and sensitive measure of everyday functioning.

It is important to acknowledge that some VKC-2 scores reflect visuomotor skills more than others. For example, the completion time (time) score did not remain significantly different between participants with MCI and those with healthy cognition after controlling for the digital dexterity score. This should not be viewed as a limitation, as mild upper motor dexterity difficulties contribute to functional impairments in people with MCI [20], and mild upper and lower limb difficulties are significantly associated with cognitive challenges in older adults without MCI [21,22,88]. Indeed, the VKC-2 digital dexterity score, as well as VKC-2 measures of efficiency, were associated with a measure of global cognitive abilities (mKnight-PACC) that is sensitive to preclinical Alzheimer disease. Thus, mild motor difficulties may serve as important early indicators of Alzheimer disease/Alzheimer disease–related disorder risk that could be missed by conventional cognitive tests. Additional studies, including longitudinal follow-up, are needed to identify the optimal combination of VKC-2 scores to maximize early detection of functional difficulties and risk.

Correlation analyses with conventional self- and informant-report questionnaires of everyday function provided additional support for the validity of the VKC-2 automated measures as indicators of processes that influence real-world functioning. The strength and pattern of correlations between VKC-2 scores and conventional questionnaires were similar to those reported for validated performance-based tests and questionnaires of everyday function in the existing literature [89,90]. Correlations were stronger and more consistent with informant reports than with self-reports, particularly for the questionnaire assessing cognitive/functional decline (ECog-12 [13]); this pattern has been reported in previous studies [56] and aligns with our conceptualization of the constructs measured by performance-based tests versus questionnaires. We view the VKC-2, like other performance-based tests, as a measure of everyday functional capacity, making it well-suited for between-participant comparisons, staging, and tracking change over time. Questionnaires assess real-world functioning, which is highly unconstrained, with task demands, motivation, economic resources, social support, and other factors varying widely. Thus, in clinical practice, the VKC-2 could be used alongside questionnaires to provide a comprehensive evaluation of everyday function across contexts.

Significant associations between VKC-2 scores and conventional questionnaires of everyday function support the clinical relevance of the VKC-2 measures. Differences between participants with healthy cognition and those with MCI were small to moderate, with absolute differences amounting to only a few seconds on some scores. Such differences may reflect subtle processing difficulties that lead to inefficiency and increased cognitive load, which could accumulate over the course of a day. We acknowledge, however, that direct evidence that the mild cognitive difficulties captured by the VKC-2 translate to meaningful impacts on everyday tasks is currently lacking. Further validation using ecological momentary assessment or digital phenotyping via wearables (or both) would provide more direct evidence of the VKC-2 as a measure of real-world everyday function.

In addition to validity analyses, the reliability of the VKC-2 represents an important novel contribution of this study. To our knowledge, reliability has not been examined for any prior version of the Virtual Kitchen. Retest reliability estimates (ICC) showed that the automated VKC-2 scores—except for distractor interactions, which occurred very infrequently—were highly stable over a 4-6-week period. ICCs were even stronger when participants who reported notable changes in cognitive abilities were excluded. Strong retest reliability is critical for using the VKC-2 to evaluate meaningful change over time and for clinical trial applications. The VKC-2 tasks (breakfast and lunch) also demonstrated strong internal consistency, supporting the coherence of the combined total VKC-2 scores. Furthermore, correlations and PCA indicate that VKC-2 automated scores reflect a single underlying dimension and can be combined into a composite score representing task efficiency.

Several strengths of the study are worth noting. First, the sample size and inclusion of a substantial proportion of participants (106/236, 44.9%) identifying as Black or African American addresses a critical gap in cognitive assessment research and enhances the generalizability of our findings across the US population. Second, the VKC-2's portability, automated scoring, and standardized administration protocol offer clear advantages over current functional measures and existing regulatory-approved outcome measures for clinical trials, which often require specialized training, lengthy administration times, and access to informants. The VKC-2 does not require Wi-Fi and can be administered on any commercially available, budget-friendly touchscreen computer. Finally, the efficiency of the VKC-2 compared with conventional cognitive test batteries makes it particularly suitable for busy clinical settings where comprehensive neuropsychological assessments are impractical.

Study limitations also warrant consideration. First, although our sample included substantial racial diversity, the predominance of highly educated participants (mean 15.7 years of education) may limit generalizability to populations with lower educational attainment. Additionally, the sample was majority female (156/236, 66.1%), and participants from racial groups other than Black/African American or White or diverse ethnicities were underrepresented. Second, the community-based sample primarily included older adults with healthy cognition, with only 64 of 236 (27.1%) participants meeting criteria for cognitive impairment. The imbalance in subgroup sizes between participants with healthy versus those with impaired cognition limited statistical power for between-group comparisons and AUC/classification analyses. Therefore, additional studies are needed to replicate these findings in samples with larger groups of individuals with MCI or mild dementia. Third, although the virtual task environment is ecologically valid, it may not capture all real-world functional demands, such as physical fatigue, environmental distractions, or competing task requirements. Fourth, the cross-sectional design limits conclusions about the VKC-2's ability to detect meaningful change over time or predict clinical outcomes (predictive validity). Finally, direct validation against regulatory outcome measures is necessary before the VKC-2 can be considered an alternative end point in clinical trials.

As noted, future research on the VKC-2 should include longitudinal studies to determine the predictive validity of its scores. It will be important to evaluate whether the VKC-2 outperforms conventional measures in identifying individuals at risk for cognitive and functional decline. However, even if the VKC-2 performs comparably to traditional cognitive tests or questionnaires, it offers important advantages, including greater efficiency and independence from the need for a reliable informant. Another important future direction is validation of the VKC-2 against biomarkers of neurodegenerative disease. Holmqvist and colleagues [48] demonstrated strong correlations between VKC-2 scores and magnetic resonance imaging–derived measures of cerebral white matter hyperintensities, a biomarker of small vessel disease associated with brain aging and neurodegeneration. Ongoing studies are examining associations between VKC-2 scores and additional biomarkers, including Alzheimer disease–specific positron emission tomography and blood markers. Finally, automated VKC-2 scores that capture task accomplishment are under development, which will further enhance the utility of the VKC-2 by providing a detailed characterization of everyday task performance patterns [85]. Future implementation research should examine and address potential barriers to VKC-2 adoption, including variability in technology literacy, digital skill levels, and computer-related anxiety among diverse older adults [91], as well as strategies for seamless integration into existing clinical workflows.

In conclusion, there is growing interest in the development of digital assessments of cognition, including digitized versions of traditional cognitive tests, smartphone- and tablet-based cognitive assessments, and VR [92]. Digital, performance-based assessments of everyday tasks, such as the VKC-2, extend this trend to meaningful measures of everyday function. The VKC was designed to address weaknesses of conventional functional measures by providing an objective, standardized, and highly efficient assessment that does not rely on informant reports. The VKC-2 requires approximately 15-20 minutes to administer, is suitable for the full spectrum of cognitive aging—from healthy aging to mild dementia—and includes tasks (breakfast and lunch) that have been extensively studied and shown to be highly familiar to older adults [27,28,32,47,93]. The VKC-2 can be administered on a portable laptop without the need for additional objects or supplies, including a VR headset, avoiding limitations associated with cybersickness and confusion. The touchscreen interface provides a more natural interaction than a mouse or joystick for older adults [94]. Finally, the VKC-2 provides sensitive and detailed performance analysis, including time to completion and measures of performance efficiency derived from the touchscreen, eliminating the need for video recording and human coders. Older adult participants in this study were able to use the touchscreen interface, understood the instructions, did not require extensive training, and performed the tasks consistently with expectations based on data from the Real Kitchen [35]. The VKC-2 shows strong potential as an ecologically valid and scalable tool for capturing everyday functional capabilities in people with healthy cognition, MCI, and mild dementia across various settings, including large longitudinal studies, health clinics, and clinical trials.

Acknowledgments

The authors acknowledge Anna Callahan, Julina Hossfeld, Nicole Lloyd, Julia Manganti, Shrey Patel, Emma Pinsky, and Yuki Tsuchiya for their contributions to the Virtual Kitchen Challenge Project, including participant recruitment, assistance with data collection, and database management. The authors also thank Ms. Sherry Hill, Director of the Community Engagement Fitness Network, for her support with community outreach and participant recruitment. Generative artificial intelligence was not used in the creation of this manuscript.

Funding

This study was funded by grants from the National Institute on Aging (R01AG062503 to TG; F31AG089944 to SH).

Data Availability

The dataset analyzed for this study will be made available in the Virtual Kitchen Challenge Repository. In the interim, data requests may be sent to the corresponding author (TG).

Authors' Contributions

Conceptualization: TG (lead), TY (equal)
Data curation: TG (lead), MK (equal), MM (equal)
Formal analysis: TG (lead), DAGD (equal)
Funding acquisition: TG
Investigation: TG, MK, MM, SMS, MBT, K Hackett, SH, REM, K Halberstadter, RC, MR, GV, MDS
Methodology: TG (lead), TY (equal)
Project administration: TG (lead), MK (equal), MM (equal)
Resources: TG
Supervision: TG
Validation: TG
Visualization: TG (lead), MK (supporting)
Writing – original draft: TG (lead), MK (supporting)
Writing – review & editing: TG (lead), MK (equal), MDS (supporting), SMS (supporting), MBT (supporting), REM (supporting), K Hackett (supporting), K Halberstadter (supporting), DAGD (supporting), MR (supporting), SH (supporting), TY (supporting)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Additional analysis.
[DOCX File , 36 KB - aging_v9i1e82092_app1.docx]

References

1. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Behavioral and Social Sciences and Education , Board on Health Sciences Policy, Board on Behavioral, Cognitive, and Sensory Sciences, Committee on

Research Priorities for Preventing and Treating Alzheimer's Disease and Related Dementias. In: Yost OC, Downey A, Powell T, editors. Preventing and Treating Dementia: Research Priorities to Accelerate Progress. United States: National Academies Press; 2025. URL: https://doi.org/10.17226/28588
2. Pérès K, Chrysostome V, Fabrigoule C, Orgogozo JM, Dartigues JF, Barberger-Gateau P. Restriction in complex activities of daily living in MCI: impact on outcome. Neurology 2006 Aug 08;67(3):461-466. [doi: 10.1212/01.wnl.0000228228.70065.f1] [Medline: 16894108]
3. Pérès K, Helmer C, Amieva H, Orgogozo J, Rouch I, Dartigues J, et al. Natural history of decline in instrumental activities of daily living performance over the 10 years preceding the clinical diagnosis of dementia: a prospective population-based study. J Am Geriatr Soc 2008 Jan 20;56(1):37-44. [doi: 10.1111/j.1532-5415.2007.01499.x] [Medline: 18028344]
4. Lau KM, Parikh M, Harvey DJ, Huang C, Farias ST. Early cognitively based functional limitations predict loss of independence in instrumental activities of daily living in older adults. J Int Neuropsychol Soc 2015 Oct;21(9):688-698 [FREE Full text] [doi: 10.1017/S1355617715000818] [Medline: 26391766]
5. Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE, Alzheimer's Disease Neuroimaging Initiative. Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to Alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative. Arch Gen Psychiatry 2011 Sep 05;68(9):961-969. [doi: 10.1001/archgenpsychiatry.2011.96] [Medline: 21893661]
6. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004 Sep;256(3):183-194 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01388.x] [Medline: 15324362]
7. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May 22;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
8. Edgar CJ, Vradenburg G, Hassenstab J. The 2018 revised FDA guidance for early Alzheimer's disease: establishing the meaningfulness of treatment effects. J Prev Alzheimers Dis 2019;6(4):223-227. [doi: 10.14283/jpad.2019.30] [Medline: 31686092]
9. Mintun M, Ritchie CW, Solomon P, Sims JR, Salloway S, Hansson O, et al. Donanemab in early symptomatic Alzheimer's disease: efficacy and safety in TRAILBLAZER‐ALZ 2, a phase 3 randomized clinical trial. Alzheimer's & Dementia 2023 Dec 25;19(S24):e082733. [doi: 10.1002/alz.082733]
10. Rabinovici GD, La Joie R. Amyloid-targeting monoclonal antibodies for Alzheimer disease. JAMA 2023 Aug 08;330(6):507-509. [doi: 10.1001/jama.2023.11703] [Medline: 37459124]
11. Harvey PD, Cosentino S, Curiel R, Goldberg TE, Kaye J, Loewenstein D, et al. Performance-based and observational assessments in clinical trials across the Alzheimer's disease spectrum. Innov Clin Neurosci 2017;14(1-2):30-39 [FREE Full text] [Medline: 28386519]
12. Stoeckel LE, Fazio EM, Hardy KK, Kidwiler N, McLinden KA, Williams B. Clinically meaningful outcomes in Alzheimer's disease and Alzheimer's disease related dementias trials. Alzheimers Dement (N Y) 2025 Feb 19;11(1):e70058. [doi: 10.1002/trc2.70058] [Medline: 39975465]
13. Farias ST, Mungas D, Reed BR, Cahn-Weiner D, Jagust W, Baynes K, et al. The measurement of everyday cognition (ECog): scale development and psychometric properties. Neuropsychology 2008 Jul;22(4):531-544 [FREE Full text] [doi: 10.1037/0894-4105.22.4.531] [Medline: 18590364]
14. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982 May 01;37(3):323-329. [doi: 10.1093/geronj/37.3.323] [Medline: 7069156]
15. Schmitter-Edgecombe M, Parsey C, Lamb R. Development and psychometric properties of the instrumental activities of daily living: compensation scale. Arch Clin Neuropsychol 2014 Dec 25;29(8):776-792 [FREE Full text] [doi: 10.1093/arclin/acu053] [Medline: 25344901]
16. Hackett K, Mis R, Drabick DA, Giovannetti T. Informant reporting in mild cognitive impairment: sources of discrepancy on the Functional Activities Questionnaire. J Int Neuropsychol Soc 2020 Jan 22;26(5):503-514. [doi: 10.1017/s1355617719001449]
17. Wadley VG, Harrell LE, Marson DC. Self- and informant report of financial abilities in patients with Alzheimer's disease: reliable and valid? J Am Geriatr Soc 2003 Nov 24;51(11):1621-1626. [doi: 10.1046/j.1532-5415.2003.51514.x] [Medline: 14687393]
18. Okonkwo OC, Wadley VG, Griffith HR, Belue K, Lanza S, Zamrini EY, et al. Awareness of deficits in financial abilities in patients with mild cognitive impairment: going beyond self-informant discrepancy. The American Journal of Geriatric Psychiatry 2008 Aug;16(8):650-659. [doi: 10.1097/jgp.0b013e31817e8a9d]
19. Tassoni MB, Drabick DAG, Giovannetti T. The frequency of self-reported memory failures is influenced by everyday context across the lifespan: Implications for neuropsychology research and clinical practice. Clin Neuropsychol 2023 Aug;37(6):1115-1135. [doi: 10.1080/13854046.2022.2112297] [Medline: 36000515]
20. Rycroft SS, Quach LT, Ward RE, Pedersen MM, Grande L, Bean JF. The relationship between cognitive impairment and upper extremity function in older primary care patients. J Gerontol A Biol Sci Med Sci 2019 Mar 14;74(4):568-574 [FREE Full text] [doi: 10.1093/gerona/gly246] [Medline: 30358815]

21. Hesseberg K, Tangen G, Pripp A, Bergland A. Associations between cognition and hand function in older people diagnosed with mild cognitive impairment or dementia. Dement Geriatr Cogn Dis Extra 2020 Dec 15;10(3):195-204 [FREE Full text] [doi: 10.1159/000510382] [Medline: 33569075]
22. Teruya SL, Dimino C, Silverman KD, Mielenz T. Poor lower extremity functioning is associated with modest increased incidence of probable dementia. Geriatrics (Basel) 2021 Aug 10;6(3):77 [FREE Full text] [doi: 10.3390/geriatrics6030077] [Medline: 34449642]
23. Giovannetti T, Seligman SC, Britnell P, Brennan L, Libon DJ. Differential effects of goal cues on everyday action errors in Alzheimer's disease versus Parkinson's disease dementia. Neuropsychology 2015 Jul;29(4):592-602. [doi: 10.1037/neu0000167] [Medline: 25495833]
24. Brennan L, Giovannetti T, Libon DJ, Bettcher BM, Duey K. The impact of goal cues on everyday action performance in dementia. Neuropsychol Rehabil 2009 Aug;19(4):562-582. [doi: 10.1080/09602010802405623] [Medline: 18923960]
25. Schwartz MF, Segal M, Veramonti T, Ferraro M, Buxbaum LJ. The Naturalistic Action Test: a standardised assessment for everyday action impairment. Neuropsychological Rehabilitation 2010 Oct 22;12(4):311-339. [doi: 10.1080/09602010244000084]
26. Bruce I, Ntlholang O, Crosby L, Cunningham C, Lawlor B. The clinical utility of naturalistic action test in differentiating mild cognitive impairment from early dementia in memory clinic. Int J Geriatr Psychiatry 2016 Mar 11;31(3):309-315. [doi: 10.1002/gps.4331] [Medline: 26264127]
27. Giovannetti T, Libon DJ, Buxbaum LJ, Schwartz MF. Naturalistic action impairments in dementia. Neuropsychologia 2002 Jan;40(8):1220-1232. [doi: 10.1016/s0028-3932(01)00229-9] [Medline: 11931925]
28. Giovannetti T, Bettcher BM, Brennan L, Libron DJ, Kessler RK, Duey K. Coffee with jelly or unbuttered toast: commissions and omissions are dissociable aspects of everyday action impairment in Alzheimer's disease. Neuropsychology 2008 Mar;22(2):235-245. [doi: 10.1037/0894-4105.22.2.235] [Medline: 18331166]
29. Giovannetti T, Bettcher BM, Libon DJ, Brennan L, Sestito N, Kessler RK. Environmental adaptations improve everyday action performance in Alzheimer's disease: empirical support from performance-based assessment. Neuropsychology 2007 Jul;21(4):448-457. [doi: 10.1037/0894-4105.21.4.448] [Medline: 17605578]
30. Giovannetti T, Libon DJ, Hart T. Awareness of naturalistic action errors in dementia. J Int Neuropsychol Soc 2002 Jul 14;8(5):633-644. [doi: 10.1017/s135561770280131x] [Medline: 12164673]
31. Giovannetti T, Schmidt KS, Gallo JL, Sestito N, Libon DJ. Everyday action in dementia: evidence for differential deficits in Alzheimer's disease versus subcortical vascular dementia. J Int Neuropsychol Soc 2006 Jan 23;12(1):45-53. [doi: 10.1017/S1355617706060012] [Medline: 16433943]
32. Giovannetti T, Bettcher BM, Brennan L, Libon DJ, Burke M, Duey K, et al. Characterization of everyday functioning in mild cognitive impairment: a direct assessment approach. Dement Geriatr Cogn Disord 2008 Mar 14;25(4):359-365. [doi: 10.1159/000121005] [Medline: 18340108]
33. Cornelis E, Gorus E, Van Weverbergh K, Beyer I, De Vriendt P. Convergent and concurrent validity of a report- versus performance-based evaluation of everyday functioning in the diagnosis of cognitive disorders in a geriatric population. Int Psychogeriatr 2018 Dec;30(12):1837-1848 [FREE Full text] [doi: 10.1017/S1041610218000327] [Medline: 29564999]
34. Bettcher BM, Giovannetti T, Libon DJ, Eppig J, Wambach D, Klobusicky E. Improving everyday error detection, one picture at a time: a performance-based study of everyday task training. Neuropsychology 2011 Nov;25(6):771-783. [doi: 10.1037/a0024107] [Medline: 21639639]
35. McKniff M, Holmqvist S, Kaplan M, Simone SM, Tassoni MB, Mis RE, et al. Subtle inefficiencies in everyday tasks indicate early functional difficulties in older adults: implications for clinical practice and research. Clin Neuropsychol 2025 Apr 30:1-20. [doi: 10.1080/13854046.2025.2497381] [Medline: 40304615]
36. Rycroft SS, Giovannetti T, Divers R, Hulswit J. Sensitive performance-based assessment of everyday action in older and younger adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2018 Mar 07;25(2):259-276. [doi: 10.1080/13825585.2017.1287855] [Medline: 28270012]
37. Seligman SC, Giovannetti T, Sestito J, Libon DJ. A new approach to the characterization of subtle errors in everyday action: implications for mild cognitive impairment. Clin Neuropsychol 2014 Nov 05;28(1):97-115. [doi: 10.1080/13854046.2013.852624] [Medline: 24191855]
38. Divers R, Ham L, Matchanova A, Hackett K, Mis R, Howard K, et al. When and how did you go wrong? Characterizing mild functional difficulties in older adults during an everyday task. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2021 Mar 30;28(2):308-326 [FREE Full text] [doi: 10.1080/13825585.2020.1756210] [Medline: 32352347]
39. Giovannetti T, Yamaguchi T, Roll E, Harada T, Rycroft SS, Divers R, et al. The Virtual Kitchen Challenge: preliminary data from a novel virtual reality test of mild difficulties in everyday functioning. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2019 Nov 28;26(6):823-841. [doi: 10.1080/13825585.2018.1536774] [Medline: 30370825]
40. Diehl M, Willis SL, Schaie KW. Everyday problem solving in older adults: observational assessment and cognitive correlates. Psychol Aging 1995 Sep;10(3):478-491. [doi: 10.1037//0882-7974.10.3.478] [Medline: 8527068]
41. McAlister C, Schmitter-Edgecombe M. Naturalistic assessment of executive function and everyday multitasking in healthy older adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2013 Nov;20(6):735-756 [FREE Full text] [doi: 10.1080/13825585.2013.781990] [Medline: 23557096]

42. Sanders C, Schmitter-Edgecombe M. Identifying the nature of impairment in planning ability with normal aging. J Clin Exp Neuropsychol 2012 Aug;34(7):724-737. [doi: 10.1080/13803395.2012.670210] [Medline: 22506736]
43. Schmitter-Edgecombe M, Parsey C, Cook DJ. Cognitive correlates of functional performance in older adults: comparison of self-report, direct observation, and performance-based measures. J Int Neuropsychol Soc 2011 Jun 27;17(5):853-864. [doi: 10.1017/s1355617711000865]
44. Schmitter-Edgecombe M, McAlister C, Weakley A. Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: the day-out task. Neuropsychology 2012 Sep;26(5):631-641 [FREE Full text] [doi: 10.1037/a0029352] [Medline: 22846035]
45. Griffith H, Belue K, Sicola A, Krzywanski S, Zamrini E, Harrell L, et al. Impaired financial abilities in mild cognitive impairment: a direct assessment approach. Neurology 2003 Feb 11;60(3):449-457. [doi: 10.1212/wnl.60.3.449] [Medline: 12578926]
46. Gold DA, Park NW, Murphy KJ, Troyer AK. Naturalistic action performance distinguishes amnestic mild cognitive impairment from healthy aging. J Int Neuropsychol Soc 2015 Jul 06;21(6):419-428. [doi: 10.1017/s135561771500048x]
47. Giovannetti T, Britnell P, Brennan L, Siderowf A, Grossman M, Libon DJ, et al. Everyday action impairment in Parkinson's disease dementia. J Int Neuropsychol Soc 2012 May 24;18(5):787-798. [doi: 10.1017/s135561771200046x]
48. Holmqvist SL, Jobson K, Desalme D, Simone SM, Tassoni M, McKniff M, et al. Preliminary validation of the Virtual Kitchen Challenge as an objective and sensitive measure of everyday function associated with cerebrovascular disease. Alzheimers Dement (Amst) 2024 Feb 05;16(1):e12547 [FREE Full text] [doi: 10.1002/dad2.12547] [Medline: 38318469]
49. Giovannetti T, Schwartz MF, Buxbaum LJ. The Coffee Challenge: a new method for the study of everyday action errors. J Clin Exp Neuropsychol 2007 Oct 24;29(7):690-705. [doi: 10.1080/13803390600932286] [Medline: 17891679]
50. Kessler RK, Giovannetti T, MacMullen LR. Everyday action in schizophrenia: performance patterns and underlying cognitive mechanisms. Neuropsychology 2007 Jul;21(4):439-447. [doi: 10.1037/0894-4105.21.4.439] [Medline: 17605577]
51. Devlin KN, Giovannetti T, Kessler RK, Fanning MJ. Commissions and omissions are dissociable aspects of everyday action impairment in schizophrenia. J Int Neuropsychol Soc 2014 Jul 30;20(8):812-821. [doi: 10.1017/s1355617714000654]
52. Mis RE, Ando T, Yamaguchi T, Brough C, Michalski L, Hoffman LJ, et al. Assessing everyday action in young adult athletes using the Virtual Kitchen Challenge: relations with conventional cognitive tests. J Int Neuropsychol Soc 2025 Sep 01;31(5-6):423-429. [doi: 10.1017/s135561772510101x]
53. Loewenstein DA, Amigo E, Duara R, Guterman A, Hurwitz D, Berkowitz N, et al. A new scale for the assessment of functional status in Alzheimer's disease and related disorders. J Gerontol 1989 Jul 01;44(4):P114-P121. [doi: 10.1093/geronj/44.4.p114] [Medline: 2738312]
54. Weiner MF, Gehrmann HR, Hynan LS, Saine KC, Cullum CM. Comparison of the test of everyday functional abilities with a direct measure of daily function. Dement Geriatr Cogn Disord 2006 Jun 19;22(1):83-86. [doi: 10.1159/000093388] [Medline: 16710087]
55. Binegar D, Hynan L, Lacritz L, Weiner M, Cullum C. Can a direct IADL measure detect deficits in persons with MCI? Curr Alzheimer Res 2009 Feb 01;6(1):48-51 [FREE Full text] [doi: 10.2174/156720509787313880] [Medline: 19199874]
56. Goldberg TE, Koppel J, Keehlisen L, Christen E, Dreses-Werringloer U, Conejero-Goldberg C, et al. Performance-based measures of everyday function in mild cognitive impairment. Am J Psychiatry 2010 Jul 01;167(7):845-853 [FREE Full text] [doi: 10.1176/appi.ajp.2010.09050692] [Medline: 20360320]
57. Allain P, Foloppe DA, Besnard J, Yamaguchi T, Etcharry-Bouyx F, Le Gall D, et al. Detecting everyday action deficits in Alzheimer's disease using a nonimmersive Virtual Reality Kitchen. J Int Neuropsychol Soc 2014 Mar 10;20(5):468-477. [doi: 10.1017/s1355617714000344]
58. Park MS, Kang KJ, Jang SJ, Lee JY, Chang SJ. Evaluating test-retest reliability in patient-reported outcome measures for older people: a systematic review. Int J Nurs Stud 2018 Mar;79:58-69. [doi: 10.1016/j.ijnurstu.2017.11.003] [Medline: 29178977]
59. Jak AJ, Bondi MW, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon DP, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. The American Journal of Geriatric Psychiatry 2009 May;17(5):368-375. [doi: 10.1097/jgp.0b013e31819431d5]
60. Bondi MW, Edmonds EC, Jak AJ, Clark LR, Delano-Wood L, McDonald CR, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. JAD 2014 Aug 11;42(1):275-289. [doi: 10.3233/jad-140276]
61. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May 22;7(3):263-269 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
62. Schretlen D, Testa S, Pearlson GD. Calibrated Neuropsychological Normative System Professional Manual. Lutz, FL: Psychological Assessment Resources; 2010.
63. Manly JJ, Byrd DA, Touradji P, Stern Y. Acculturation, reading level, and neuropsychological test performance among African American elders. Appl Neuropsychol 2004 Mar;11(1):37-46. [doi: 10.1207/s15324826an1101_5] [Medline: 15471745]

64. Manly JJ, Jacobs DM, Touradji P, Small SA, Stern Y. Reading level attenuates differences in neuropsychological test performance between African American and White elders. J Int Neuropsychol Soc 2002 Mar 16;8(3):341-348. [doi: 10.1017/s1355617702813157] [Medline: 11939693]
65. Schretlen DJ, Winicki JM, Meyer SM, Testa SM, Pearlson GD, Gordon B. Development, psychometric properties, and validity of the Hopkins Adult Reading Test (HART). Clin Neuropsychol 2009 Aug;23(6):926-943. [doi: 10.1080/13854040802603684] [Medline: 19191072]
66. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
67. Brandt J, Benedict HRB. Hopkins Verbal Learning Test-Revised Professional Manual. Lutz, FL: Psychologica Assessment Resources, Inc; 2001.
68. Benedict RHB, Schretlen D, Groninger L, Dobraski M, et al. Revision of the Brief Visuospatial Memory Test: studies of normal performance, reliability, and validity. Psychological Assessment 1996;8(2):145-153. [doi: 10.1037//1040-3590.8.2.145]
69. Goodglass H, Kaplan E. Boston Diagnostic Aphasia Examination. Philadelphia, PA: Lea and Febiger; 1983.
70. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept Mot Skills 1958 Dec 01;8(3):271-276. [doi: 10.2466/pms.1958.8.3.271]
71. Wechsler D. Wechsler Memory Scale-Revised Manual. San Antonio, TX: Psychological Corporation; 1987.
72. Salthouse TA. Mediation of adult age differences in cognition by reductions in working memory and speed of processing. Psychol Sci 1991 May 01;2(3):179-183. [doi: 10.1111/j.1467-9280.1991.tb00127.x]
73. McKay NS, Millar PR, Nicosia J, Aschenbrenner AJ, Gordon BA, Benzinger TLS, et al. Pick a PACC: comparing domain-specific and general cognitive composites in Alzheimer disease research. Neuropsychology 2024 Jul;38(5):443-464 [FREE Full text] [doi: 10.1037/neu0000949] [Medline: 38602816]
74. Schwartz MF, Montgomery MW, Buxbaum LJ, Lee SS, Carew TG, Coslett HB, et al. Naturalistic action impairment in closed head injury. Neuropsychology 1998;12(1):13-28. [doi: 10.1037/0894-4105.12.1.13]
75. Schwartz MF, Buxbaum LJ, Montgomery MW, Fitzpatrick-DeSalme E, Hart T, Ferraro M, et al. Naturalistic action production following right hemisphere stroke. Neuropsychologia 1999 Jan;37(1):51-66. [doi: 10.1016/s0028-3932(98)00066-9] [Medline: 9920471]
76. Buxbaum LJ. Ideational apraxia and naturalistic action. Cogn Neuropsychol 1998 Sep 01;15(6-8):617-643. [doi: 10.1080/026432998381032] [Medline: 22448839]
77. Marson D, Sawrie SM, Snyder S, McInturff B, Stalvey T, Boothe A, et al. Assessing financial capacity in patients with Alzheimer disease: a conceptual model and prototype instrument. Arch Neurol 2000 Jun;57(6):877-884. [doi: 10.1001/archneur.57.6.877] [Medline: 10867786]
78. Farias S, Mungas D, Reed B, Harvey D, Cahn-Weiner D, Decarli C. MCI is associated with deficits in everyday functioning. Alzheimer Dis Assoc Disord 2006;20(4):217-223 [FREE Full text] [doi: 10.1097/01.wad.0000213849.51495.d9] [Medline: 17132965]
79. BM SPSS Statistics for Windows (Version 29). Armonk, NY: IBM Corp; 2024.
80. Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. Acta Paediatr 2007 May 21;96(5):644-647. [doi: 10.1111/j.1651-2227.2006.00178.x] [Medline: 17376185]
81. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychological Bulletin 1979;86(2):420-428. [doi: 10.1037/0033-2909.86.2.420]
82. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163 [FREE Full text] [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
83. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159. [doi: 10.2307/2529310]
84. Eisinga R, Grotenhuis MT, Pelzer B. The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown? Int J Public Health 2013 Aug 23;58(4):637-642. [doi: 10.1007/s00038-012-0416-3] [Medline: 23089674]
85. Giovannetti T, Mis R, Hackett K, Simone SM, Ungrady MB. The goal-control model: An integrated neuropsychological framework to explain impaired performance of everyday activities. Neuropsychology 2021 Jan;35(1):3-18 [FREE Full text] [doi: 10.1037/neu0000714] [Medline: 33393796]
86. Cooper RP, Shallice T. Hierarchical schemas and goals in the control of sequential behavior. Psychological Review 2006;113(4):887-916. [doi: 10.1037/0033-295x.113.4.887]
87. Chirico A, Giovannetti T, Neroni P, Simone S, Gallo L, Galli F, et al. Virtual reality for the assessment of everyday cognitive functions in older adults: an evaluation of the virtual reality action test and two interaction devices in a 91-year-old woman. Front Psychol 2020;11:123 [FREE Full text] [doi: 10.3389/fpsyg.2020.00123] [Medline: 32116928]
88. Dodge H, Mattek N, Austin D, Hayes T, Kaye J. In-home walking speeds and variability trajectories associated with mild cognitive impairment. Neurology 2012 Jun 12;78(24):1946-1952 [FREE Full text] [doi: 10.1212/WNL.0b013e318259e1de] [Medline: 22689734]
89. Mcalister C, Schmitter-Edgecombe M, Lamb R. Examination of variables that may affect the relationship between cognition and functional status in individuals with mild cognitive impairment: a meta-analysis. Arch Clin Neuropsychol 2016 Mar 06;31(2):123-147 [FREE Full text] [doi: 10.1093/arclin/acv089] [Medline: 26743326]

90. Royall DR, Lauterbach EC, Kaufer D, Malloy P, Coburn KL, Black KJ. The cognitive correlates of functional status: a review from the Committee on Research of the American Neuropsychiatric Association. Journal of Neuropsychiatry 2007 Aug 01;19(3):249-265. [doi: 10.1176/appi.neuropsych.19.3.249]
91. Raihan MM, Subroto S, Chowdhury N, Koch K, Ruttan E, Turin TC. Dimensions and barriers for digital (in)equity and digital divide: a systematic integrative review. DTS 2024 Aug 13;4(2):111-127. [doi: 10.1108/dts-04-2024-0054]
92. Öhman F, Hassenstab J, Berron D, Schöll M, Papp K. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. Alzheimers Dement (Amst) 2021;13(1):e12217 [FREE Full text] [doi: 10.1002/dad2.12217] [Medline: 34295959]
93. Roll EE, Giovannetti T, Libon DJ, Eppig J. Everyday task knowledge and everyday function in dementia. J Neuropsychol 2019 Mar 26;13(1):96-120. [doi: 10.1111/jnp.12135] [Medline: 28949080]
94. Holzinger A. Finger instead of mouse: touch screens as a means of enhancing universal access. In: Carbonell C, Stephanidis N, editors. Universal Access: Theoretical Perspectives, Practice and Experience. Berlin/Heidelberg, Germany: Springer; 2003:387-397.

Abbreviations

ANCOVA: analysis of covariance
AUC: area under the curve
ECog-12: 12-item Everyday Cognition Scale
FAQ: Functional Activity Questionnaire
IADL-C: Instrumental Activities of Daily Living—Compensation
ICC: intraclass correlation coefficient
IPS: In-Plane Switching
Knight-PACC: Knight-Preclinical Alzheimer Cognitive Composite
MCI: mild cognitive impairment
mKnight-PACC: modified Knight-Preclinical Alzheimer Cognitive Composite
NAT: Naturalistic Action Test
PCA: principal component analysis
QHD+: Quad High Definition Plus
VKC: Virtual Kitchen Challenge
VKC-2: Virtual Kitchen Challenge—Version 2
VR: virtual reality

Edited by M O'Connell; submitted 12.Aug.2025; peer-reviewed by Z Kunicki, F Akhter; comments to author 19.Oct.2025; revised version received 09.Nov.2025; accepted 05.Dec.2025; published 07.Jan.2026.

Please cite as:
Kaplan M, McKniff M, Simone SM, Tassoni MB, Hackett K, Holmqvist S, Mis RE, Halberstadter K, Chaturvedi R, Rosahl M, Vallecorsa G, Serruya MD, Drabick DAG, Yamaguchi T, Giovannetti T
The Virtual Kitchen Challenge—Version 2: Validation of a Digital Assessment of Everyday Function in Older Adults
JMIR Aging 2026;9:e82092
URL: https://aging.jmir.org/2026/1/e82092
doi:10.2196/82092
PMID:41349042

Original Paper

Developing Consumer Consensus on Remote Assessment and Management of Physical Function in Older Adults (RAMP): International Modified Delphi Process

Elsa Dent[1], PhD; Christopher Hurst[2,3], PhD; Jack Dalla Via[4,5], PhD; Jackson J Fyfe[4], PhD; Paul Jansons[4,6], PhD; Eleanor J Hayes[7], PhD; Gary Skinner[4], MSc; Marc Sim[5,8], PhD; Mylene Aubertin-Leheudre[9,10], PhD; Sabine Britting[11], PhD; Fanny Buckinx[9,10,12], PhD; Gavin Connolly[13], PhD; Ruth Dignam[14†]; Lora Giangregorio[15,16], PhD; Jennifer R A Jones[17,18,19], PhD; Pauline Kelly[14]; Robert Kob[11], PhD; Suzanne N Morin[20], MD; Girish Nandakumar[21], PhD; Lucas B R Orssatto[22], PhD; Maria Pearson[23], MPH; Daniel Pinto[24], PT, PhD; Esmee M Reijnierse[25], PhD; Catherine M Said[17,26,27], PhD; Mohamed Salem[28], MB; Vina PS Tan[29], PhD; Rosanna Tran[30], BAPSC; Jesse Zanker[31], PhD; Robin M Daly[4], PhD; David Scott[4,6], PhD

[1]Division of Health, University of Waikato, Hamilton, New Zealand

[2]AGE Research Group, Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom

[3]NIHR Newcastle Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Cumbria Northumberland Tyne and Wear NHS Foundation Trust and Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom

[4]Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences, Deakin University, Burwood, Australia

[5]Nutrition & Health Innovation Research Institute, School of Medical and Health Sciences, Edith Cowan University, Joondalup, Australia

[6]School of Clinical Sciences at Monash Health, Monash University, Clayton, Australia

[7]Department of Sport, Exercise and Rehabilitation, Northumbria University, Newcastle Upon Tyne, United Kingdom

[8]School of Medicine, The University of Western Australia, Perth, Australia

[9]Département des Sciences de l'Activité Physique, Faculté des Sciences, Groupe de Recherche en Activité Physique Adaptée (GRAPA), Université du Québec à Montréal, Montréal, QC, Canada

[10]Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal (CRIUGM), Montréal, QC, Canada

[11]Institute for Biomedicine of Aging, Friedrich-Alexander-Universität Erlangen-Nürnberg, Nuremberg, Germany

[12]Research Unit in Public Health, Epidemiology and Health Economics, University of Liège, Liège, Belgium

[13]Metabolism and Healthy Aging Research Center (MHARC), Division of Endocrinology, Diabetes, and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA, United States

[14]IPAN Consumer Network, Deakin University, Geelong, Australia

[15]Schlegel-UW Research Institute for Aging, University of Waterloo, Waterloo, ON, Canada

[16]Department of Kinesiology and Health Sciences, University of Waterloo, Waterloo, ON, Canada

[17]Physiotherapy Department, The University of Melbourne, Parkville, Australia

[18]Physiotherapy Department, Division of Allied Health, Austin Health, Heidelberg, Australia

[19]Institute of Breathing and Sleep, Heidelberg, Australia

[20]Division of General Internal Medicine, Department of Medicine, McGill University, Montréal, QC, Canada

[21]Department of Physiotherapy, Manipal College of Health Professions, Manipal Academy of Higher Education, Manipal, India

[22]Centre for Sensorimotor Performance, School of Human Movement and Nutrition Sciences, Faculty of Health, Medicine and Behavioural Sciences, The University of Queensland, Brisbane, Australia

[23]Maria Pearson Coaching, Havelock North, New Zealand

[24]Department of Physical Therapy, Marquette University, Milwaukee, WI, United States

[25]Department of Nutrition and Dietetics, Faculty of Health, Sport and Physical Activity, Amsterdam University of Applied Sciences, Amsterdam, The Netherlands

[26]Physiotherapy, Western Health, St Albans, Australia

[27]Australian Institute for Musculoskeletal Science (AIMSS), The University of Melbourne and Western Health, St Albans, Australia

[28]Geriatrics, Active Ageing and Community Care, Ministry for Health and Active Ageing, Valletta, Malta

[29]Exercise and Sports Science Programme, School of Health Sciences, Universiti Sains Malaysia, Kubang Kerian, Malaysia

[30]Rehabilitation and Aged Care Services, Northern Sydney Local Health District, Sydney, Australia

[31]Department of Medicine - Royal Melbourne Hospital, The University of Melbourne, Parkville, Australia

[†]deceased

Corresponding Author:
David Scott, PhD
Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences
Deakin University
221 Burwood Highway
Burwood, 3125
Australia
Phone: 61 9246 8438
Email: d.scott@deakin.edu.au

Abstract

Background: Remote health care delivery, including the use of digital health interventions, is emerging as a tool for assessing and managing physical function, but its design and implementation often overlook the needs and preferences of older adult end users.

Objective: The primary aim of this modified Delphi process was to develop consumer consensus on preferences for remote assessment and management of physical function in older adults.

Methods: Research and consumer experts of the Remote Assessment and Management of Physical Function in Older Adults (RAMP) Working Group co-developed the Round 1 Delphi survey, which was advertised to consumers (adults aged ≥60 years) via international clinical and research networks and social media between August and November 2023. The online survey presented 23 Delphi statements for which respondents reported their level of agreement using an 11-point Likert scale (0-10; scores ≥7 indicated agreement). Statements were classified as having "strong agreement" and achieving consensus if ≥80% of participants indicated agreement. Statements classified as having "moderate" (70%-80% of participants indicated agreement) or "low" (<70% of participants indicated agreement) agreement were revised or rejected. Revised statements were presented to participants in Round 2 (January to February 2024), and the final consensus statements were consolidated into recommendations.

Results: A total of 654 consumers (75.7% female) with a mean age of 69.0 (SD 6.0) years from 15 countries (5 continents) were included in analyses in Round 1. Of 23 statements, 13 achieved consensus, with the strongest agreement observed for statements relating to the importance of physical function for quality of life and performing activities of daily living (6 statements; agreement 97.6%-99.5%). Two statements regarding privacy and security concerns when using technology (agreement 20.8%) and the inability to perform physical function assessments or exercise at home (agreement 15.5%) were rejected with low agreement. The remaining 8 statements (agreement 49.5%-79.5%) were modified into 7 new statements for the Round 2 survey, which was completed by 526 (80.4%) respondents from Round 1. Five of seven Round 2 statements were accepted with strong agreement (agreement 80%-82.7%), including the importance of addressing personal preferences for self- versus clinician-led remote interventions, group versus individual exercise, and availability of necessary resources (eg, technology and exercise equipment).

Conclusions: Eighteen statements achieved consensus and were translated into 7 recommendations highlighting that older adults recognize physical function as a health priority, would value more information about it, and are willing to participate in remote assessment and management interventions (including via digital health) to maintain or improve it. These recommendations also reinforce that interventions should be easily accessible and meet individual preferences of consumers.

(JMIR Aging 2026;9:e75791) doi:10.2196/75791

KEYWORDS

aged; Delphi technique; exercise; remote consultation; digital health; physical function

Introduction

Maintaining physical function is crucial for older adults to live independently and maintain a good quality of life [1]. Depending on the task examined, up to 50% of older adults may report difficulty with physical function, and more than 10% use walking aids, with an increasing prevalence of functional limitation in older age (eg, ≥80 years) [2]. Older adults with functional limitations are half as likely to engage with their communities, family, and friends [3] and have significantly increased health care costs compared to those without functional limitations [4]. However, assessment and management of physical function receive limited attention in many clinical settings [5-7]. While multimodal exercise is recommended for improving physical function [8], older adults report that lack of access to interventions, facilities, and relevant health care professionals are barriers to participation [9].

Remote health care services enable patient assessment and monitoring without physical co-location [10] and, especially since the COVID-19 pandemic, have been commonly used by older adults [11]. Emerging evidence suggests digital health interventions using technologies such as telephone calls, videoconferencing, wearable devices, and web applications are feasible and effective for supporting older adults to maintain or improve their physical function [12-14]. However, both consumers (ie, older adult end users) and health care professionals face barriers to engagement with digital health interventions, such as low digital literacy, lack of equipment, and unsatisfactory social interactions [15,16]. Furthermore, the design and implementation of digital health technologies for remote care have largely failed to consider the perceptions and experiences of consumers in their design and implementation, impeding large-scale sustained adoption [17].

The Delphi method, a rigorous consensus-building approach [18], is effective for developing recommendations for the delivery of care related to maintaining and improving physical function in older adults based on the insights of both experts and consumers (ie, older adults themselves) [19,20]. Incorporating consumer participation in consensus-building processes helps ensure that subsequent recommendations address end users' priorities, maximizing the potential for wide adoption and adherence among consumers in health care and community settings [21]. However, no consumer-focused consensus Delphi process has explored the priorities, acceptability, enablers, and barriers for remote assessment and management of physical function from older adults' perspectives, and there are no consensus guidelines on the delivery of remote care in older adult populations. Thus, this study aimed to develop consumer-informed recommendations for remote assessment and management of physical function in older adults via a modified Delphi process.

Methods

Study Design and Population

The Remote Assessment and Management of Physical Function in Older Adults (RAMP) International Consumer Delphi Process was a modified (2-round) Delphi study. Eligible participants were aged 60 years and older with internet access, residing in any country, and able to complete the survey in the English language.

RAMP was advertised to potential participants via email (consumer mailing lists of the RAMP Working Group), direct invitations (investigator-led conversations with consumers), and social media (Facebook [Meta Platforms, Inc] advertisements; posts on X [X Corp] and LinkedIn [LinkedIn Corp]) between August and November 2023.

Delphi Process

This study's methodology adhered to that of a previous modified Delphi process where 2 rounds, as demonstrated by common practice in previous Delphi studies [22,23], were sufficient to reach consensus on sarcopenia management [19]. Multimedia Appendix 1 summarizes the study timeline. The Round 1 survey (Multimedia Appendix 2) was co-developed by RAMP Working Group research experts (DS, CH, JF, PJ, and RMD) and consumer experts (RD and PK) and included questions on demographics, health status (including the SARC-F questionnaire [24]), and health care experiences, plus 23 Delphi statements related to physical function and its assessment and management. The first 13 statements relate to physical function generally, while the latter 10 statements refer to remote provision of health care related to physical function. Respondents were asked to report their level of agreement with each Delphi statement via an 11-point Likert scale (0=Strongly disagree, 10=Strongly agree) and could optionally provide a free-text comment explaining their response to each statement.

Round 1 survey responses were analyzed by DS, and the results and proposed Round 2 survey were shared with RAMP Working Group members for feedback. In January 2024, an invitation to participate in the Round 2 survey, which included seven new or modified statements, was emailed to respondents who completed Round 1, along with a summary of the Round 1 results and an explanation of decisions taken in developing the Round 2 statements (Multimedia Appendix 3). Further details on reasons for low or moderate agreement to Round 1 statements and decisions taken are provided in Multimedia Appendix 4. Upon completion of the Round 2 survey (Multimedia Appendix 5), DS analyzed the results, and ED and DS initially developed the final recommendations. Finally, this manuscript, including the developed recommendations, was revised (2 rounds of revisions) and approved by RAMP Working Group members.

Statistical Analyses

Survey data were assessed for completeness, and all respondents who completed ≥50% of survey questions were included in analyses. Where duplicate responses were identified, the most complete and/or first response was included. Descriptive characteristics were reported as frequencies or percentages for categorical variables and means and SDs or medians and IQR for continuous variables.

Participants with a response ≥7 out of 10 were considered to have agreed with a given statement, and the level of consensus was determined by the proportion of participants who agreed. Round 1 and 2 Delphi statements were classified as having "strong agreement" if ≥80% of participants responded with a score ≥7 out of 10 [19]. These statements were considered to have achieved consensus and were not further modified. In Round 1 only, statements with "moderate" (70%-80% of participants responded ≥7) or "low" (<70% of participants responded ≥7) agreement were revised or rejected. DS reviewed the associated free-text comments and revised the statements based on common reasons for lack of agreement. The revised statements were shared with the RAMP Working Group for approval. In Round 2, all statements that did not achieve "strong agreement" were rejected. All analyses were performed using SPSS Statistics 28 (IBM Corp).

Ethical Considerations

The study was approved by the Deakin University Human Ethics Advisory Group (Reference number: HEAG-H 111_2023) and complied with the Declaration of Helsinki. Potential participants were directed to an online plain language statement defining physical function (ie, "the ability of a person to perform everyday activities"), as well as the study aims and methods. Potential participants provided an electronic signature confirming their informed consent. Those who consented were subsequently emailed a link to the Round 1 Delphi survey (hosted by Qualtrics XM). Participants who completed the Round 1 survey were invited to participate in the Round 2 survey between January and February 2024. Participants who consented but had not completed the Round 1 and 2 surveys were sent a reminder 2 weeks prior to the closing date. Participation was voluntary, and respondents did not receive any form of reimbursement.

Results

Round 1

A total of 861 complete consent forms were received, and the Round 1 Delphi survey subsequently received 716 responses during the Round 1 survey period. Of these responses, 50 (7%) completed less than 50% of the survey, and 12 (2%) were duplicate responses. Thus, 654 of 861 (76%) consented consumers were included in the Round 1 survey analyses. Most of the included respondents (n=644, 98%) completed the entire Round 1 survey, while 10 respondents rated their agreement with the first 13 statements only. Respondents' mean age was 69.0 (SD 6.0) years, and three-quarters were female (Table 1). Respondents were residing in 15 countries across 5 continents (Australia, Africa, Asia, Europe, and North America), although the majority (81.5%) were from Australia and the United Kingdom. Three-quarters of respondents had completed higher education and around 70% were retired. Over 87% rated their health as good to excellent, but more than 70% perceived their current physical function was "somewhat worse" or "much worse" than when they were 40 years old. The SARC-F instrument demonstrated that approximately 54% of respondents had some functional limitation, and around one-quarter had experienced at least one fall in the past year (Multimedia Appendix 6).

Table 1. Demographic characteristics of Remote Assessment and Management of Physical (RAMP) participants in Round 1.

Characteristic	All respondents (n=654)
Age (years), mean (SD)	69 (6.0)
Sex (female), n (%)	495 (75.7)
Country of residence, n (%)	
Australia	281 (43.0)
United Kingdom	252 (38.5)
Canada	32 (4.9)
Ireland	22 (3.4)
United States	20 (3.1)
New Zealand	15 (2.3)
Germany	7 (1.1)
Singapore	7 (1.1)
Malta	6 (0.9)
India	5 (0.8)
Netherlands	3 (0.5)
Belgium	1 (0.2)
France	1 (0.2)
Georgia	1 (0.2)
South Africa	1 (0.2)
Highest level of education, n (%)	
Infants or primary school	3 (0.5)
Secondary or high school	154 (23.5)
University, college, or other higher education	497 (76.0)
Employment status, n (%)	
Employed full-time	67 (10.2)
Employed part-time	90 (13.8)
Home duties	3 (0.5)
Pension	18 (2.8)
Retired	462 (70.6)
Student	3 (0.5)
Unemployed	11 (1.7)
General health, n (%)	
Excellent	89 (13.6)
Very good	255 (39.0)
Good	228 (34.9)
Fair	67 (10.2)
Poor	15 (2.3)
Physical function now compared with that at age 40 years, n (%)	
Much better	18 (2.8)
Somewhat better	44 (6.7)
Neither better nor worse	117 (17.9)
Somewhat worse	348 (53.2)
Much worse	127 (19.4)

Characteristic	All respondents (n=654)
SARC-F score, median (IQR)	1 (0-2)

Less than one-third of respondents reported that a health care professional had started a conversation with them about their physical function in the past 5 years, while 38% had initiated a conversation themselves (Table 2). More than 60% had sought information on physical function from sources other than a health care professional. Around 40% of respondents reported having ever completed a physical function assessment with a health care professional, and almost 50% had commenced a supervised exercise program aimed at improving their physical function. The majority of respondents (>60%) had commenced an unsupervised exercise program aimed at improving their physical function.

Table 2. Health care experiences related to physical function for Remote Assessment and Management of Physical (RAMP) participants in Round 1 (n= 654).

Question	Yes, n (%)	No, n (%)	Don't Know, n (%)
In the past five years, have you started a conversation with a health professional (eg, doctor, physiotherapist, and nurse) about your physical function (eg, asking how you can continue to stay independent as you get older, or why you might not be as strong as you were when you were younger)?	245 (37.5)	399 (61)	9 (1.4)
In the past 5 years, has a health professional (eg, doctor, physiotherapist, and nurse) started a conversation with you about your physical function?	209 (32)	436 (66.7)	7 (1.1)
Have you ever tried to find information about physical function from sources other than a health professional (eg, by asking a friend or family member, visiting a website, or reading a book or magazine)?	395 (60.4)	255 (39)	3 (0.5)
Have you ever completed a physical function test (eg, walking speed test, hand grip strength test, and chair stand test) under the supervision of a health professional (eg, where a health professional asked you to perform a specific test while under their supervision to determine whether your physical function was poor)?	270 (41.3)	376 (57.5)	7 (1.1)
Have you ever completed a physical function test (eg, walking speed test, hand grip strength test, and chair stand test) while NOT under the supervision of a health professional (eg, performing a specific test designed to determine whether your physical function is poor after you read or viewed instructions in a document or online)?	101 (15.4)	543 (83)	9 (1.4)
Has a health professional ever prescribed you an exercise program aimed at improving your physical function?	315 (48.2)	333 (50.9)	5 (0.8)
Have you ever commenced an exercise program aimed at improving your physical function while NOT under the supervision of a health professional (eg, an exercise program that you created for yourself, with or without the help of a friend or family member, or using a website/book/magazine)?	402 (61.5)	245 (37.5)	6 (0.9)

Over 58% of respondents had participated in a remote health care service, while less than 5% had participated in a remote physical function test, and less than 20% had engaged in remote physical function management such as an exercise program (Figure 1). For respondents who had participated in any form of remote care, telephone calls were most commonly used (77%), with video calls and emails/text messages each used by over 30% of respondents. For respondents who had participated in a remote physical function assessment, video calls were most commonly used (75%), followed by telephone calls (25%). For respondents who had participated in a remote treatment for their physical function, written documents (eg, flyers, brochures, magazines, and books) were most common (46%), followed by video calls (38%) and telephone calls (31%). Regardless of the type of remote care, respondents generally reported positive experiences; 75%-81% rated their experience as somewhat or very positive.

Only 5% of respondents reported that they would not be willing to participate in remote physical function tests and treatments (Multimedia Appendix 7), with optional comments indicating that this was due to concerns regarding lack of supervision, safety, lack of necessary technologies and technological familiarity, data privacy, hearing and visual problems, potential costs, and insufficient space in the home. Amongst those willing to participate in such interventions, videos (eg, on a website or DVD) were the most preferred delivery method (78%). Approximately 4% were willing to use other approaches; free-text comments indicated that these could include digital voice assistants, wearable devices, webinars, smart televisions, and wall charts. The most common perceived positives of remote tests and treatments, selected by over 80% of respondents, were the convenience of not needing to travel to appointments and the flexibility to perform assessments and exercises when suitable. Some (<4%) respondents nominated other positives in free-text comments, including avoiding body shaming or embarrassment when exercising, self-motivation, empowerment and autonomy for managing exercise, and better use of health care professionals' time. Regarding negatives of remote care, the lack of personalized guidance during exercises was reported by half (51%) of respondents, and concerns regarding the potential ineffectiveness of interventions and a lack of social interaction and motivation were each reported by more than

40%. Approximately 14% of respondents cited other potential negatives, including costs and/or lack of access to necessary equipment and technological devices, lack of confidence in performing exercises correctly, low motivation to exercise, lack of engagement with programs and health care professionals, and insufficient space to perform exercises at home.

Figure 1. Remote health care experiences of RAMP participants in Round 1, related to (A) any type of remote care, (B) remote care to assess physical function, and (C) remote care to manage physical function. Sublabels (1) describe the proportion of respondents who reported having participated in that type of remote care, (2) report the modalities by which respondents engaged with that type of remote care, and (3) report the perceived experience of respondents participating in that type of remote care. Questions 2 and 3 were only asked to participants who responded yes to having participated in that type of remote care. For question 2, participants could choose multiple responses, so the total does not equal 100%.

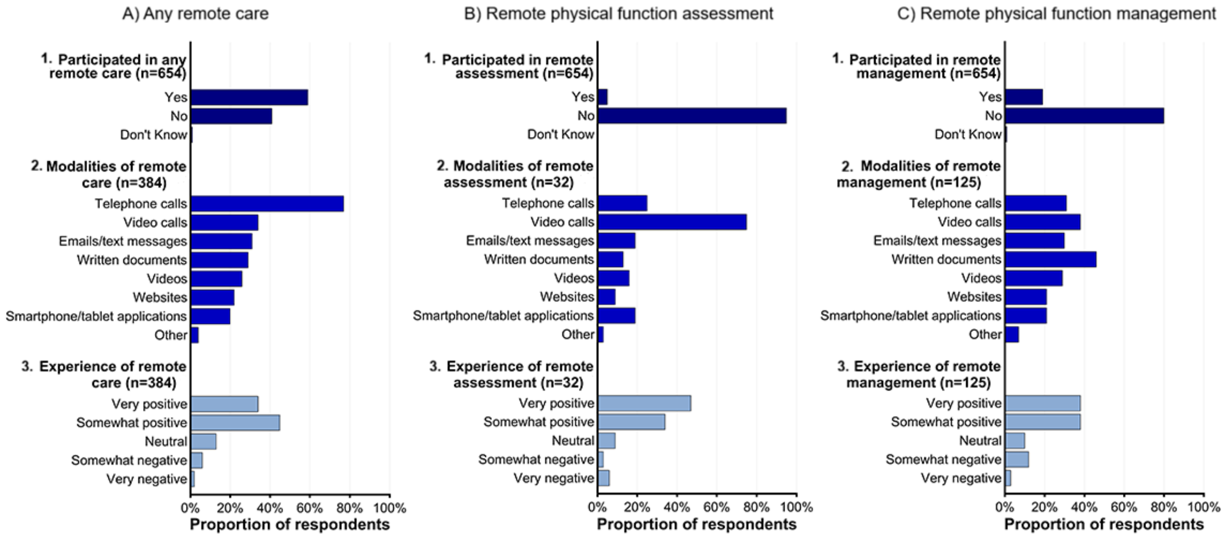


Figure 2 and Table 3 present respondents' reported agreement with the 23 Delphi statements from Round 1. Statements related to the importance of physical function for maintaining overall quality of life, participating in activities in the community and home, and engaging with family and friends (Statements 1.01-1.06) had the highest proportion (97%-100%) of respondents who reported agreement (response ≥7 out of 10) in this round and were accepted. Consensus was also achieved that poor physical function can be prevented and reversed (Statements 1.07 and 1.08; both 86% agreement) and that respondents would discuss with their health care professionals if concerned about their physical function (Statement 1.09; 80% agreement). There was moderate agreement that respondents would like access to information on how to test their physical function (Statement 1.10; 78% agreement) and low agreement that having better access to information would assist in conversations with health care professionals (Statement 1.12; 69% agreement). However, there was strong agreement that respondents would like access to information on how to improve their physical function (Statement 1.11; 86% agreement) and that having better access to information would assist in managing their physical function independently (Statement 1.13; 81% agreement).

Regarding remote provision of health care, almost 80% of respondents agreed that they would be willing to participate in remote physical function tests, but this statement (Statement 1.14) did not meet the criterion for consensus. There was strong agreement that physical function tests would be safe to perform without direct supervision (Statement 1.15; 86% agreement) but only low agreement that respondents would be willing to participate in a remote exercise program that was always (Statement 1.16; 50% agreement) or sometimes (Statement 1.17; 63% agreement) supervised, and moderate agreement that they would be willing to participate in an unsupervised exercise program (Statement 1.18; 72% agreement). There was also low or moderate agreement that respondents would be happy to participate in a remote exercise program in a group setting (Statement 1.19; 45% agreement) or individually (Statement 1.20; 77% agreement). Finally, low proportions of respondents agreed that they would be concerned about their privacy and security when using technology to participate in remote care (Statement 1.22; 21% agreement) or that remote physical function tests or exercise programs would be difficult to perform in their home (Statement 1.23; 16% agreement).

Figure 2. Representation of participants' level of agreement (rated on an 11-point Likert scale; 0-10) with each of 23 Round 1 Delphi statements. n=654 for statements 1.01-1.13, n=644 for statements 1.14-1.23. Participants with a response ≥7 were considered to have agreed with the statement. Statements with strong agreement (≥80%; black dotted line) were accepted. Statements with 70%-<80% agreement (grey dotted line) were considered to have moderate agreement. Statements with <70% agreement were considered to have low agreement.

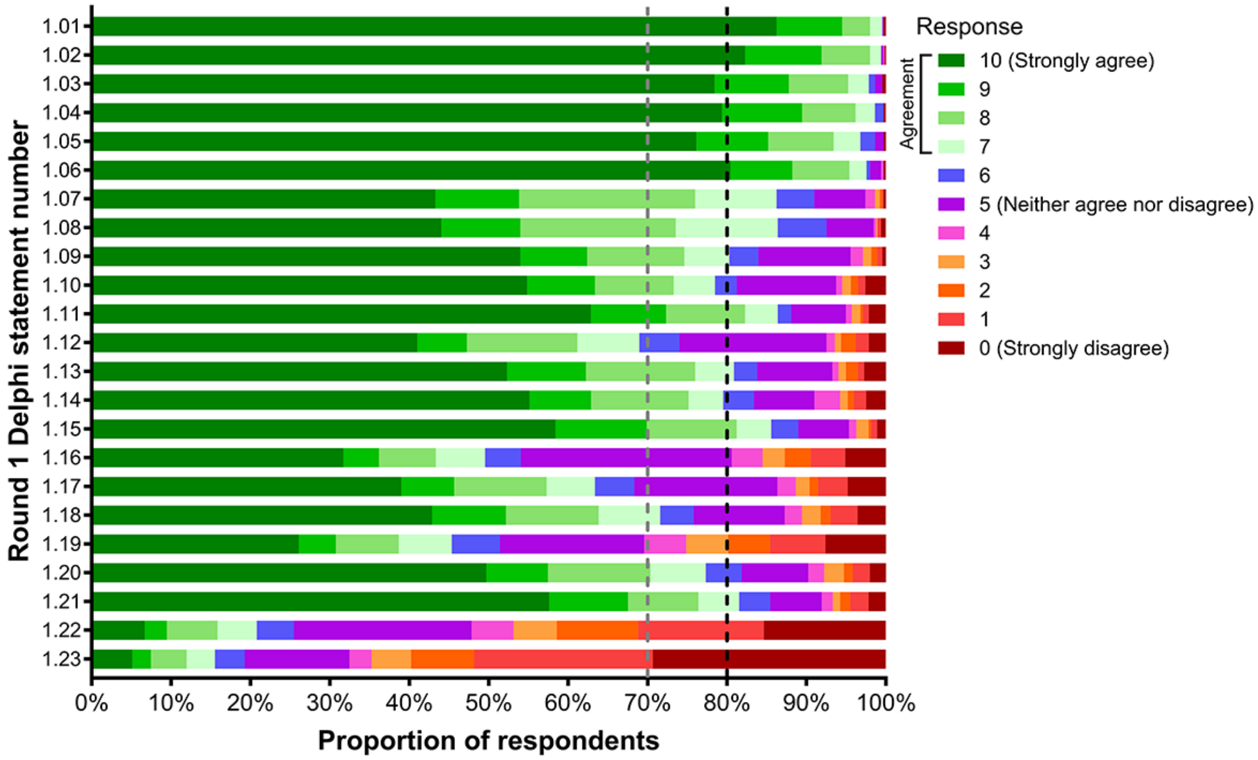


Table 3. List of Round 1 Delphi statements with levels of agreement and outcome decisions.

Statement number	Statement	Agreement (%)[a]	Agreement rating[b]	Outcome
1.01	Having good physical function is important to the overall quality of life of older adults	99.5	Strong agreement	Accept
1.02	Having good physical function is important for activities involving moving around the community (eg, going shopping or to a restaurant or cafe, visiting your neighbors, friends and family, or the doctor)	99.4	Strong agreement	Accept
1.03	Having good physical function is important for participating in activities with family and friends (eg, playing with grandchildren)	97.9	Strong agreement	Accept
1.04	Having good physical function is important for participating in activities like work, household duties (eg, cooking, cleaning, and gardening), and volunteering	98.6	Strong agreement	Accept
1.05	Having good physical function is important for participating in hygiene activities (eg, showering, dressing, and using the toilet)	96.8	Strong agreement	Accept
1.06	Having good physical function is important for participating in exercise (eg, walking, swimming, dancing, golf, and other types of physical activity)	97.6	Strong agreement	Accept
1.07	It is possible to slow down or prevent poor physical function that occurs as we get older	86.2	Strong agreement	Accept
1.08	If someone already has poor physical function, it is possible to improve it	86.4	Strong agreement	Accept
1.09	If I was concerned about my physical function, I would discuss it with my health professional	80.3	Strong agreement	Accept
1.10	I would like access to information about how to test my physical function myself to determine if it is poor	78.4	Moderate agreement	Modify statement for Round 2
1.11	I would like access to information about things that I can do myself to improve my physical function	86.4	Strong agreement	Accept
1.12	Having better access to information on physical function would help me to have conversations about this with health professionals	69.0	Low agreement	Modify statement for Round 2
1.13	Having better access to information on physical function would help me to take care of my own physical function	80.9	Strong agreement	Accept
1.14	I would be willing to participate in remote tests of my physical function (eg, on a video call with a health professional, or by myself using written instructions and/or video demonstrations provided to me)	79.5	Moderate agreement	Modify statement for Round 2
1.15	I am confident that it would be safe for me to perform physical function tests at home without direct supervision by a health professional if I was provided with instructions (eg, written information or video demonstrations)	85.6	Strong agreement	Accept
1.16	I would be willing to participate in a remote exercise program to improve my physical function if it was ALWAYS supervised (eg, exercising while on a live video call with a health professional for all exercise sessions)	49.5	Low agreement	Modify statement for Round 2
1.17	I would be willing to participate in a remote exercise program if it was SOMETIMES supervised (eg, exercising on a live video call with a health professional for some exercise sessions, but exercising by myself unsupervised using instructions provided by the health professional for other sessions)	63.4	Low agreement	Modify statement for Round 2
1.18	I would be willing to participate in a remote exercise program if it was NOT supervised (eg, exercising by myself unsupervised using instructions provided by a health professional)	71.6	Moderate agreement	Modify statement for Round 2

Statement number	Statement	Agreement (%)[a]	Agreement rating[b]	Outcome
1.19	If I was to participate in a remote exercise program I would be happy to do so with a group (eg, exercising by myself at home but while on a video call with other people like me who are also exercising at home, with or without the supervision of a health professional)	45.3	Low agreement	Modify statement for Round 2
1.20	If I was to participate in a remote exercise program to improve my physical function, I would be happy to do so alone without other people like me involved in the exercise sessions (eg, exercising by myself at home with or without supervision by a health professional)	77.3	Moderate agreement	Modify statement for Round 2
1.21	I would be comfortable using technology (eg, computers, smartphones, and tablets) to participate in remote tests and treatments for my physical function	81.5	Strong agreement	Accept
1.22	I would be concerned about the privacy and security of my personal information when participating in remote tests and treatments for physical function using technology (eg, computer, smartphone, or tablet)	20.8	Low agreement	Reject
1.23	Remote physical function tests or exercise programs would be difficult to perform in my home (eg, because there is limited space)	15.5	Low agreement	Reject

[a]Proportion of participants who rated statement ≥7 out of 10.

[b]Statement classification based on the following criteria: Strong agreement (>80% of respondents rated statement ≥7 out of 10); Moderate agreement (70% to 80% of respondents rated statement ≥7 out of 10); Low agreement (<70% of respondents rated statement ≥7 out of 10).

Round 2

Following Round 1, 13 of 23 statements were accepted to have achieved consensus with strong agreement. It was determined that statements 1.22 and 1.23 had such low agreement that they should be rejected rather than modified and presented again in Round 2. Other statements that achieved moderate or low agreement (Statements 1.10, 1.12, 1.14, 1.16, 1.17, 1.18, 1.19, and 1.20) were modified. In brief, Statements 1.10, 1.12, and 1.14 were revised into Statements 2.10, 2.12, and 2.14, respectively. Statements 1.16, 1.17, and 1.18, which covered similar concepts regarding remote exercise supervision, were merged into a single revised statement (Statement 2.16). Statements 1.19 and 1.20 covered similar concepts regarding preferences for participating in remote exercise individually or in a group setting and were revised into a single statement (Statement 2.19). Finally, based on common themes identified in free-text responses to several statements across Round 1, 2 new statements were introduced for Round 2: Statement 2.24 explored the importance for consumers that remote physical function tests have been demonstrated to be safe and accurate, and Statement 2.25 explored the importance of having access to necessary information and resources to support participation in remote exercise programs. Multimedia Appendices 3 and 4 provide further details on the development of Round 2 Delphi statements.

Among participants who completed Round 1, 526 (80%) completed Round 2. No notable differences were observed between the Round 2 and Round 1 respondents, respectively, for age (mean = 69.3, SD 5.7 vs 69.0, SD 6.0 years) or SARC-F score (median 1.0, IQR 0.0-2.0 vs median 1.0, IQR 0.0-2.0). In Round 2, respondents were asked to rate their physical function compared with when they completed the Round 1 survey (mean interval between survey completions was 117.2, SD 21.5 days). The majority (68%) reported that their physical function was neither better nor worse, while 13% reported their physical function was somewhat or much better, and 19% reported that it was somewhat or much worse.

Figure 3 and Table 4 summarize participant responses to the Round 2 Delphi statements. Respondents strongly agreed that they would like access to instructions on how to test their own physical function and monitor changes (Statement 2.10; 85% agreement) and that having access to information on physical function would allow them to have more informed conversations with health care professionals (Statement 2.12; 82% agreement). Respondents also strongly agreed they would be willing to participate in a remote test of their physical function (Statement 2.14; 83% agreement) and to participate in a remote exercise program suited to their preferences at the time (Statement 2.16; 82% agreement). However, there was only moderate agreement that respondents would be willing to participate in remote exercise programs with a group or individually (Statement 2.19; 72% agreement) and that they would be more likely to participate in a remote physical function test if they were confident that the test was safe and accurate to perform alone (Statement 2.24; 77% agreement). Finally, there was strong agreement that respondents would be more likely to participate in a remote exercise program if they had access to necessary information and resources, including technology and exercise equipment (Statement 2.25; 80% agreement).

Figure 3. Representation of participants' level of agreement (rated on an 11-point Likert scale; 0-10) with each of 7 Round 2 Delphi statements (n=526). Participants with a response ≥7 were considered to have agreed with the statement. Statements with strong agreement (≥80%; black dotted line) were accepted.

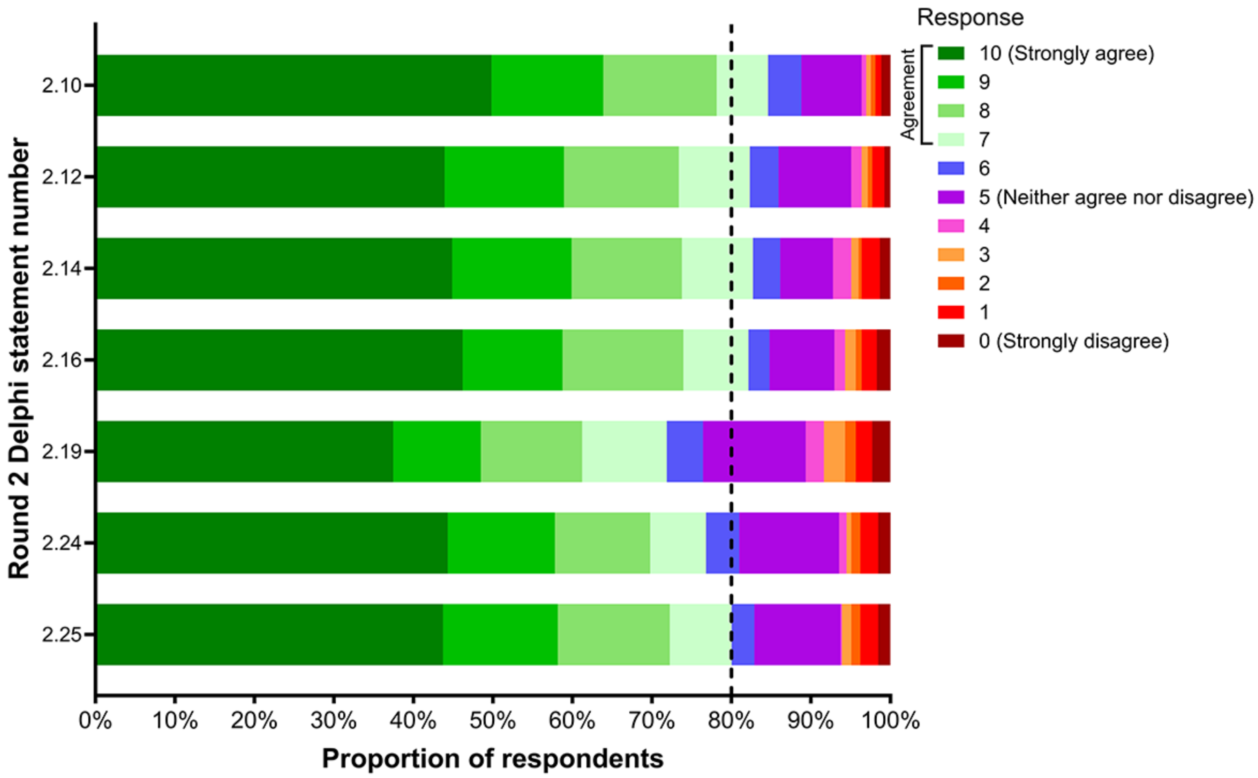


Table 4. List of Round 2 Delphi statements with levels of agreement and outcome decisions.

Statement Number	Statement	Agreement (%)[a]	Agreement Rating[b]	Outcome
2.10	I would like access to simple and reliable instructions on how to test my physical function myself so that I can monitor how it changes over time	84.6	Strong agreement	Accept
2.12	If I felt I needed help to improve or maintain my physical function, having access to simple information about this (including advice on appropriate health professionals to discuss it with) would help me to have more informed conversations with health professionals about my physical function	82.3	Strong agreement	Accept
2.14	If I felt I needed help to improve or maintain my physical function, I would be willing to participate in a remote test (eg, supervised on a live video call with a health professional or unsupervised using printed instructions and/or video demonstrations provided to me)	82.7	Strong agreement	Accept
2.16	If I felt I needed help to improve or maintain my physical function, I would be willing to participate in a remote exercise program suited to my preferences at the time which may include exercise supervised by a health professional, and/or exercise led by myself	82.1	Strong agreement	Accept
2.19	If I felt I needed help to maintain or improve my physical function, I would be willing to participate in a remote exercise program suited to my preferences at the time which may include exercise performed by myself, and/or exercise performed with a group of people	71.9	Moderate agreement	Reject
2.24	I would be more likely to participate in a remote test of physical function if I was confident that the test was safe and accurate to perform by myself, and I had access to the necessary information and resources, including technology and equipment, to perform the test myself	76.8	Moderate agreement	Reject
2.25	I would be more likely to participate in a remote exercise program if I was confident that I had access to the necessary information and resources, including technology and exercise equipment, to exercise safely and effectively	80.0	Strong agreement	Accept

[a]Proportion of participants who rated statement ≥7 out of 10.

[b]Statement classification based on the following criteria: Strong agreement (>80% of respondents rated statement ≥7 out of 10); Moderate agreement (70% to 80% of respondents rated statement ≥7 out of 10); Low agreement (<70% of respondents rated statement ≥7 out of 10).

Summary of Delphi Outcomes and Recommendations

Five of seven Round 2 statements achieved consensus (ie, ≥80% agreement), and 2 (2.19 and 2.24) were rejected with moderate agreement. Thus, including the 13 statements accepted in Round 1, a total of 18 statements achieved consensus in this Delphi process and provided a basis for the 7 recommendations relevant to health care professionals, researchers, and policymakers developed by the RAMP Working Group and presented in Table 5.

Table 5. Recommendations for remote assessment and management of physical function in older adults.

Recommendation number	Recommendation	Supporting Delphi statements
1	Recognize that physical function is an important health priority for older adults	1.01, 1.02, 1.03, 1.04, 1.05, 1.06
2	Recognize that older adults are aware that it is possible to prevent, delay, and reverse declines in physical function	1.07, 1.08
3	Regularly engage older adults in discussions about their physical function	1.09, 2.12
4	Provide older adults with accessible and reliable information on how to monitor and maintain their physical function	1.11, 1.13, 2.10
5	For older adults who have concerns about their physical function, facilitate physical function assessments in-person and/or remotely considering feasibility and consumer preferences	2.14
6	Facilitate in-person and/or remote exercise programs for improving older adults' physical function considering feasibility, consumer preferences, and supervision requirements	2.16
7	Ensure that older adults participating in remote assessments or exercise for physical function have access to, or are provided with, appropriate support, technology, and/or equipment to perform assessments and exercises safely and effectively	1.15, 1.21, 2.25

Discussion

Principal Results

The World Health Organization's global strategy on digital health stresses the importance of person-centered approaches where end users are engaged in the design and development phases of digital health approaches [25]. This consumer-focused Delphi process demonstrated that older adults recognize physical function as a health priority and are generally accepting of remote assessment and management, including the use of digital health approaches. Based on these findings, the RAMP Working Group has developed 7 recommendations for researchers, health care professionals, and policymakers to guide remote assessment and management of physical function.

Comparison with Prior Work

Consistent with previous health care professional surveys showing that assessment and management of physical function are infrequently performed in clinical settings [7], our Round 1 survey demonstrated less than half of older adult respondents had received any health care professional-led assessment or intervention for physical function. Even fewer had participated in a remote physical function assessment (<5%) or intervention (<20%). This is despite almost 60% of respondents having participated in some form of remote care, likely due to increased general digital health use during and since the COVID-19 pandemic [11]. For those respondents who had participated in remote physical function care, several methods were reported, likely influenced by their appropriateness for the desired outcome. For example, video calls were most commonly used for remote physical function assessments whereas written documents were most commonly used for remote interventions. Respondents reported receptiveness to a range of different remote care approaches, and less than 5% stated they would not be willing to participate in any remote care for physical function. Key facilitators for remote care included convenience and flexibility in scheduling, while barriers included lack of guidance and motivation as reported previously [26]. Based on the Delphi statements that achieved consensus, seven recommendations for remote care of physical function in older adults were developed. The above barriers and facilitators should be considered when seeking to implement these recommendations in research, clinical care, and policy. Our findings are consistent with a recent position statement on telehealth policy for older adults, which highlights the need for dedicated policies to address common barriers to telehealth among older adults [27].

Recommendations for Researchers, Health Care Professionals, and Policymakers

Recognize That Physical Function Is an Important Health Priority for Older Adults

Over 97% of respondents strongly agreed physical function was integral to their overall quality of life and their ability to participate in activities with family and friends and in the community, to exercise, and to complete self-care tasks. This recommendation is in line with the World Report on Ageing and Health [28], which emphasizes that maximizing functional ability is a priority for older adults.

Recognize That Older Adults are Aware it Is Possible to Prevent, Delay, and Reverse Declines in Physical Function

Over 86% of respondents strongly agreed that it is possible to slow down or prevent poor physical function and that if someone already has poor physical function, it is possible to improve it. This knowledge may encourage older adults to engage in physical function care [29] and health care professionals should leverage it to promote and implement individually tailored strategies to maintain or improve physical function.

Regularly Engage Older Adults in Discussions About Their Physical Function

Respondents strongly agreed that they would raise concerns about their physical function with health care professionals and that access to information about physical function would help them initiate these conversations. Previous research has highlighted the importance that older adults place on mutual goal setting with health care professionals regarding their physical function [29], and that aligning care with patient priorities can lead to better health outcomes [30]. Health care professionals may require upskilling to ensure effective collaboration with consumers on identifying causes and symptoms of poor physical function and promoting benefits and strategies for maintaining physical function [8].

Provide Older Adults With Accessible and Reliable Information on How to Monitor and Maintain Their Physical Function

In addition to conversations with health care professionals, respondents reported a desire for information on how to independently monitor and maintain their physical function. Previous research has highlighted the importance of providing physical function information and advice to patients [30]. It is necessary to develop and promote appropriate resources that empower older adults to monitor and maintain their physical function.

For Older Adults Who Have Concerns About Their Physical Function, Facilitate Physical Function Assessments In-Person and/or Remotely Considering Feasibility and Consumer Preferences

There is a lack of data on the acceptability and appropriateness of remote physical function assessment [31]. In this study, however, respondents agreed they would be willing to participate in a remote physical function assessment if they were concerned about it. Remote assessment of physical function for older adults can be as reliable as face-to-face assessments [32,33] although further research is required to identify the most appropriate physical function tests, protocols, and communication platforms to support reliable remote physical function assessment, as well as facilitators and barriers to implementation in home and community settings.

Facilitate In-Person and/or Remote Exercise Programs for Improving Older Adults' Physical Function

Considering Feasibility, Consumer Preferences, and Supervision Requirements

Round 1 statements regarding remote exercise programs with set levels of supervision (ie, always, sometimes, or not supervised) or a set format (ie, individual or group-based) achieved only low to moderate agreement, reflecting the varied preferences for exercise among older adults [34]. Remotely delivered exercise programs with varying levels of synchronous and asynchronous exercise may be beneficial for community-dwelling older adults who do not have access to exercise facilities or prefer exercising alone and/or at home [35]. Remote programs can support novel and engaging exercise approaches such as integrating exercise into everyday activities [36], "exercise snacking" [12,13] and gamification [37], and may also incorporate other behavioral and educational interventions such as nutrition counseling [38].

Many medical professionals lack knowledge on exercise prescription [39]. Referral of patients to an exercise professional with experience in supporting older adults to exercise via remote care should be a consideration for clinicians who do not feel qualified to prescribe exercise. Upskilling exercise professionals in effectively delivering exercise via remote care and providing access to requisite resources and equipment is also important to build capacity for remote management of physical function.

Ensure That Older Adults Participating in Remote Assessments or Exercise for Physical Function Have Access to, or Are Provided With, Appropriate Support, Technology, and/or Equipment

Respondents believed it was safe and feasible for them to perform remote physical function assessments and exercise programs, and contrary to previous research [40], had relatively low agreement (<21%) that privacy was a concern. Nevertheless, ensuring remote care is administered by secure technologies and adheres to privacy laws is an important consideration for health care professionals to reduce barriers for those who do have concerns. Most of our respondents strongly agreed that they were comfortable using technology to participate in remote tests and treatments for their physical function. Technological literacy can be a barrier to participation in remote health assessments [26] and further research is required to identify approaches to overcoming technological barriers to remote care. Our results demonstrate that older adults agree that access to appropriate instructions, technology, and equipment would increase the likelihood of participation in remote programs. These findings suggest that older adults can successfully participate in remotely delivered exercise programs if appropriate support is provided. This can include standby technical assistance and technology orientation sessions, especially in the early stages of the program [41]. This support can potentially be integrated into existing funding models to support remote care which have increased internationally, particularly following the COVID-19 pandemic, although further development of policy and reimbursement mechanisms is needed for sustainable integration [42].

Recommendations for Future Studies

Based on the current results, it is recommended that future research, informed by collaboration with consumers, caregivers, health care professionals, and policymakers should focus on, but not be limited to, the following outcomes:

- Identifying and addressing barriers to accessibility of remote physical function assessment and management for older populations, particularly in culturally and linguistically diverse populations, those with socioeconomic disadvantage, those from low-middle-income countries, and/or those with low technology literacy or with limited access to technology
- Determining optimal approaches (including protocols and technologies) to delivering remote care for physical function to ensure validity and reliability of assessments and effectiveness and safety of interventions
- Exploring cost-effectiveness and implementation processes to embed remote care for physical function across varying levels of health care internationally
- Developing evidence-based guidelines and health promotion strategies for remote physical function assessment and management in older adults

Strengths and Limitations

Our modified Delphi study was co-developed with health care consumers and incorporated 2 rounds of iterative and anonymous questionnaires and controlled feedback to create consensus. Our study adhered to quality evaluation metrics for Delphi methodology [22] and included an international population of older adults. A high level of agreement (80%) was set a priori for acceptance of statements. There was low attrition of participants between Delphi rounds (<20%), suggesting the respondents were engaged and interested in sharing their views on physical function assessment and management.

Despite these strengths, there were limitations to our study. Given our survey was electronic, the participants required internet access and were likely technologically savvy, and thus a selection bias may be present. Respondents were from 15 generally high-income countries, with the majority residing in Australia and the United Kingdom, a large proportion (76%) were tertiary-educated, and all were English-speaking. It is not known if results are generalizable to those with lower socioeconomic status and/or non–English-speaking individuals. Our respondents also included a large proportion of women (76%), so it may not accurately reflect the views and experiences of older men. Similarly, less than 5% of respondents were aged 80 years or older, and less than 7% had a SARC-F score ≥4 (symptomatic of poor physical function) [24]. However, comparable SARC-F data in our study (median 1, IQR 0-2, proportion with score ≥4=7%) and other similar cohorts (median 0, IQR 1-2, proportion with score ≥4=6%-15%) [24] suggests the level of poor physical function in our sample is generally representative of community-dwelling older adults. Overall, further research investigating remote physical function care preferences in more diverse populations is required to understand the unique and common barriers, enablers, and needs, which would help to enable more widespread adoption of remote methods. This should include research in men, non–English speakers, the oldest old, those with poorer health and/or digital literacy, those with poorer access to health care and/or technology, and those with poor physical function. Furthermore, effective strategies to address technology access, digital literacy, and support needs for these vulnerable populations need to be explored. This may include reducing access barriers through device and internet connectivity provision, improving digital health literacy through ongoing human coaching and troubleshooting support, and co-designing content, interfaces, and delivery models with underserved communities [43].

The current study captures respondents' preferences and intentions, but not their behaviors or health outcomes. Future research is therefore needed to evaluate engagement with remote physical function interventions among older adults in the real world, as well as the effectiveness of such interventions for relevant health outcomes. The study was also focused on consumer perspectives and does not capture the perspectives of other stakeholders involved in delivering care to older adults (eg, health care professionals and policymakers), which are critical to translating research into practice. To address this, the RAMP working group has recently completed a Delphi process investigating the views of experts involved in the care of older adults (manuscript under review).

Conclusions

This international consumer Delphi process achieved consensus on 18 Delphi statements, which were synthesized into 7 recommendations for health care professionals, researchers, and policymakers to inform remote assessment and management of physical function in older adults. Further research on the feasibility and integration of remote delivery of physical function assessment and exercise programs is required, and this should be co-designed with older adults and other relevant stakeholders. Furthermore, given the recommendations reflect the sample of predominantly highly educated and digitally literate volunteers from high-income countries, further research is also required to explore their generalizability to more diverse older adult populations.

Acknowledgments

The Remote Assessment and Management of Physical Function in Older Adults Working Group gratefully acknowledges the contributions of all consumer participants of this Delphi process.

Data Availability

The study data may be made available upon reasonable request to the corresponding author.

Funding

This research was supported by a Deakin University Institute for Physical Activity and Nutrition (IPAN) Seed Grant (2023). DS is supported by an NHMRC Investigator Grant (GNT1174886). MS is supported by an Emerging Leader Fellowship from the Future Health Research & Innovation Fund, Department of Health, Western Australia. CH is supported by the National Institute for Health and Care Research (NIHR) Newcastle Biomedical Research Centre (reference: NIHR203309). The views expressed in this publication are those of the authors and do not necessarily reflect the views of the National Institute for Health and Care Research or the Department of Health and Social Care. MAL is supported by a T1 Canadian Research Chair.

Conflicts of Interest

DS has received speaker and consulting fees and an educational grant from Abbott, Amgen and Pfizer. RMD has received speaker and consulting fees and an educational grant from Abbott. No other authors have conflicts of interest to declare.

Multimedia Appendix 1
Study design of the RAMP Consumer Delphi Process.
[PDF File (Adobe PDF File), 271 KB - aging_v9i1e75791_app1.pdf]

Multimedia Appendix 2
RAMP Delphi Survey: Round 1. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 364 KB - aging_v9i1e75791_app2.pdf]

Multimedia Appendix 3
RAMP Round 1 results as presented to consumers. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 3165 KB - aging_v9i1e75791_app3.pdf]

Multimedia Appendix 4
Summary of changes to Round 1 Delphi statements with moderate or low agreement.
[PDF File (Adobe PDF File), 152 KB - aging_v9i1e75791_app4.pdf]

Multimedia Appendix 5
RAMP Delphi Survey: Round 2. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 222 KB - aging_v9i1e75791_app5.pdf]

Multimedia Appendix 6
SARC-F scores for RAMP participants in Round 1. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 80 KB - aging_v9i1e75791_app6.pdf]

Multimedia Appendix 7
Delivery preferences and perceived positives and negatives of remote tests and treatments for physical function among RAMP participants in Round 1. RAMP: Remote Assessment and Management of Physical Function in Older Adults.
[PDF File (Adobe PDF File), 112 KB - aging_v9i1e75791_app7.pdf]

References

1. Webber SC, Porter MM, Menec VH. Mobility in older adults: a comprehensive framework. Gerontologist 2010;50(4):443-450. [doi: 10.1093/geront/gnq013] [Medline: 20145017]
2. Ostchega Y, Harris TB, Hirsch R, Parsons VL, Kington R. The prevalence of functional limitations and disability in older persons in the US: data from the national health and nutrition examination survey III. J Am Geriatr Soc 2000;48(9):1132-1135. [doi: 10.1111/j.1532-5415.2000.tb04791.x] [Medline: 10983915]
3. Rosso AL, Taylor JA, Tabb LP, Michael YL. Mobility, disability, and social engagement in older adults. J Aging Health 2013;25(4):617-637 [FREE Full text] [doi: 10.1177/0898264313482489] [Medline: 23548944]
4. Hoffman JM, Ciol MA, Huynh M, Chan L. Estimating transition probabilities in mobility and total costs for medicare beneficiaries. Arch Phys Med Rehabil 2010;91(12):1849-1855 [FREE Full text] [doi: 10.1016/j.apmr.2010.08.010] [Medline: 21112425]
5. Offord NJ, Clegg A, Turner G, Dodds RM, Sayer AA, Witham MD. Current practice in the diagnosis and management of sarcopenia and frailty - results from a UK-wide survey. J Frailty Sarcopenia Falls 2019;4(3):71-77 [FREE Full text] [doi: 10.22540/JFSF-04-071] [Medline: 32300721]
6. Reijnierse EM, de van der Schueren MAE, Trappenburg MC, Doves M, Meskers CGM, Maier AB. Lack of knowledge and availability of diagnostic equipment could hinder the diagnosis of sarcopenia and its management. PLoS One 2017;12(10):e0185837 [FREE Full text] [doi: 10.1371/journal.pone.0185837] [Medline: 28968456]
7. Daly RM, Scott D, Kiss N, Tieland M, Baguley B, Fyfe JJ. Knowledge, awareness, behaviours, beliefs, attitudes, and perceptions of older Australians regarding muscle health and sarcopenia: a national survey. Arch Gerontol Geriatr 2025;135:105835 [FREE Full text] [doi: 10.1016/j.archger.2025.105835] [Medline: 40347780]
8. Izquierdo M, Merchant R, Morley J, Anker S, Aprahamian I, Arai H, et al. International exercise recommendations in older adults (ICFSR): expert consensus guidelines. J Nutr Health Aging 2021;25(7):824-853 [FREE Full text] [doi: 10.1007/s12603-021-1665-8] [Medline: 34409961]
9. Hurst C, Dismore L, Granic A, Tullo E, Noble JM, Hillman SJ, et al. Attitudes and barriers to resistance exercise training for older adults living with multiple long-term conditions, frailty, and a recent deterioration in health: qualitative findings from the lifestyle in later life - older people's medicine (LiLL-OPM) study. BMC Geriatr 2023;23(1):772 [FREE Full text] [doi: 10.1186/s12877-023-04461-5] [Medline: 38001414]
10. Nugent CD, Finlay D, Davies R, Donnelly M, Hallberg J, Black ND, et al. Remote healthcare monitoring and assessment. Technol Health Care 2011;19(4):295-306. [doi: 10.3233/THC-2011-0626] [Medline: 21849740]
11. Haimi M, Sergienko R. Adoption and use of telemedicine and digital health services among older adults in light of the COVID-19 pandemic: repeated cross-sectional analysis. JMIR Aging 2024;7:e52317 [FREE Full text] [doi: 10.2196/52317] [Medline: 38656768]
12. Jansons P, Dalla Via J, Daly R, Fyfe J, Gvozdenko E, Scott D. Delivery of home-based exercise interventions in older adults facilitated by amazon alexa: a 12-week feasibility trial. J Nutr Health Aging 2022;26(1):96-102 [FREE Full text] [doi: 10.1007/s12603-021-1717-0] [Medline: 35067710]
13. Fyfe JJ, Dalla Via J, Jansons P, Scott D, Daly RM. Feasibility and acceptability of a remotely delivered, home-based, pragmatic resistance 'exercise snacking' intervention in community-dwelling older adults: a pilot randomised controlled trial. BMC Geriatr 2022;22(1):521 [FREE Full text] [doi: 10.1186/s12877-022-03207-z] [Medline: 35751032]
14. Buckinx F, Aubertin-Leheudre M, Daoust R, Hegg S, Martel D, Martel-Thibault M, et al. Feasibility and acceptability of remote physical exercise programs to prevent mobility loss in pre-disabled older adults during isolation periods such as the COVID-19 pandemic. J Nutr Health Aging 2021;25(9):1106-1111 [FREE Full text] [doi: 10.1007/s12603-021-1688-1] [Medline: 34725669]

15. Brown RCC, Keating SE, Owen PJ, Jansons PS, McVicar J, Askew CD, et al. Client and clinician experiences and perspectives of exercise physiology services during the COVID-19 pandemic: qualitative study. J Med Internet Res 2023;25:e46370 [FREE Full text] [doi: 10.2196/46370] [Medline: 38127430]
16. Saaei F, Klappa SG. Rethinking telerehabilitation: attitudes of physical therapists and patients. J Patient Exp 2021;8:23743735211034335 [FREE Full text] [doi: 10.1177/23743735211034335] [Medline: 34377773]
17. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
18. Hsu CC, Sandford BA. The Delphi technique: making sense of consensus. Practical assessment, research, and evaluation 2007;12(1). [doi: 10.4018/978-1-4666-0074-4.ch011]
19. Zanker J, Sim M, Anderson K, Balogun S, Brennan-Olsen SL, Dent E, et al. Consensus guidelines for sarcopenia prevention, diagnosis and management in Australia and New Zealand. J Cachexia Sarcopenia Muscle 2023;14(1):142-156 [FREE Full text] [doi: 10.1002/jcsm.13115] [Medline: 36349684]
20. Zanker J, Sim M, Anderson K, Balogun S, Brennan-Olsen SL, Dent E, et al. The Australian and New Zealand society for sarcopenia and frailty research (ANZSSFR) sarcopenia diagnosis and management task force: findings from the consumer expert Delphi process. Australas J Ageing 2023;42(1):251-257 [FREE Full text] [doi: 10.1111/ajag.13164] [Medline: 36480154]
21. Barrington H, Young B, Williamson PR. Patient participation in Delphi surveys to develop core outcome sets: systematic review. BMJ Open 2021;11(9):e051066 [FREE Full text] [doi: 10.1136/bmjopen-2021-051066] [Medline: 34475183]
22. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. World J Methodol 2021;11(4):116-129 [FREE Full text] [doi: 10.5662/wjm.v11.i4.116] [Medline: 34322364]
23. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 2014;67(4):401-409. [doi: 10.1016/j.jclinepi.2013.12.002] [Medline: 24581294]
24. Malmstrom TK, Miller DK, Simonsick EM, Ferrucci L, Morley JE. SARC-F: a symptom score to predict persons with sarcopenia at risk for poor functional outcomes. J Cachexia Sarcopenia Muscle 2016;7(1):28-36 [FREE Full text] [doi: 10.1002/jcsm.12048] [Medline: 27066316]
25. WHO. Global Strategy on Digital Health 2020-2025. Geneva: World Health Organization; 2021.
26. Jansons P, Fyfe J, Via JD, Daly RM, Gvozdenko E, Scott D. Barriers and enablers for older adults participating in a home-based pragmatic exercise program delivered and monitored by Amazon Alexa: a qualitative study. BMC Geriatr 2022;22(1):248 [FREE Full text] [doi: 10.1186/s12877-022-02963-2] [Medline: 35337284]
27. Holliday AM, Hashmi AZ, Okoli-Umeweni AO, Khan A, Jindal SK, Gaur S, et al. American geriatrics society position statement: telehealth policy for older adults. J Am Geriatr Soc 2025;73(12):3646-3654. [doi: 10.1111/jgs.70004] [Medline: 40650623]
28. Beard JR, Officer A, de Carvalho IA, Sadana R, Pot AM, Michel J, et al. The world report on ageing and health: a policy framework for healthy ageing. Lancet 2016;387(10033):2145-2154 [FREE Full text] [doi: 10.1016/S0140-6736(15)00516-4] [Medline: 26520231]
29. Chan E, Samsudin SA, Lim YJ. Older patients' perception of engagement in functional self-care during hospitalization: a qualitative study. Geriatr Nurs 2020;41(3):297-304. [doi: 10.1016/j.gerinurse.2019.11.009] [Medline: 31787364]
30. Ong M, Soh K, Saimon R, Wai M, Mortell M, Soh K. Fall prevention education to reduce fall risk among community-dwelling older persons: a systematic review. J Nurs Manag 2021;29(8):2674-2688 [FREE Full text] [doi: 10.1111/jonm.13434] [Medline: 34331491]
31. Heslop P, Hurst C, Sayer A, Witham M. Remote collection of physical performance measures for older people: a systematic review. Age Ageing 2023;52(1):2023 [FREE Full text] [doi: 10.1093/ageing/afac327] [Medline: 36721962]
32. Buckinx F, Rezoulat M, Lefranc C, Reginster J, Bruyere O. Comparing remote and face-to-face assessments of physical performance in older adults: a reliability study. Geriatr Nurs 2024;55:71-78. [doi: 10.1016/j.gerinurse.2023.11.004] [Medline: 37976558]
33. Peyrusqué E, Granet J, Pageaux B, Buckinx F, Aubertin-Leheudre M. Assessing physical performance in older adults during isolation or lockdown periods: web-based video conferencing as a solution. J Nutr Health Aging 2022;26(1):52-56 [FREE Full text] [doi: 10.1007/s12603-021-1699-y] [Medline: 35067703]
34. Cohen-Mansfield J, Marx M, Biddison J, Guralnik J. Socio-environmental exercise preferences among older adults. Prev Med 2004;38(6):804-811. [doi: 10.1016/j.ypmed.2004.01.007] [Medline: 15193902]
35. Aaltonen S, Waller K, Vähä-Ypyä H, Rinne J, Sievänen H, Silventoinen K, et al. Motives for physical activity in older men and women: a twin study using accelerometer-measured physical activity. Scand J Med Sci Sports 2020;30(8):1409-1422 [FREE Full text] [doi: 10.1111/sms.13673] [Medline: 32259351]
36. Clemson L, Fiatarone Singh MA, Bundy A, Cumming RG, Manollaras K, O'Loughlin P, et al. Integration of balance and strength training into daily life activity to reduce rate of falls in older people (the LiFE study): randomised parallel trial. BMJ 2012;345:e4547 [FREE Full text] [doi: 10.1136/bmj.e4547] [Medline: 22872695]

37. Martel D, Lauzé M, Agnoux A, Fruteau de Laclos L, Daoust R, Émond M, et al. Comparing the effects of a home-based exercise program using a gerontechnology to a community-based group exercise program on functional capacities in older adults after a minor injury. Exp Gerontol 2018;108:41-47. [doi: 10.1016/j.exger.2018.03.016] [Medline: 29577975]
38. Rodrigues IB, Wang E, Keller H, Thabane L, Ashe MC, Brien S, et al. The moveStrong program for promoting balance and functional strength training and adequate protein intake in pre-frail older adults: a pilot randomized controlled trial. PLoS One 2021;16(9):e0257742 [FREE Full text] [doi: 10.1371/journal.pone.0257742] [Medline: 34559837]
39. Abreu F, Rodrigues A, Baptista F. Low-volume resistance training: a feasible, cost-effective strategy for musculoskeletal frailty in older adults attending daycare centers. Front Sports Act Living 2025;7:1542188 [FREE Full text] [doi: 10.3389/fspor.2025.1542188] [Medline: 40264931]
40. Pool J, Akhlaghpour S, Fatehi F, Gray LC. Data privacy concerns and use of telehealth in the aged care context: an integrative review and research agenda. Int J Med Inform 2022;160:104707. [doi: 10.1016/j.ijmedinf.2022.104707] [Medline: 35131698]
41. Gell N, Hoffman E, Patel K. Technology support challenges and recommendations for adapting an evidence-based exercise program for remote delivery to older adults: exploratory mixed methods study. JMIR Aging 2021;4(4):e27645 [FREE Full text] [doi: 10.2196/27645] [Medline: 34889743]
42. van Kessel R, Srivastava D, Kyriopoulos I, Monti G, Novillo-Ortiz D, Milman R, et al. Digital health reimbursement strategies of 8 European countries and israel: scoping review and policy mapping. JMIR Mhealth Uhealth 2023;11:e49003. [doi: 10.2196/49003] [Medline: 37773610]
43. Oudbier SJ, Souget-Ruff SP, Chen BSJ, Ziesemer KA, Meij HJ, Smets EMA. Implementation barriers and facilitators of remote monitoring, remote consultation and digital care platforms through the eyes of healthcare professionals: a review of reviews. BMJ Open 2024;14(6):e075833 [FREE Full text] [doi: 10.1136/bmjopen-2023-075833] [Medline: 38858155]

Abbreviations

RAMP: Remote Assessment and Management of Physical Function in Older Adults

Edited by D Liu, DRPH, MS, FGSA; submitted 13.Apr.2025; peer-reviewed by N Misser, N Paramita, R Miranda; comments to author 20.Oct.2025; revised version received 02.Nov.2025; accepted 15.Dec.2025; published 06.Feb.2026.

Please cite as:
Dent E, Hurst C, Dalla Via J, Fyfe JJ, Jansons P, Hayes EJ, Skinner G, Sim M, Aubertin-Leheudre M, Britting S, Buckinx F, Connolly G, Dignam R, Giangregorio L, Jones JRA, Kelly P, Kob R, Morin SN, Nandakumar G, Orssatto LBR, Pearson M, Pinto D, Reijnierse EM, Said CM, Salem M, Tan VPS, Tran R, Zanker J, Daly RM, Scott D
Developing Consumer Consensus on Remote Assessment and Management of Physical Function in Older Adults (RAMP): International Modified Delphi Process
JMIR Aging 2026;9:e75791
URL: https://aging.jmir.org/2026/1/e75791
doi:10.2196/75791
PMID:

Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study

Xiyao Yang[1*], MSc; Juan Ren[1*], MSc; Dan Su[2,3*], PhD; Manzhen Bao[1], MSc; Miao Zhang[1], MSc; Xiaoming Chen[1], BSc; Yanhua Li[1], BSc; Zonggui Wang[1], BSc; Xiujing Dai[1], BSc; Zengzeng Wei[1], BSc; Shuiyu Zhang[2], BSc; Yuxin Zhang[1], MSc; Juan Li[1], BSc; Xiaolin Li[1], BSc; Junjin Xu[1], BSc; Nan Mo[1], MSc

[1]Second Hospital of Anhui Medical University, No.678 Furong Road, Shushan District, Hefei, China

[2]School of Nursing, Anhui Medical University, Hefei, China

[3]The Taikang Health and Wellness Industry Research Institute, Anhui Medical University, Hefei, China

[*]these authors contributed equally

Corresponding Author:
Nan Mo, MSc
Second Hospital of Anhui Medical University, No.678 Furong Road, Shushan District, Hefei, China

Abstract

Background: Falls are one of the leading causes of injury or death among older adults. Falls occurring in individuals during hospitalization, as an adverse event, are a key concern for health care institutions. Identifying older adults at high risk of falls in clinical settings enables early interventions, thereby reducing the incidence of falls.

Objective: This study aims to develop and validate machine learning models to predict the risk of falls among hospitalized older adults.

Methods: This study retrospectively analyzed data from a tertiary general hospital in China, including 342 older adults who experienced falls and 684 randomly matched nonfallers, between January 2018 and December 2024, encompassing demographic information, comorbidities, laboratory parameters, and medication use, among other variables. The dataset was randomly split into training and testing sets in a 7:3 ratio. Predictors were selected from the training set using stepwise regression, least absolute shrinkage and selection operator, and random forest-recursive feature elimination. Seven machine learning algorithms were employed to develop predictive models in the training set, and their performance was compared in the testing set. The optimal model was interpreted using Shapley Additive Explanations (SHAP).

Results: The gradient boosting machine model demonstrated the best predictive performance (C-index 0.744, 95% CI 0.688 - 0.799). The 8 most important variables associated with fall risk were dizziness, epilepsy, fall history within the past 3 months, use of walking assistance, emergency admission, Morse Fall Scale scores, modified Barthel Index scores, and the number of indwelling catheters. The model was interpreted using SHAP to enhance the clinical utility of the predictive model.

Conclusions: The gradient boosting machine model was identified as the optimal predictive model. The SHAP method enhanced its integration into clinical workflows.

(JMIR Aging 2026;9:e80602) doi:10.2196/80602

KEYWORDS

machine learning; risk prediction; older adults; fall prediction; gradient boosting machine; random forest

Introduction

Falls, the second leading cause of global unintentional injury deaths, are a significant public health concern. They are defined as "an event that leads to a person inadvertently coming to rest on the ground, floor, or other lower surface than their original position [1]." Age is one of the main risk factors for falls [2], and statistics indicate that the incidence of falls among older adults is approximately 26.5% [3]. Among individuals aged >60 years globally, falls are one of the most common causes of injury or death, with one out of every 5 falls resulting in a fracture or head injury [2,4]. In addition, falls generate substantial medical costs, imposing a heavy economic burden worldwide [5].

Notably, falls are adverse events in hospitals, and the prevention of falls is also a priority for improving the quality of nursing care [6]. The incidence of falls in hospitals is typically in the range of 2 to 16 per 1000 bed days [7,8]. Despite a declining incidence of falls among hospitalized older adults, the increasing number of older adults admitted to hospitals, driven by an expanding aging population, suggests that falls prevention will remain a critical concern in hospitals [8,9]. Falls are preventable adverse events in hospitals, and implementing fall prevention programs can avoid costs of US $14,600 per 1000 patient-days [6]. Therefore, identifying individuals at high risk of falls in hospitals to take preventive measures is particularly important, especially among older adults.

The MFS (Morse Fall Scale) and STRATIFY (St. Thomas's Risk Assessment Tool in Falling Elderly Inpatients), widely used in hospitals to identify individuals at high risk of falls, have drawbacks such as low specificity [10,11]. Several studies have developed predictive models for fall risk in older inpatients [12-17]. While some employed traditional regression methods [12-15,17], these conventional approaches often struggle with complex, multidimensional data [18]. Other models exhibit limited applicability, being restricted to specific clinical settings or units [12,13,17]. Additionally, certain models rely solely on clinical texts for prediction, a methodology constrained by single-variable limitations that compromise performance [14].

In recent years, machine learning (ML) algorithms have attracted considerable interest in health care predictive modeling due to their capacity to develop highly accurate prediction models at low cost [19]. The capacity of ML algorithms to process high-dimensional data not only enhances the accuracy and efficiency of predictive models but also enables personalized risk prediction [4,20]. Although existing studies have employed ML algorithms to develop fall prediction models for hospitalized older adults, these models exhibit limitations, including suboptimal performance, applicability restricted to specific geriatric subpopulations, and reliance on environmental detection systems that hinder their widespread clinical adoption [16,21-23]. Critically, limited studies have offered comprehensive explanations or analyses of model predictions, restricting clinical applicability and diminishing the practical value of these models.

Therefore, the objective of our study is to develop and validate multiple ML models utilizing clinically accessible data to predict fall risk of hospitalized older adults. We seek to identify the optimal model while interpreting its predictions through the Shapley Additive Explanations (SHAP) method.

Methods

Data Source and Participants

Using an adverse event reporting system integrated into electronic nursing workstations of a tertiary general hospital, researchers retrieved fall incident records for hospitalized older adults (aged ≥60 y) occurring between January 2018 and December 2024, extracting hospitalization identifiers and fall timestamps. An electronic health record (EHR) system was used to record admission and discharge dates along with hospitalization identifiers for older adults without a history of falls hospitalized between January 2018 and December 2024. The fall timestamp of each case patient was used to anchor the index time for the matched controls. For each case, 2 controls were matched. Specifically, we first preprocessed the data by removing duplicate records from individuals with multiple hospitalizations (retaining only the first admission). From this refined pool of potential controls, we then used a Visual Basic algorithm in Microsoft Excel (version 16.0) to identify patients whose entire hospitalization period (from admission to discharge) encompassed the fall timestamp. This approach ensures that both cases and controls were exposed to similar time-dependent clinical factors at the same specific time point, thereby minimizing potential time-dependent bias. The case and control groups were not matched on demographics such as age or gender in order to maintain the natural distribution found in real-world clinical settings. With the aim of capturing all relevant information, variables with clinical or predictive relevance were included as model features for the ML algorithm to parse their associations with the outcome.

Matched controls identified as day cases were excluded and replaced until a 1:2 case-control ratio was maintained. This ratio was selected based on considerations of statistical power, cost-effectiveness, and practical constraints, as increasing the control-to-case ratio beyond 2:1 yields diminishing returns in power while substantially increasing costs and workload [24,25]. Cases were initially identified from the adverse event reporting system as any patient with a documented fall event occurring within the hospital premises and were excluded if they were aged <60 years old at admission, experienced subsequent falls occurring during the same hospitalization, or were nonhospitalized patients or day cases. Controls were selected from the EHR system as hospitalized patients aged ≥60 years with no record of an in-hospital fall and were excluded for having duplicate admission records (only the first was retained), day-case status, or if they could not be matched to a case. We excluded day case patients because more than 20% of the data were missing in EHR. The sample size was estimated using the "pmsampsize" package in R software (version 4.5.0). According to other researchers, the c-statistic is 0.73, the number of predictor parameters chosen for our study is 17, and the prevalence is 0.33 (1/3), with a required sample size of 992 for the calculation.

Ethical Considerations

The study was approved by the Ethics Committee Board of the Second Affiliated Hospital of Anhui Medical University (approval number: YX2025-162). This study adheres strictly to privacy protection principles. Nonessential identifying information is omitted during data processing, and informed consent is obtained when necessary. Informed consent was waived for patients who died or were disconnected. No financial compensation is provided to participants. This study conforms to the principles outlined in the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis) statement.

Data Collection and Processing

All records containing timestamps and hospitalization identifiers were randomly split into 2 datasets. Two uniformly trained data collectors independently extracted variables through the EHR system using these identifiers, followed by cross-verification upon completion. Five categories of variables were collected: demographic characteristics, comorbidities, medications, laboratory indicators, and other variables. Table S1 in

Multimedia Appendix 1 provides the list of 64 extracted variables. Demographic characteristics and sleep duration data were extracted from hospital admission records. The absence of BMI values was directly attributable to practical barriers in anthropometric data collection for patients with mobility limitations (bedridden or wheelchair-dependent status). Comorbidities were identified by integrating inpatient diagnoses from admission summaries with discharge diagnoses in corresponding discharge records. Medication administration records were retrieved from both permanent and temporary medical orders to capture all medications administered within the 24-hour period preceding the timestamp. Polypharmacy was defined as taking 5 or more medications daily. Laboratory indicators were collected from laboratory test reports. For indicators with repeated measurements, data within the 7 days before and after the timestamp were selected for analysis. The remaining variables were extracted from nursing records within 1 week before and after the timestamp. Given that at least 2 nursing records are documented weekly, there is no missing data for these variables. A total of 64 variables were initially extracted. With 27.49% missing values, BMI was removed from analysis. For the remaining variables, only albumin and hemoglobin contained missing values (0.03% and 0.04%, respectively). The missing values for albumin and hemoglobin were imputed using the random forest (RF) imputation method, implemented via the “missForest” package in R software (version 4.5.0). This approach offers the advantage of handling mixed data types (continuous and categorical) and effectively capturing nonlinear relationships among variables [26].

Feature Selection

The dataset was randomly split into a training set (70%) and a testing set (30%). A three-step selection strategy was implemented in the training set to identify optimal predictors. First, univariate (LR) was applied for preliminary screening (P<.05) to retain statistically significant variables. Second, 5 feature selection methods were integrated: stepwise regression (SR) comprises 3 variants—forward selection, backward selection, and bidirectional elimination; least absolute shrinkage and selection operator (LASSO); and random forest-recursive feature elimination (RF-RFE). Predictors were determined by the overlap among the results of these methods. This approach aimed to mitigate high correlation among predictors while capturing their complex relationships with the outcome variable [27]. SR iteratively adjusts variables based on statistical significance, LASSO addresses high dimensionality and multicollinearity while preventing overfitting, and RF-RFE captures nonlinear patterns and variable interactions. Both LASSO and RF-RFE incorporated 10-fold cross-validation. Finally, clinical experts validated the selected predictors to ensure clinical applicability.

Models Development and Validation

To comprehensively evaluate predictive performance and ensure robust results, we employed multiple algorithms to construct predictive models in the training set, including seven ML models: LR, support vector machines, RF, gradient boosting machine (GBM), extreme gradient boosting (XGBoost), k-nearest neighbor (KNN), and neural network (NN). Grounded in distinct modeling philosophies, each algorithm offers unique advantages. LR establishes an optimal linear decision boundary, valued for its conceptual simplicity and high interpretability, serving as a reliable performance benchmark [20]. Support vector machines aim to determine a separating hyperplane that maximizes the geometric margin for robust classification. They address nonlinear problems by employing kernel functions to project data into a higher-dimensional feature space where the maximum-margin principle is applied [20]. As a representative bagging ensemble, RF enhances predictive stability and captures complex feature interactions by aggregating numerous decorrelated decision trees, also providing inherent resistance to overfitting and enabling feature importance evaluation [20,28]. GBM employs a sequential modeling strategy that iteratively corrects errors from preceding models, often achieving high predictive accuracy [20]. XGBoost, an optimized implementation of gradient boosting, incorporates regularization and advanced algorithmic techniques to further improve computational efficiency and performance [20]. KNN is an instance-based learning method operating on the principle of local similarity. Predictions are derived from the majority label or average value of a sample’s KNNs in the feature space, offering an intuitive perspective on the local data structure [29]. NN, or deep learning models, function as universal approximators by leveraging multiple layers of tunable nonlinear transformations. This architecture enables them to automatically learn hierarchical data representations and extract complex, high-level features through training [30]. This systematic selection of algorithms, encompassing linear models, kernel methods, bagging and boosting ensembles, instance-based learning, and NNs, ensures our evaluation is comprehensive and avoids bias toward any single modeling strategy.

To mitigate class imbalance, we applied random upsampling to the training dataset, which involves duplicating instances from the minority class at random to balance the class distribution. Subsequently, to rigorously tune hyperparameters and guard against overfitting, we performed a grid search with 10-fold cross-validation on this processed training set to identify the optimal parameters. The test set was used to evaluate model performance. The area under the receiver operating characteristic curve (AUROC) in the testing set served as the primary metric for assessing discriminative ability. Model discrimination was primarily assessed using the AUROC. This metric is considered a standard method for evaluating ranking ability, as it provides a threshold-independent assessment of a model’s inherent discriminative power [31]. Additionally, model performance was comprehensively evaluated using the area under the precision-recall curve (AUPRC), which is particularly informative for imbalanced datasets, along with sensitivity, specificity, accuracy, recall, F_1-score, positive predictive value, and negative predictive value. Calibration curves were plotted to assess prediction accuracy. Decision curve analysis (DCA) was performed to quantify clinical utility. SHAP is a model interpretation tool that calculates feature contribution values to provide both global (model-level) and local (individual prediction) explanations, making models more interpretable and applicable [20,32].

Therefore, we employed the SHAP method to elucidate how individual features influence fall risk predictions in hospitalized older adults within the optimal model.

Statistical Analysis

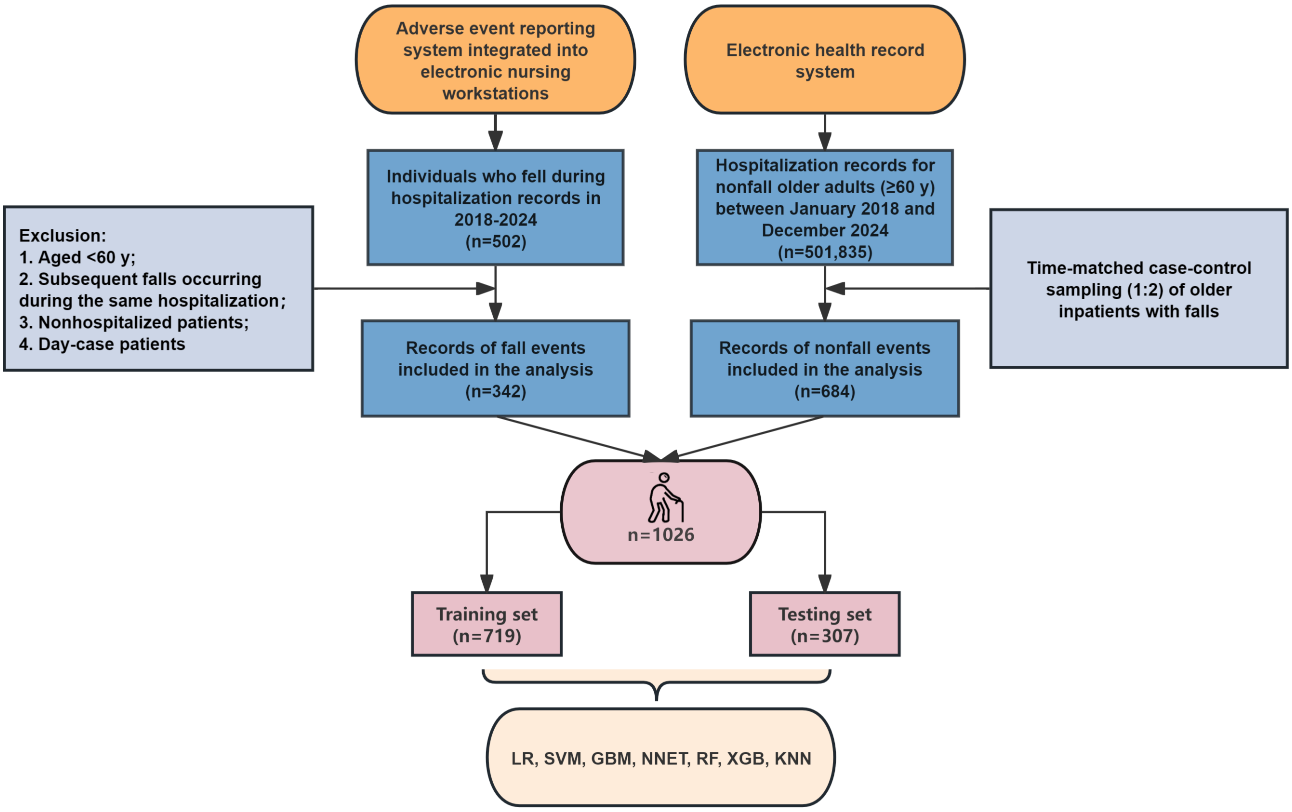
All statistical analyses were performed using R software (version 4.5.0), with categories merged when necessary to address sparse data. Use of walking assistance (UWA) was classified into 4 groups: no assistance, wheelchair or bedridden, support by others or furniture, and walker/crutches/cane. Continuous variables were categorized as follows: age into 60 to 69, 70 to 79, and ≥80 years; serum albumin into ≥34 and <34 g/L [33]; MFS scores into <45 points and ≥45 points [34]; modified Barthel Index (mBI) [35] scores into 0 to 20 points, 21 to 60 points, 61 to 90 points, 91 to 99 points, and 100 points; Nutritional Risk Screening 2002 (NRS 2002) scores into <3 points (no nutritional risk) and ≥3 points (at risk) [36]; and Numeric Pain Rating Scale scores into 0 points (no pain) and ≥1 point (pain) [37]. Continuous variables, none of which followed a normal distribution, were expressed as medians and IQR (M, Q1-Q3). Categorical variables were reported as numbers and percentages (n, %). Differences between groups were analyzed using the Mann-Whitney *U* test for nonnormally distributed continuous variables and the Chi-square test (or Fisher exact test for sparse data) for categorical variables. A 2-sided $P<.05$ was considered statistically significant.

Results

Baseline Characteristics

Ultimately, 1026 older adults were included in the study. Figure 1 illustrates the process of patient screening. The comparison between fallers and nonfallers in the overall dataset is presented in Table S2 in Multimedia Appendix 1. Among the 1026 patients, 40.84% (419/1026) were aged 60 to 69 years and 55.65% (571/1026) were male. Among the 342 fallers, 40.06% (137/342) were aged 70 to 79 years and 52.05% (178/342) were male. Significant differences were observed between fallers and nonfallers in the following variables: age, blood pressure, diabetes, chronic kidney disease, heart failure, hypothyroidism, cancer, Parkinson's disease, dizziness, stroke, gait abnormality, epilepsy, visual impairment, hearing impairment, polypharmacy, antiplatelet drugs, statins, α-blockers, vasodilators, antidiabetic drugs, anti-Parkinson's disease drugs, antiepileptic drugs, benzodiazepines, Z-drugs, albumin levels, fall history in the past 3 months, UWA, emergency admission (EA), MFS scores, NRS 2002 scores, mBI scores, number of indwelling catheters (Indw Cath), and departments.

Figure 1. Flowchart of the patient screening. GBM: gradient boosting machine; KNN: k-nearest neighbor; LR: logistic regression; NNET: neural network; RF: random forest; SVM: support vector machine; XGB: extreme gradient boosting.



Selection of Predictor Variables

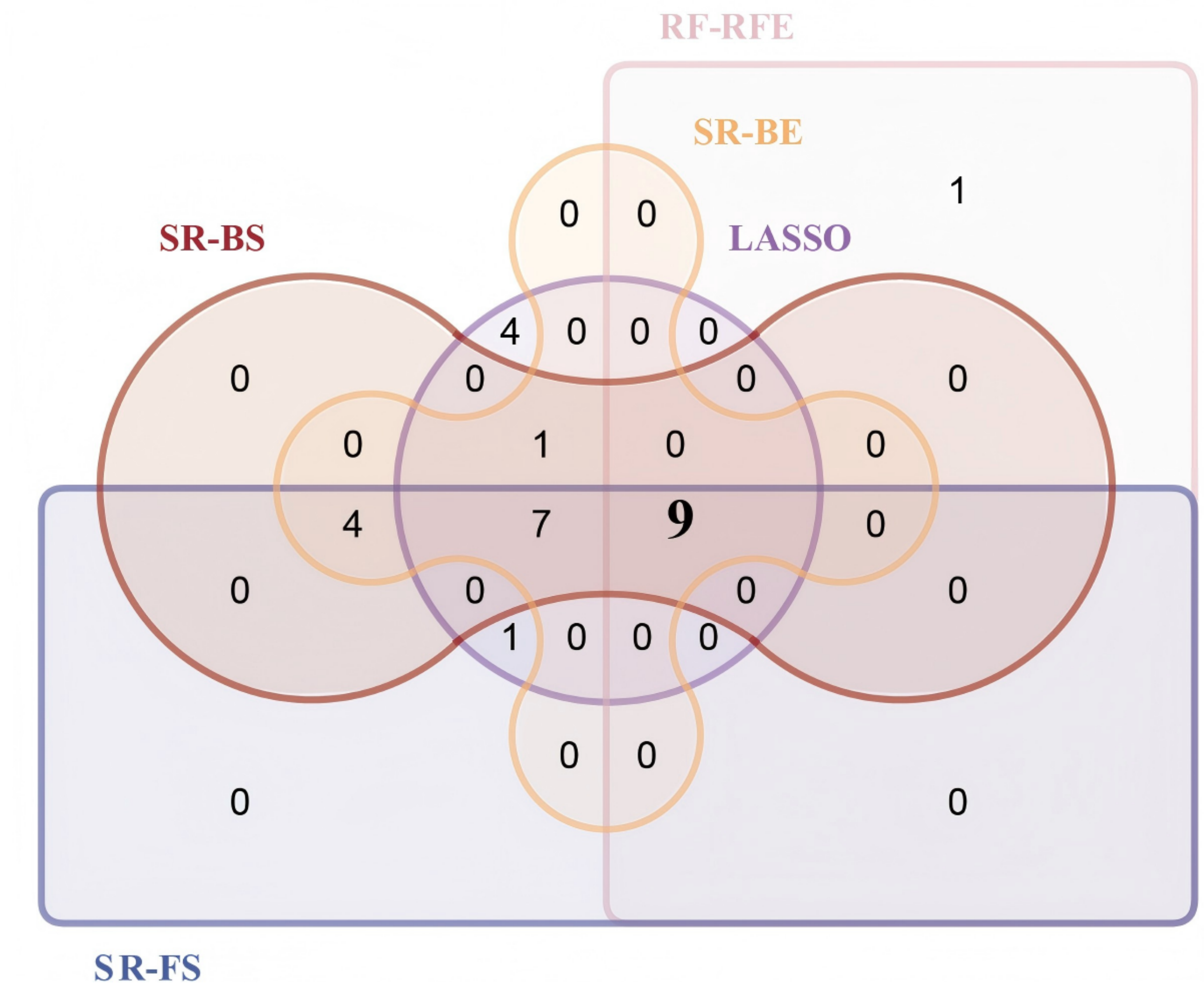
Univariate LR identified 27 potential predictors ($P<.05$) in the training set, as detailed in Table S3 of Multimedia Appendix 1. Table 1 displays the predictors identified by the 5 methods (SR-forward selection, SR-backward selection, SR-bidirectional elimination, LASSO, and RF-RFE). Table S4, Figure S1, and Figure S2 in Multimedia Appendix 1 provide detailed information. Figure 2 visualizes the overlap of predictors selected across 5 methods. The intersecting predictors from these methods formed the final predictor set, comprising dizziness, epilepsy, fall history in the past 3 months, UWA, EA, MFS scores, mBI scores, and Indw Cath. After expert consultation, no predictors were added or removed. The final development model included these 8 predictor variables.

Table . The predictors obtained through 5 selection methods.

Methods	Number of predictors (categories)	Predictor variables
SR-FS	15 (21)	Hypothyroidism, OP[a], dizziness, stroke, epilepsy, polypharmacy, AC[b], BZDs[c], Alb[d], FH-3M[e], UWA.1[f], UWA.2[g], UWA.3[h], EA[i], MFS[j] scores, mBI.1[k], mBI.2[l], mBI.3[m], mBI.4[n], Indw Cath.1[o], Indw Cath.2[p]
SR-BS	15 (21)	Hypothyroidism, OP, dizziness, stroke, epilepsy, AC, BZDs, Zdrugs, Alb, FH-3M, UWA.1, UWA.2, UWA.3, EA, MFS scores, mBI.1, mBI.2, mBI.3, mBI.4, Indw Cath.1, Indw Cath.2
SR-BE	15 (21)	Hypothyroidism, OP, dizziness, stroke, epilepsy, AC, BZDs, Zdrugs, Alb, FH-3M, UWA.1, UWA.2, UWA.3, EA, MFS scores, mBI.1, mBI.2, mBI.3, mBI.4, Indw Cath.1, Indw Cath.2
LASSO	19 (22)	Gender, hypothyroidism, OP, dizziness, stroke, epilepsy, polypharmacy, AC, antidiabetics, BZDs, Zdrugs, Alb, FH-3M, UWA.1, UWA.2, UWA.3, EA, MFS scores, mBI.0[q], mBI.4, Indw Cath.2, Department.2[r]
RF-RFE	9 (10)	CA[s], dizziness, epilepsy, FH-3M, UWA.2, UWA.3, EA, MFS scores, mBI.4, Indw Cath.2

[a]OP: osteoporosis.
[b]AC: anticoagulants.
[c]BZDs: benzodiazepines.
[d]Alb: albumin.
[e]FH-3M: fall history in the past 3 months.
[f]UWA.1: use of walking assistance category 1 (wheelchair or bedridden).
[g]UWA.2: use of walking assistance category 2 (support by others or furniture).
[h]UWA.3: use of walking assistance category 3 (walker/crutches/cane).
[i]EA: emergency admission.
[j]MFS: Morse fall scale.
[k]mBI.1: modified Barthel Index scores category 1 (21-60 points).
[l]mBI.2: modified Barthel Index scores category 2 (61-90 points).
[m]mBI.3: modified Barthel Index scores category 3 (91-99 points).
[n]mBI.4: modified Barthel Index scores category 4 (100 points).
[o]Indw Cath.1: number of indwelling catheters 1 (1).
[p]Indw Cath.2: number of indwelling catheters 2 (≥2).
[q]mBI.0: modified Barthel Index scores category 0 (0-20 points).
[r]Department.2: department category 2 (department of rehabilitation medicine).
[s]CA: cancer.

Figure 2. Upset plot of the overlap of predictors selected across 5 methods. BE: bidirectional elimination; BS: backward selection; FS: forward selection, LASSO: least absolute shrinkage and selection operator; RF-RFE: random forest-recursive feature elimination; SR: stepwise regression.



Models Development and Validation

Table S5 in Multimedia Appendix 1 compares the characteristics of the training and testing sets. The training set comprised 719 (70%) older adults, while the testing set included 307 (30%). All final model predictors and fall status (yes/no) were balanced between the training and testing sets, as shown in Table 2. The AUROC for the 7 models in the testing set is shown in Figure 3B. Among these, the GBM model demonstrated the highest discrimination with an AUROC of 0.744 (95% CI 0.688 - 0.799) compared to the other 6 models. The LR model followed closely with an AUC of 0.742 (95% CI 0.685 - 0.798). The NN and RF models had the lowest AUROCs, at 0.705 (95% CI 0.646 - 0.765) and 0.715 (95% CI 0.657 - 0.772), respectively. Table 3 displays the detailed predictive performance of the 7 ML models. In the testing set, the LR model achieved the highest AUPRC of 0.570 (0.475 - 0.663), while the RF model showed the lowest AUPRC of 0.477 (0.386 - 0.580). Regarding other performance metrics, the NN models had the best sensitivity (0.931), the XGBoost model had the best specificity (0.644), and the LR model showed the highest accuracy (0.687). The calibration curves for the predictive models in the testing set are shown in Figure 3D. The LR model demonstrated the best calibration ability in the testing set. The DCA curves for the predictive models in the testing set are shown in Figure 3F. The DCA curves suggest that the 7 models have certain clinical utility, generating net benefits within the threshold range of 0 to 0.5. Considering AUROC, sensitivity, and specificity, the GBM was determined to be the best-performing model.

Table . Characteristics of the predictors in the training and testing sets.

Predictors	Training set (n=719)	Testing set (n=307)	*P* value
Fall, n (%)			>.99
No	479 (66.62)	205 (66.78)	
Yes	240 (33.38)	102 (33.22)	
Dizziness, n (%)			.93
No	705 (98.05)	302 (98.37)	
Yes	14 (1.95)	5 (1.63)	
Epilepsy, n (%)			>.99
No	706 (98.19)	302 (98.37)	
Yes	13 (1.81)	5 (1.63)	
FH-3M[a], n (%)			.99
No	660 (91.79)	281 (91.53)	
Yes	59 (8.21)	26 (8.47)	
UWA[b], n (%)			.08
No assistance	475 (66.06)	179 (58.31)	
Wheelchair or bedridden	192 (26.7)	100 (32.57)	
Support by others or furniture	10 (1.39)	3 (0.98)	
Walker/crutches/cane	42 (5.84)	25 (8.14)	
EA[c], n (%)			.64
No	701 (97.50)	297 (96.74)	
Yes	18 (2.50)	10 (3.26)	
MFS[d] (points), n (%)			.42
<45	196 (27.26)	92 (29.97)	
≥45	523 (72.74)	215 (70.03)	
mBI[e] (points), n (%)			.24
0-20	36 (5.01)	13 (4.23)	
21-60	172 (23.92)	93 (30.29)	
61-90	322 (44.78)	121 (39.41)	
91 - 99	79 (10.99)	30 (9.77)	
100	110 (15.3)	50 (16.29)	
Indw Cath[f], n (%)			.81
0	566 (78.72)	247 (80.46)	
1	95 (13.21)	38 (12.38)	
≥2	58 (8.07)	22 (7.17)	

[a]FH-3M: fall history in the past 3 months.
[b]UWA: use of walking assistance.
[c]EA: emergency admission.
[d]MFS: Morse fall scale.
[e]mBI: modified Barthel Index.
[f]Indw Cath: number of indwelling catheters.

Figure 3. Receiver operating characteristic (ROC) curves, calibration curves, and decision curve analysis (DCA) curves of different machine learning (ML) models in the training and testing sets. (A) ROC curves and area under the ROC curve (AUC) values of different ML prediction models in the training set. (B) ROC curves and AUC values of different ML prediction models in the testing set.(C) Calibration curves of different ML prediction models in the training set. (D) Calibration curves of different ML prediction models in the testing set. (E) DCA curves of different ML prediction models in the training dataset. (F) DCA curves of different ML prediction models in the testing dataset. GBM: gradient boosting machine; KNN: k-nearest neighbor; LR: logistic regression; NNET: neural network; RF: random forest; SVMs: support vector machines; XGBoost: extreme gradient boosting.

Table . The performance of 7 machine learning models for predicting falls in hospitalized older adults.

Model	AUPRC[a] (95% CI)	Sensitivity	Specificity	F_1-score	Accuracy	PPV[b]	NPV[c]	Recall
Training set								
LR[d]	0.614 (0.555 - 0.672)	0.708	0.656	0.591	0.673	0.508	0.818	0.708
SVMs[e]	0.599 (0.534 - 0.661)	0.767	0.585	0.591	0.645	0.480	0.833	0.767
GBM[f]	0.620 (0.559 - 0.681)	0.654	0.714	0.588	0.694	0.534	0.805	0.654
NN[g]	0.646 (0.590 - 0.705)	0.629	0.741	0.586	0.704	0.549	0.800	0.629
RF[h]	0.580 (0.517 - 0.643)	0.771	0.608	0.604	0.662	0.496	0.841	0.771
XGBoost[i]	0.637 (0.579 - 0.696)	0.654	0.718	0.590	0.697	0.538	0.806	0.654
KNN[j]	0.626 (0.567 - 0.685)	0.783	0.553	0.586	0.630	0.468	0.836	0.783
Testing set								
LR	0.570 (0.475 - 0.663)	0.794	0.634	0.628	0.687	0.519	0.861	0.794
SVMs	0.537 (0.437 - 0.640)	0.873	0.561	0.634	0.665	0.497	0.898	0.873
GBM	0.560 (0.464 - 0.654)	0.873	0.561	0.634	0.665	0.497	0.898	0.873
NN	0.509 (0.409 - 0.610)	0.931	0.424	0.603	0.593	0.446	0.926	0.931
RF	0.477 (0.386 - 0.580)	0.863	0.576	0.635	0.671	0.503	0.894	0.863
XGBoost	0.535 (0.435 - 0.635)	0.745	0.644	0.606	0.678	0.510	0.835	0.745
KNN	0.547 (0.446 - 0.643)	0.794	0.590	0.607	0.658	0.491	0.852	0.794

[a]AUPRC: area under the precision recall curve.

[b]PPV: positive predictive value.

[c]NPV: negative predictive value.

[d]LR: logistic regression.

[e]SVM: support vector machine.

[f]GBM: gradient boosting machine.

[g]NN: neural network.

[h]RF: random forest.

[i]XGBoost: extreme gradient boosting.

[j]KNN: k-nearest neighbor.

Interpretability Analysis

SHAP was utilized to illustrate how the features predict the occurrence of falls in old adults during hospitalization within the GBM model. Figure 4A displays the 17 features sorted by their average absolute SHAP values, and higher absolute SHAP indicates greater contribution to fall risk. Figure 4B shows the impact values and explanations of these features, and yellow dots represent high risk, while purple dots indicate low risk. An MFS score of ≥45, an mBI score that is not 100 points, an mBI score not between 0 and 20 points, having fewer than 2 indwelling tubes, a history of falls in the past 3 months, EA, epilepsy, dizziness, use of a walker/cane/crutch, requiring assistance from others/furniture for walking, and not using a wheelchair or not being bedridden are associated with a higher risk of falls in old adults during hospitalization. Beyond global SHAP interpretations, local interpretability was demonstrated. Figure 5A and B visualizes how the GBA model makes predictions about falls in older adults during hospitalization; yellow arrows indicate risk-increasing features and purple arrows risk-decreasing features. The *f(x)* values inside arrows quantify each feature's contribution. Summing these yields the model's final prediction, which is represented by the *f(x)* value outside arrows.

Figure 4. Interpretation of the gradient boosting machine model by the Shapley Additive Explanations (SHAP) method. (A) A bar summary of the most important features according to the SHAP values. (B) Summary and explanation of the most influential features. Yellow dots indicate high-risk values, and purple dots indicate low-risk values. EA: emergency admission; FH-3M: fall history in the past 3 months; Indw Cath.0: number of indwelling catheters 0 (0); Indw Cath.1: number of indwelling catheters 1 (1); Indw Cath.2: number of indwelling catheters 2 (≥2); mBI.0: modified Barthel Index scores category 0 (0-20 points); mBI.1: modified Barthel Index scores category 1 (21-60 points); mBI.2: modified Barthel Index scores category 2 (61 - 90 points); mBI.3: modified Barthel Index scores category 3 (91 - 99 points); mBI.4: modified Barthel Index scores category 4 (100 points); MFS: Morse Fall Scale; UWA.0: use of walking assistance category 0 (no assistance); UWA.1: use of walking assistance category 1 (wheelchair or bedridden); UWA.2: use of walking assistance category 2 (support by others or furniture); UWA.3: use of walking assistance category 3 (walker/crutches/cane).

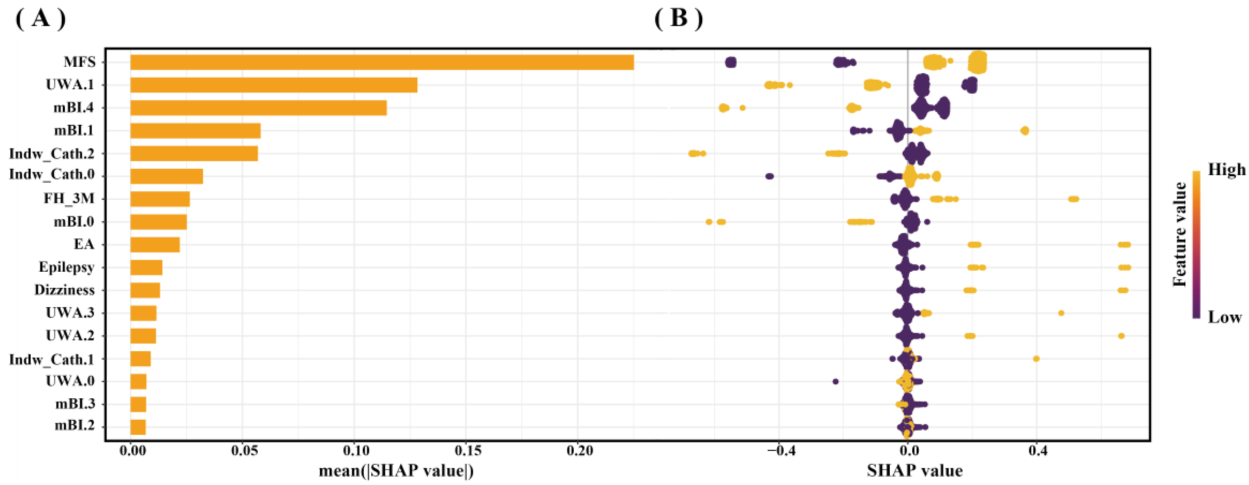
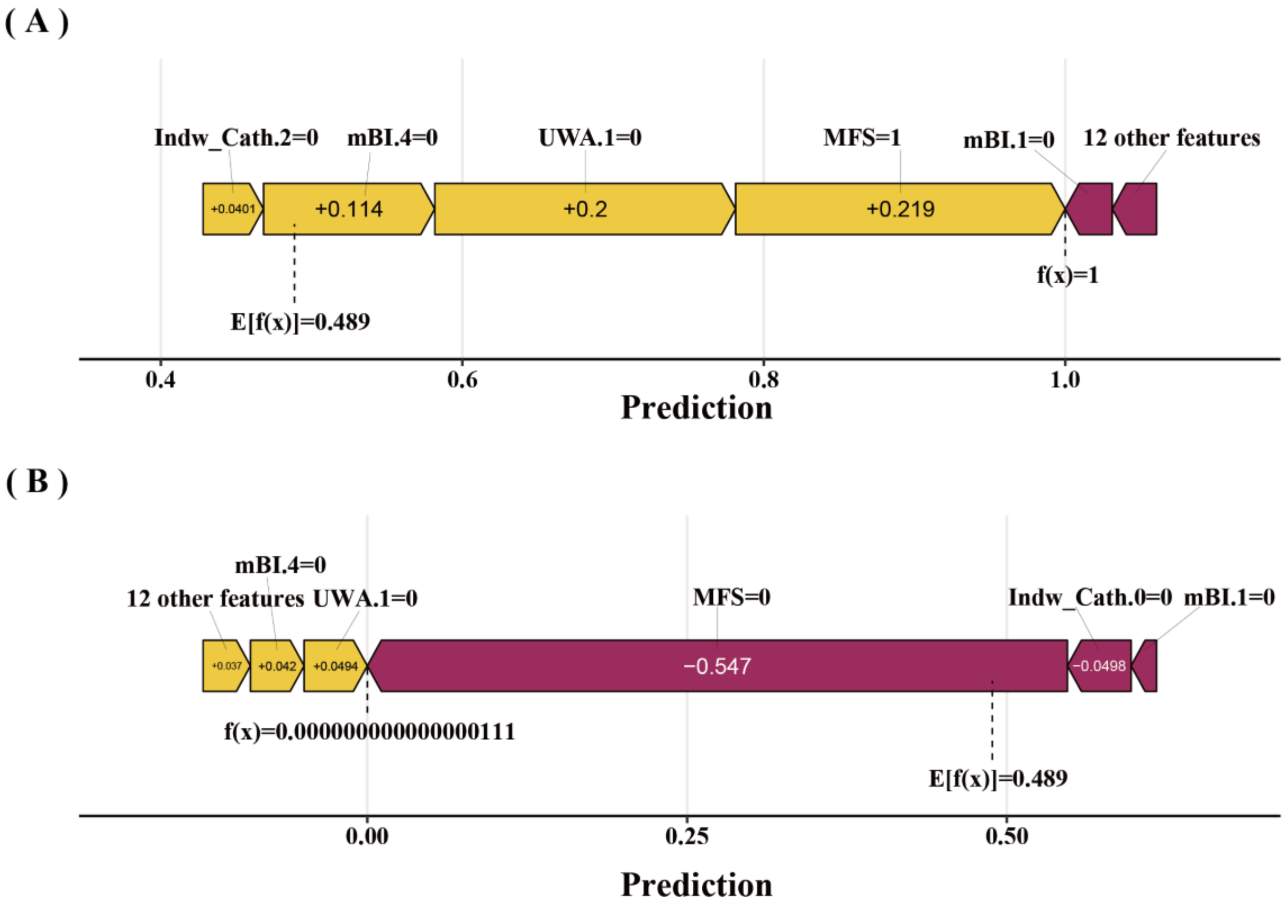


Figure 5. Compositional risk of feature contributions for predicting the occurrence of falls in 2 older adults during hospitalization. Purple arrows denote factors that decrease the risk of falls for old adults during hospitalization, while yellow arrows indicate features that increase the risk. (A) An older adult with fewer than 2 indwelling catheters, a modified Barthel Index (mBI) score not equal to 100 and not within the range of 21 - 60 points, not using a wheelchair or bedridden, and a Morse Fall Scale (MFS) score ≥45 points experienced a fall during hospitalization. (B) An older adult with an mBI score not equal to 100 or within the range of 21 - 60 points, not using a wheelchair or bedridden, with the presence of indwelling catheters, and an MFS score <45 points did not experience a fall during hospitalization. Indw Cath.0: number of indwelling catheters 0 (0); Indw cath.2: number of indwelling catheters 2 (≥2); mBI.1: modified Barthel Index scores category 1 (21 - 60 points); mBI.4: modified Barthel Index scores category 4 (100 points); UWA.1: use of walking assistance category 1 (wheelchair or bedridden).



Discussion

Principal Results

In this study, we used 7 ML algorithms to predict in-hospital falls among hospitalized older adults based on clinically accessible data, including demographic characteristics, comorbidities, laboratory parameters, and medications. The GBM algorithm demonstrated the optimal predictive performance. Model interpretability was achieved at both global and individualized patient levels using SHAP [38]. The SHAP approach bridges the gap between ML models and realistic clinical decision-making, enabling health care providers to understand the model's predictive process and trust its predictive power [39].

In our study, multiple ML algorithms based on distinct principles were employed to develop predictive models. However, the predictive performance across these methods showed limited variation, consistent with prior research [20], which suggests that advanced ML algorithms generally perform well on relatively small and low-dimensional datasets. Through a comprehensive evaluation of the AUROC, AUPRC, sensitivity, and specificity, the GBM model was ultimately selected as the optimal model. Notably, LR also demonstrated competitive performance, and DCA curves indicated that the LR model could provide favorable clinical net benefit. Nevertheless, compared to LR, GBM offers distinct advantages in handling nonlinear relationships and complex data patterns [27].

Fall risk factors among hospitalized older adults encompass multiple domains, including demographic characteristics, comorbidities, laboratory parameters, and medications [40-42]. Previous studies have relied on subjective nursing documentation texts, comprehensive geriatric assessments, or environmental detection systems to develop fall prediction models for hospitalized individuals [16,21], which limits their clinical utility. Identifying predictors is a critical step in building predictive models. It is notable that the predictors identified in our study are aligned with routinely collected clinical data, ensuring practical accessibility in health care settings. Conventional approaches often select predictors using a single method, such as regression models, whereas combining multiple feature selection techniques may yield simplified models with higher generalizability [43]. Different from previous studies, we used multiple methods such as SR, RF-RFE, and LASSO to identify predictors, which is one of the advantages of our study.

Eight variables were ultimately identified: dizziness, epilepsy, fall history within the past 3 months, UWA, EA, MFS scores, mBI scores, and the number of Indw Cath. They are also key predictors in other predictive models [11,15,16,44]. Our study identified MFS scores ≥45, nonbedridden and not using a wheelchair, and scores of mBI not 100 as the 3 strongest predictors of falls in hospitalized older adults. These findings align with previous studies [11,15,16]. MFS is widely used for fall risk assessment in hospitals. Previous research has shown that MFS exhibits lower sensitivity than ML models [11,44]. Nevertheless, including it as a predictor in ML models permits the evaluation of its predictive contribution relative to other variables. MFS remains a valid predictor of falls among hospitalized patients [11]. Similarly, in our study, MFS emerged as a strong feature in the ML model.

Moreover, since patients' clinical data often include MFS scores, an integrated model that incorporates MFS can better simulate real-world decision-making, providing a more practical foundation for clinical decision. One of the key strengths of our study lies in integrating a simple, widely used assessment tool with a high-performance ML method, leveraging the advantages of both methodologies to develop and validate a simple, easily generalizable predictive model. This study found that older adults who are not bedridden or not using a wheelchair had a higher fall risk during hospitalization. This may occur as over half of falls happen during daily activities [45], whereas bedridden or wheelchair-bound patients have very low activities of daily living (ADL) ability, limiting activity engagement and thereby reducing fall risk. Similarly, patients with mBI scores >0 had higher fall risk, where higher scores indicate better ADL ability [35]. Notably, those with mBI scores <100 or 21 to 60 also showed increased risk, implying a nonlinear relationship between mBI scores and fall risk. This contradicts findings by Dormosh et al [15] and Chu et al [16] that low ADL ability predicts falls but aligns with Nagarkar et al's [45] longitudinal study linking difficulty with >4 ADL to higher fall odds. Functional decline impairs musculoskeletal integrity and body composition, reducing mobility and increasing fall risk [45-47]. However, the relationship between functional ability and fall risk in elderly patients requires further investigation. Identifying functional states associated with the highest fall risk and implementing dynamic interventions are crucial for preventing falls in this population.

Despite the growing number of ML-based clinical prediction models being developed, most studies lack interpretability of these models, limiting their clinical understanding and practical adoption. The interpretability of ML predictions requires attention from researchers so that physicians can understand, trust, and ultimately apply these predictive models to guide their clinical practice [28,38,39]. SHAP is a model-agnostic interpretation framework grounded in cooperative game theory. Its core lies in computing Shapley values to quantify the marginal contribution of each input feature to individual predictions. This approach provides consistent and locally accurate explanations for every prediction made by the model [38]. In this study, we addressed the "black-box" nature of ML models by implementing SHAP to interpret the GBM model at both global and individualized levels. This means that in a clinical setting, the model can calculate a patient's fall risk in real time and simultaneously provide the primary clinically interpretable factors contributing to that risk, thereby enabling rapid screening and informed decision-making. SHAP improves the clinical utility of prediction models, providing fall risk prediction and interpretable descriptions for older adults during hospitalization, thereby demonstrating its potential to address the "black-box" problem [28,39].

Limitations

This study has several limitations. First, the predictive model was developed using single-center retrospective data, which may introduce potential biases and limit its generalizability to other health care settings. External validation in multicenter cohorts is required to confirm broader applicability. Second, incorporating environmental variables (eg, ward layout, lighting conditions) was challenging due to constraints in single center data collection. Lastly, the exclusion of additional laboratory parameters may have overlooked potential predictors. Future research should prioritize integrating environmental variables, expanding laboratory indicators, and leveraging multicenter datasets for model development and validation.

Conclusions

In this study, multiple ML models were developed and validated using multifaceted clinical data to identify the risk of falls among hospitalized older adults. The GBM model demonstrated the optimal predictive performance. By SHAP, the clinical utility of the predictive model was significantly enhanced. In the future, this GBM fall prediction model could be integrated into the hospital EHR system as an embedded decision support module to dynamically assess fall risk among inpatients and generate real-time alerts. Simultaneously, based on the SHAP values provided by the model, the system could offer evidence to support health care providers in developing personalized intervention measures, thereby translating risk prediction into clinical actions aimed at reducing the incidence of falls in hospitalized older adults.

Funding

This study was supported by Nursing Project of Anhui Institite of Translational Medicine (No. 2024zhyx-hl-A04).

Authors' Contributions

XY, JR, and DS cooperatively designed the study, analyzed the data, and drafted the manuscript. NM conceived and designed the study. MB and MZ contributed to the revision of the manuscript. XC, YL, ZW, XD, ZW, SZ, YZ, JL, XL, and JX contributed to the acquisition of data.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The variable list, patient characteristics, regression analysis results, and model development process.
[DOCX File, 456 KB - aging_v9i1e80602_app1.docx]

References

1. Falls. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/falls [accessed 2025-12-20]
2. Colón-Emeric CS, McDermott CL, Lee DS, Berry SD. Risk assessment and prevention of falls in older community-dwelling adults: a review. JAMA 2024 Apr 23;331(16):1397-1406. [doi: 10.1001/jama.2024.1416] [Medline: 38536167]
3. Salari N, Darvishi N, Ahmadipanah M, Shohaimi S, Mohammadi M. Global prevalence of falls in the older adults: a comprehensive systematic review and meta-analysis. J Orthop Surg Res 2022 Jun 28;17(1):334. [doi: 10.1186/s13018-022-03222-1] [Medline: 35765037]
4. Chen X, He L, Shi K, Wu Y, Lin S, Fang Y. Interpretable machine learning for fall prediction among older adults in China. Am J Prev Med 2023 Oct;65(4):579-586. [doi: 10.1016/j.amepre.2023.04.006] [Medline: 37087076]
5. Jung YS, Suh D, Kim E, Park HD, Suh DC, Jung SY. Medications influencing the risk of fall-related injuries in older adults: case-control and case-crossover design studies. BMC Geriatr 2023 Jul 22;23(1):452. [doi: 10.1186/s12877-023-04138-z] [Medline: 37481554]
6. Dykes PC, Curtin-Bowen M, Lipsitz S, et al. Cost of inpatient falls and cost-benefit analysis of implementation of an evidence-based fall prevention program. JAMA Health Forum 2023 Jan 6;4(1):e225125. [doi: 10.1001/jamahealthforum.2022.5125] [Medline: 36662505]
7. Morris ME, Webster K, Jones C, et al. Interventions to reduce falls in hospitals: a systematic review and meta-analysis. Age Ageing 2022 May 1;51(5):afac077. [doi: 10.1093/ageing/afac077] [Medline: 35524748]
8. Cho I, Kwon JM, Choe W, Cho J, Park SH, Bates DW. Under-reporting of falls in hospitals: a multisite study in South Korea. BMJ Qual Saf 2025 Jul 18;34(8):491-498. [doi: 10.1136/bmjqs-2024-017993] [Medline: 40139777]
9. Moreland B, Kakara R, Henry A. Trends in nonfatal falls and fall-related injuries among adults aged ≥65 years - United States, 2012-2018. MMWR Morb Mortal Wkly Rep 2020 Jul 10;69(27):875-881. [doi: 10.15585/mmwr.mm6927a5] [Medline: 32644982]
10. Shim S, Yu JY, Jekal S, et al. Development and validation of interpretable machine learning models for inpatient fall events and electronic medical record integration. Clin Exp Emerg Med 2022 Dec;9(4):345-353. [doi: 10.15441/ceem.22.354] [Medline: 36128798]
11. Lindberg DS, Prosperi M, Bjarnadottir RI, et al. Identification of important factors in an inpatient fall risk prediction model to improve the quality of care using EHR and electronic administrative data: a machine-learning approach. Int J Med Inform 2020 Nov;143:104272. [doi: 10.1016/j.ijmedinf.2020.104272] [Medline: 32980667]
12. Peel NM, Jones LV, Berg K, Gray LC. Validation of a falls risk screening tool derived from InterRAI acute care assessment. J Patient Saf 2021 Dec 1;17(8):e1152-e1156. [doi: 10.1097/PTS.0000000000000462] [Medline: 29360675]
13. Vratsistas-Curto A, Tiedemann A, Treacy D, Lord SR, Sherrington C. External validation of approaches to prediction of falls during hospital rehabilitation stays and development of a new simpler tool. J Rehabil Med 2018 Feb 13;50(2):216-222. [doi: 10.2340/16501977-2290] [Medline: 29260235]
14. Kawazoe Y, Shimamoto K, Shibata D, Shinohara E, Kawaguchi H, Yamamoto T. Impact of a clinical text-based fall prediction model on preventing extended hospital stays for elderly inpatients: model development and performance evaluation. JMIR Med Inform 2022 Jul 27;10(7):e37913. [doi: 10.2196/37913] [Medline: 35896017]
15. Dormosh N, Damoiseaux-Volman BA, van der Velde N, Medlock S, Romijn JA, Abu-Hanna A. Development and internal validation of a prediction model for falls using electronic health records in a hospital setting. J Am Med Dir Assoc 2023 Jul;24(7):964-970. [doi: 10.1016/j.jamda.2023.03.006] [Medline: 37060922]
16. Chu WM, Kristiani E, Wang YC, et al. A model for predicting fall risks of hospitalized elderly in Taiwan-a machine learning approach based on both electronic health records and comprehensive geriatric assessment. Front Med (Lausanne) 2022;9:937216. [doi: 10.3389/fmed.2022.937216] [Medline: 36016999]
17. Satoh M, Miura T, Shimada T. Development and evaluation of a simple predictive model for falls in acute care setting. J Clin Nurs 2023 Sep;32(17-18):6474-6484. [doi: 10.1111/jocn.16680] [Medline: 36899476]
18. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods 2018 Apr;15(4):233-234. [doi: 10.1038/nmeth.4642] [Medline: 30100822]
19. González-Castro A, Leirós-Rodríguez R, Prada-García C, Benítez-Andrades JA. The applications of artificial intelligence for assessing fall risk: systematic review. J Med Internet Res 2024 Apr 29;26:e54934. [doi: 10.2196/54934] [Medline: 38684088]

20. Liu C, Zhang K, Yang X, et al. Development and validation of an explainable machine learning model for predicting myocardial injury after noncardiac surgery in two centers in China: retrospective study. JMIR Aging 2024 Jul 26;7:e54872. [doi: 10.2196/54872] [Medline: 39087583]
21. Adeli V, Korhani N, Sabo A, et al. Ambient monitoring of gait and machine learning models for dynamic and short-term falls risk assessment in people with dementia. IEEE J Biomed Health Inform 2023 Jul;27(7):3599-3609. [doi: 10.1109/JBHI.2023.3267039] [Medline: 37058371]
22. Kim DW, Seo J, Kwon S, et al. Predicting in-hospital fall risk using machine learning with real-time location system and electronic medical records. J Cachexia Sarcopenia Muscle 2025 Feb;16(1):e13713. [doi: 10.1002/jcsm.13713] [Medline: 39994910]
23. Lim ZK, Connie T, Goh MKO. Fall risk prediction using temporal gait features and machine learning approaches. Front Artif Intell 2024;7:1425713. [doi: 10.3389/frai.2024.1425713] [Medline: 39263525]
24. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology, 3rd edition: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008. URL: https://books.google.com/books/about/Modern_Epidemiology.html?id=Z3vjT9ALxHUC [accessed 2025-12-20]
25. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. I. Principles. Am J Epidemiol 1992 May 1;135(9):1019-1028. [doi: 10.1093/oxfordjournals.aje.a116396] [Medline: 1595688]
26. Stekhoven DJ, Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics 2012 Jan 1;28(1):112-118. [doi: 10.1093/bioinformatics/btr597] [Medline: 22039212]
27. Qi W, Wang Y, Wang Y, et al. Prediction of postpartum depression in women: development and validation of multiple machine learning models. J Transl Med 2025 Mar 7;23(1):291. [doi: 10.1186/s12967-025-06289-6] [Medline: 40055720]
28. Song Y, Yuan Q, Liu H, Gu K, Liu Y. Machine learning algorithms to predict mild cognitive impairment in older adults in China: a cross-sectional study. J Affect Disord 2025 Jan 1;368:117-126. [doi: 10.1016/j.jad.2024.09.059] [Medline: 39271065]
29. Sadat-Hosseini M, Arab MM, Soltani M, Eftekhari M, Soleimani A, Vahdati K. Predictive modeling of Persian walnut (Juglans regia L.) in vitro proliferation media using machine learning approaches: a comparative study of ANN, KNN and GEP models. Plant Methods 2022 Apr 11;18(1):48. [doi: 10.1186/s13007-022-00871-5] [Medline: 35410228]
30. Butler LR, Chen KA, Hsu J, Kapadia MR, Gomez SM, Farrell TM. Predicting readmission after bariatric surgery using machine learning. Surg Obes Relat Dis 2023 Nov;19(11):1236-1244. [doi: 10.1016/j.soard.2023.05.025] [Medline: 37455158]
31. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010 Jan;21(1):128-138. [doi: 10.1097/EDE.0b013e3181c30fb2] [Medline: 20010215]
32. Wang Z, Gu Y, Huang L, et al. Construction of machine learning diagnostic models for cardiovascular pan-disease based on blood routine and biochemical detection data. Cardiovasc Diabetol 2024 Sep 28;23(1):351. [doi: 10.1186/s12933-024-02439-0] [Medline: 39342281]
33. Manolis AA, Manolis TA, Melita H, Mikhailidis DP, Manolis AS. Low serum albumin: a neglected predictor in patients with cardiovascular disease. Eur J Intern Med 2022 Aug;102:24-39. [doi: 10.1016/j.ejim.2022.05.004] [Medline: 35537999]
34. Morse JM, Black C, Oberle K, Donahue P. A prospective study to identify the fall-prone patient. Soc Sci Med 1989;28(1):81-86. [doi: 10.1016/0277-9536(89)90309-2] [Medline: 2928815]
35. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. J Clin Epidemiol 1989;42(8):703-709. [doi: 10.1016/0895-4356(89)90065-6] [Medline: 2760661]
36. Jabłońska B, Mrowiec S. Nutritional status and its detection in patients with inflammatory bowel diseases. Nutrients 2023 Apr 20;15(8):1991. [doi: 10.3390/nu15081991] [Medline: 37111210]
37. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs 2005 Aug;14(7):798-804. [doi: 10.1111/j.1365-2702.2005.01121.x] [Medline: 16000093]
38. Kang CW, Yan ZK, Tian JL, Pu XB, Wu LX. Constructing a fall risk prediction model for hospitalized patients using machine learning. BMC Public Health 2025 Jan 20;25(1):242. [doi: 10.1186/s12889-025-21284-8] [Medline: 39833780]
39. Shen L, Jin Y, Pan AX, et al. Machine learning-based predictive models for perioperative major adverse cardiovascular events in patients with stable coronary artery disease undergoing noncardiac surgery. Comput Methods Programs Biomed 2025 Mar;260:108561. [doi: 10.1016/j.cmpb.2024.108561] [Medline: 39708562]
40. Machado-Duque ME, Camacho-Arteaga L, Sabaté M, Vidal-Guitart X, Machado-Alba JE. Falls in hospitalized older adults and the use of fall risk-increasing drugs and anticholinergic medications in Colombia: a case control study. Front Pharmacol 2024;15:1369200. [doi: 10.3389/fphar.2024.1369200] [Medline: 39021833]
41. Liang XZ, Chai JL, Li GZ, et al. A fall risk prediction model based on the CHARLS database for older individuals in China. BMC Geriatr 2025 Mar 13;25(1):170. [doi: 10.1186/s12877-025-05814-y] [Medline: 40082807]
42. Mao A, Su J, Ren M, Chen S, Zhang H. Risk prediction models for falls in hospitalized older patients: a systematic review and meta-analysis. BMC Geriatr 2025 Jan 14;25(1):29. [doi: 10.1186/s12877-025-05688-0] [Medline: 39810076]
43. Yang Q, Cheng H, Qin J, et al. A machine learning-based Preclinical Osteoporosis Screening Tool (POST): model development and validation study. JMIR Aging 2023 Nov 8;6:e46791. [doi: 10.2196/46791] [Medline: 37986117]

44. Yang A, Shi M, Lau ESH, et al. A novel electronic-health-record based, machine-learning model to predict 1-year risk of fall hospitalisation in older adults: a Hong Kong territory-wide cohort and modelling study. Age Ageing 2025 Aug 29;54(10):afaf285. [doi: 10.1093/ageing/afaf285] [Medline: 41066674]
45. Nagarkar A, Kulkarni S. Association between daily activities and fall in older adults: an analysis of longitudinal ageing study in India (2017-18). BMC Geriatr 2022 Mar 14;22(1):203. [doi: 10.1186/s12877-022-02879-x] [Medline: 35287596]
46. Jehu DA, Davis JC, Falck RS, et al. Risk factors for recurrent falls in older adults: a systematic review with meta-analysis. Maturitas 2021 Feb;144:23-28. [doi: 10.1016/j.maturitas.2020.10.021] [Medline: 33358204]
47. Dipietro L, Campbell WW, Buchner DM, et al. Physical activity, injurious falls, and physical function in aging: an umbrella review. Med Sci Sports Exerc 2019 Jun;51(6):1303-1313. [doi: 10.1249/MSS.0000000000001942] [Medline: 31095087]

Abbreviations

ADLs: activities of daily living
AUPRC: area under the precision-recall curve
AUROC: area under the receiver operating characteristic curve
DCA: decision curve analysis
EA: emergency admission
EHR: electronic health record
FH-3M: fall history in the past 3-months
GBM: gradient boosting machine
Indw Cath: indwelling catheters
KNN: k-nearest neighbor
LASSO: least absolute shrinkage and selection operator
LR: logistic regression
mBI: modified Barthel Index
MFS: Morse Fall Scale
ML: machine learning
NN: neural network
NRS 2002: Nutritional Risk Screening 2002
RF: random forest
RF-RFE: random forest-recursive feature elimination
SHAP: Shapley Additive Explanations
SR: stepwise regression
STRATIFY: St. Thomas's risk assessment tool in falling elderly inpatients
TRIPOD: Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis
UWA: use of walking assistance
XGBoost: extreme gradient boosting

Edited by M Gray; submitted 14.Jul.2025; peer-reviewed by S Sharma, S Shim; revised version received 13.Dec.2025; accepted 16.Dec.2025; published 05.Jan.2026.

Please cite as:
Yang X, Ren J, Su D, Bao M, Zhang M, Chen X, Li Y, Wang Z, Dai X, Wei Z, Zhang S, Zhang Y, Li J, Li X, Xu J, Mo N
Development and Validation of Machine Learning Models for Predicting Falls Among Hospitalized Older Adults: Retrospective Cross-Sectional Study
JMIR Aging 2026;9:e80602
URL: https://aging.jmir.org/2026/1/e80602
doi:10.2196/80602

An Ultra-Brief Informant Questionnaire for Case Finding of Cognitive Impairment Across Diverse Literacy: Diagnostic Accuracy Study

Tau Ming Liew[1,2,3,4], FRCPsych, PhD; King Fan Yip[5], MBBS, MRCP; Kaavya Narasimhalu[6], MD, PhD; Simon Kang Seng Ting[6], MBBS, FRCP; Weishan Li[6], MD, MRCP; Sze Yan Tay[7], MPsych; Way Inn Koay[7], MPsych

[1]Department of Psychiatry, Singapore General Hospital, Outram Road, Singapore, Singapore
[2]SingHealth Duke-NUS Medicine Academic Clinical Programme, Duke-NUS Medical School, Singapore, Singapore
[3]Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore
[4]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore
[5]Department of Geriatric Medicine, Singapore General Hospital, Singapore, Singapore
[6]Department of Neurology, National Neuroscience Institute, Singapore General Hospital, Singapore, Singapore
[7]Department of Psychology, Singapore General Hospital, Singapore, Singapore

Corresponding Author:
Tau Ming Liew, FRCPsych, PhD
Department of Psychiatry, Singapore General Hospital, Outram Road, Singapore, Singapore

Abstract

Background: Undiagnosed cognitive impairment poses a global challenge, prompting recent interest in ultra-brief screening questionnaires (comprising <2 to 3 items) to efficiently identify individuals needing further evaluation. However, evidence on ultra-brief questionnaires remains limited, particularly regarding their validity across diverse literacy levels.

Objective: This study aimed to develop an ultra-brief questionnaire that performs well in detecting mild cognitive impairment or dementia (MCI/dementia) across diverse literacy levels and to compare its performance with an established questionnaire (the 8-item Informant Interview to Differentiate Aging and Dementia [AD8]).

Methods: This diagnostic study involved 1856 participants aged ≥65 years (median education 10 y, range 0 - 23 y), prospectively recruited from community settings in Singapore. Participants and informants completed 15 cognition-related questions. MCI/dementia was diagnosed via a comprehensive assessment and consensus conference. The sample was randomly split 70/30—the training sample (70%) was used to derive an ultra-brief questionnaire from the 15 cognition-related questions (using an exhaustive search approach), and the test sample (30%) evaluated the area under the receiver operating characteristic curve (AUC).

Results: The new questionnaire comprised 2 informant questions (ie, *assistance with medications* and *worry about cognition*), plus age and years of education. It demonstrated excellent performance in detecting MCI/dementia (AUC 85%, 95% CI 80% - 90%), significantly better (*P*=.003) than a nested baseline model (comprising age and years of education; AUC 78%, 95% CI 73% - 83%). In contrast, the AD8 had an AUC of 76% (95% CI 70% - 83%), not significantly different (*P*>.99) from the baseline model. The questionnaire's performance was consistent across education subgroups and varying prevalence scenarios. Two optimal cutoffs were used—the lower cutoff provided 80% sensitivity and 96% negative predictive value, and the upper cutoff provided 99% specificity and 81% positive predictive value. A web-based calculator is available for public use.

Conclusions: This ultra-brief questionnaire enables rapid screening for cognitive impairment (in <1 min) by family members or as part of community geriatric assessments. Its excellent performance across literacy levels supports its utility for case finding in diverse populations, including underserved communities and lower- and middle-income countries.

(JMIR Aging 2026;9:e72963) doi:10.2196/72963

KEYWORDS

machine learning; Informant Questionnaire for cognitive impairment–two items plus demographics; IQ2+; informant questionnaire; subjective questionnaire; subjective cognitive decline; neurocognitive disorders

Introduction

Undiagnosed cognitive impairment is a global challenge [1], with 60% to 90% of affected individuals never receiving a formal diagnosis [2,3]. Those who remain undiagnosed miss out on timely clinical care [4], including management of reversible causes, prescription of cognitive enhancers, behavioral management, caregiver support, and advanced care planning [4-9]. All these can impact well-being [10,11] and increase the risk of premature nursing home placement [12-14]. Furthermore, undiagnosed individuals often do not receive adequate support to manage and coordinate care for their chronic diseases [15,16], resulting in suboptimal disease management, inappropriate health care utilization, and higher health care costs [17,18]. As an example, a modeling study estimated that timely management of cognitive impairment could potentially yield annual cost savings of US $13 to $41 billion in the United States [15]. Recently, the importance of early diagnosis has been heightened by emerging evidence supporting early interventions for cognitive impairment [19,20], such as risk factor modification [21,22] and antiamyloid monoclonal antibodies [23,24].

To address the challenge of undiagnosed cognitive impairment, various international bodies (eg, the Alzheimer's Disease International [25] and the International Association of Gerontology and Geriatrics) [26] have advocated for a systematic approach to case finding among high-risk individuals in the community [26]. In particular, a 2-stage strategy [27-29] has been proposed in recent literature to address resource constraints for community case finding. In the first stage, subjective reports (eg, Functional Activities Questionnaire and AD8 [the 8-item Informant Interview to Differentiate Aging and Dementia]) [30,31] are used to identify individuals with potential cognitive impairment. In the second stage, these individuals undergo brief cognitive tests (eg, Mini-Cog and short variants of Montreal Cognitive Assessment) [4,32-34] to confirm the presence of cognitive impairment. This 2-stage strategy is efficient and scalable—the first stage relies solely on subjective reports and does not draw on scarce health care resources for administration of cognitive tests, whereas the second stage reserves brief cognitive tests for a smaller subset of individuals [27]. Moreover, combining subjective reports and brief cognitive tests has been shown to improve the detection of subtle cognitive changes [27], making this approach optimal for case finding of early cognitive impairment.

In 2019, the 2-stage strategy was adopted by the World Health Organization (WHO) within the Integrated Care for Older People assessment tool to identify cognitive impairment, alongside assessments of other key components of intrinsic capacity (ie, mobility, nutrition, hearing, vision, and mood) [29]. For the first stage (subjective report), the WHO adopted an ultra-brief questionnaire based on a single question: "Do you have problems with memory or orientation (such as not knowing where one is or what day it is)?" [29] The decision to embed an ultra-brief cognitive questionnaire within Integrated Care for Older People is understandable, as it balances the need to assess a wide range of intrinsic capacity domains against the scarcity of community resources for comprehensive assessments. However, the validity of such ultra-brief questionnaires (ie, those comprising fewer than 2 - 3 items) remains unclear in the literature [28], especially when used across diverse levels of literacy [26,35]. This concern can be critical, as questionnaires with fewer items tend to have increased measurement variability and may be more susceptible to confounding factors such as educational attainment [36].

In this study, we sought to strengthen the evidence base supporting the use of ultra-brief questionnaires across diverse levels of literacy, potentially enhancing their utility in diverse populations across lower-, middle-, and higher-income countries. Specifically, we aimed to: (1) derive an ultra-brief questionnaire with high performance for detecting cognitive impairment (ie, the presence of mild cognitive impairment or dementia [MCI/dementia]) across diverse levels of literacy, using a contemporary, computationally intensive approach to identify the questions most discriminative of MCI/dementia; and (2) compare the performance of the new ultra-brief questionnaire to the well-established AD8 across participants with lower and higher educational attainment. We selected the AD8 as the benchmark because it is a widely used and extensively validated informant-based questionnaire, as demonstrated in recent systematic reviews [37,38]. Its use has been recommended by various international bodies, including the International Association of Gerontology and Geriatrics [26], the Gerontological Society of America [39], the US Alzheimer's Association [40,41], and the National Institute on Aging workgroup [42].

Of note, this study was conducted in Singapore, a city-state in Southeast Asia that provides a unique testbed of literacy diversity for developing the ultra-brief questionnaire. The current generation of older Singaporeans witnessed the country's transformation from a traditional, lower-income, Asian society to a more westernized, higher-income country [43]. Consequently, this cohort encompasses a wide range of educational backgrounds, from minimal formal education to tertiary education. By validating the ultra-brief questionnaire in such a heterogeneous population, we sought to demonstrate its potential for broader implementation in other literacy-diverse settings beyond Singapore, including populations across East and South Asia, and potentially, in some lower- and middle-income countries (LMICs).

Methods

Study Procedures

This study involved community-dwelling older persons recruited between March 2022 and September 2024, as part of a nationally funded project in Singapore aimed at developing artificial intelligence tools to detect early cognitive impairment in the community (Project PENSIEVE) [44]. Community-dwelling individuals were invited to participate if they met the following criteria: (1) higher risk of cognitive impairment (ie, aged ≥65 y [26]) and had at least one of the 3 chronic diseases (ie, diabetes mellitus, hypertension, or hyperlipidemia); this criterion was included a priori to focus on individuals with at least some risk of cognitive impairment, in line with current literature suggesting the limited benefit of screening among low-risk individuals) [26,45]; (2) ability to follow simple instructions in

English or Mandarin (due to limitations in the available assessment language); and (3) presence of an informant (eg, family member or friend) who knew the participants well. Individuals were excluded if they had severe visual impairment that would affect their ability to complete neuropsychological assessments (to ensure generalizability, participants were included as long as they could see pictures on a piece of paper held in front of them). No participants were excluded for reasons related to missing data, as we implemented strict data collection procedures (eg, mandatory data field) and routine data audits throughout the study.

Sources of recruitment included 14 community roadshows conducted by the study team, clients of our community partners, home visits by community volunteers who partnered with us, media publicity (radio, online articles, and posters), and word-of-mouth referrals from participants who had completed research assessments. To ensure the recruited samples were representative of the community, the study's publicity materials emphasized the key message of "detect dementia early" (along with direct referrals to memory clinics in the event of significant findings), rather than the conventional invitation to participate in research (which may inadvertently attract a distinct group of individuals). Examples of these publicity materials (eg, study banner, poster, and brochure) are provided in Method S1 in Multimedia Appendix 1.

All participants received comprehensive assessments, which included semistructured interviews with both participants and their informants, detailed neuropsychological testing, and observational notes of participants' behavior during assessments. Full descriptions of the comprehensive assessments are available in Method S2 in Multimedia Appendix 1, with further details on each assessment tool provided in Method S3 in Multimedia Appendix 1. Diagnoses of MCI and dementia were determined via consensus conference by 3 dementia specialists. Dementia was diagnosed using the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, criteria [46], which require the presence of cognitive concerns (reported by the individual or a knowledgeable informant), impairment in instrumental activities of daily living (eg, managing money or medications), and objective cognitive deficits. MCI was diagnosed using the modified Petersen criteria [47], which require the presence of cognitive concerns (reported by the individual or a knowledgeable informant), absence of impairment in instrumental activities of daily living, and the presence of objective cognitive deficits. Normal cognition was diagnosed when participants did not meet criteria for dementia or MCI.

Measures

AD8 [30] is an 8-item, informant-based questionnaire that assesses changes in a participant's cognition and function over the past few years. For each item, the informant rates whether there has been a change in the participant: 1=yes (a change) and 0=no (no change) or don't know. Responses to the 8 items are summed to provide a total score, with higher scores reflecting greater cognitive problems. AD8 has been shown to be useful in detecting varying severities of cognitive impairment [37,38]. Informant AD8 has also been previously validated in Singapore [48].

To derive an ultra-brief questionnaire, the study team focused on candidate questions that assess the two key criteria for diagnosing MCI/dementia [46,47]: (1) the presence of cognitive concerns as reported by the individual or a knowledgeable informant; and (2) impairment in instrumental activities of daily living (iADL). The presence of cognitive concerns was evaluated using validated questions related to subjective cognitive decline (SCD), by asking participants or informants: "Do you feel like your (or your family member's) memory is becoming worse?" and "Are you *worried* that your (or your family member's) memory is becoming worse?" These 2 sets of questions have been validated in previous studies [49,50] and have been shown to be useful for capturing early symptoms of cognitive decline [51-55].

Impairment in iADL was assessed using the locally validated modified Lawton scale [56], with informants asked about difficulties in various domains (public commuting, grocery shopping, managing money, using the telephone, taking medications, preparing meals, doing housework, and doing laundry). The original responses on the modified Lawton scale included 4 options (ie, unable to do at all, needs some help, needs no help, and never needed to do this), which were collapsed in this study into two options: (1) yes and (2) no/never needed to do this.

In total, 15 candidate questions were considered for the ultra-brief questionnaire: 12 cognition-related items and 3 basic demographic variables (eg, age, sex, and years of education), given their potential correlation with cognition. The exact wording and response options for all 15 items are provided in Table 1.

Table . A preselected list of 15 question items that are potentially related to cognition.

Question category [a]	Question item [a]	Response options
SCD[b] question—informant	Do you feel like your family member's memory is becoming worse?	1=yes; 0=no/not sure
Worry about cognition—informant	Are you *worried* that your family member's memory is becoming worse?	1=yes; 0=no/not sure
SCD question—participant	Do you feel like your memory is becoming worse?	1=yes; 0=no/not sure
Worry about cognition—participant	Are you *worried* that your memory is becoming worse?	1=yes; 0=no/not sure
iADL[c]—commute	Does your family member need help to take public transport or drive a car?	1=yes; 0=no/never needed to do this
iADL—grocery	Does your family member need help to do grocery shopping?	1=yes; 0=no/never needed to do this
iADL—money	Does your family member need help to manage money?	1=yes; 0=no/never needed to do this
iADL—telephone	Does your family member need help to use the telephone?	1=yes; 0=no/never needed to do this
iADL—medications	Does your family member need help to take medications?	1=yes; 0=no/never needed to do this
iADL—meals	Does your family member need help to prepare meals?	1=yes; 0=no/never needed to do this
iADL—housework	Does your family member need help to do housework?	1=yes; 0=no/never needed to do this
iADL—laundry	Does your family member need help to do laundry?	1=yes; 0=no/never needed to do this
Age	What is your family member's age?	Continuous variable
Sex	What is your family member's sex?	1=male; 0=female
Years of education	What is your family member's years of education?	Continuous variable. Count the years of full-time education, starting from elementary/primary school.

[a]Items in the list were selected to assess the 2 key criteria in the diagnosis of mild cognitive impairment and dementia: (1) the presence of cognitive concerns as reported by the individual or a knowledgeable informant and (2) impairment in iADL [46,47]. The presence of cognitive concerns was evaluated using validated questions related to SCD, given prior literature on the usefulness of SCD to reflect early symptoms of cognitive decline [51-55]. Impairment in iADL was assessed using the locally validated modified Lawton scale [56], with informants asked about difficulties in various domains of iADL. Three basic demographic variables (ie, age, sex, and years of education) were also included, given their potential correlation with cognition.

[b]SCD: subjective cognitive decline.

[c]iADL: instrumental activities of daily living.

Statistical Analyses

The study sample was randomly split into a 70% *training sample* and a 30% *test sample*. The *training sample* was used to develop an ultra-brief questionnaire that best distinguished MCI/dementia from normal cognition, whereas the *test sample* was used to evaluate the actual performance of this questionnaire.

In the *training sample*, a best-subset approach [57] with 5-fold cross-validation was used to select the optimal combination of items from the 15 candidate questions. The best-subset approach is an efficient, computationally intensive method for variable selection [4,20,58,59] in which logistic regression is used to exhaustively evaluate all possible combinations of the candidate questions, identifying models with the lowest prediction errors. Five-fold cross-validation was then used to select the most parsimonious model within 1 SE of the best-performing model. After identifying the optimal model from the best-subset approach, we further refined the model by considering potential inclusions of quadratic terms for continuous variables (eg, age and years of education); quadratic terms with P<.05 were incorporated into the final model to improve fit. The final selected model was then applied to the test sample to generate predicted probabilities of cognitive impairment (ie, MCI/dementia), with the model variables constituting the new ultra-brief questionnaire.

Predicted probabilities were computed from logistic regression using the following equation:

Predicted probability=eLogit1+eLogit

where

$$\text{logit}=\beta 0+\beta 1\cdot(\text{Variable 1})+\beta 2\cdot(\text{Variable 2})+\beta 3\cdot(\text{Variable 3})+\ldots$$

with each β representing the regression coefficient for its respective variable in the model.

In the *test sample*, the predicted probabilities were used to compute the area under the receiver operating characteristic curve (AUC), thereby assessing the actual performance of the ultra-brief questionnaire in discriminating MCI/dementia from normal cognition. In general, an AUC of 0.7 to 0.8 is considered acceptable discrimination, 0.8 to 0.9 is considered excellent, and more than 0.9 is considered outstanding [60]. Comparisons of AUC were conducted using the nonparametric method proposed by DeLong et al [33,59,61,62], with analyses stratified by education subgroups (ie, ≤10 and >10 y of education based on median split). A 2-cutoff approach [63-67] was adopted for the ultra-brief questionnaire. The first cutoff was chosen to yield high sensitivity and negative predictive value (>80% each) and was used to rule out MCI/dementia (ie, when probability scores fell below the first threshold). The second cutoff was selected for high specificity and positive predictive value (>80% each), identifying those very likely to have MCI/dementia. This 2-cutoff approach is recommended in recent literature [67], as it enhances the performance of cognitive assessment tools [63-66], reduces the effects of prevalence on tool performance [64], and allows prioritization of scarce health care resources for individuals who truly require further cognitive assessments [63].

As a secondary analysis, the performance of the ultra-brief questionnaire was also evaluated for distinguishing dementia from nondementia. Additionally, 2 sensitivity analyses were conducted in the test sample to evaluate the robustness of results when the prevalence of MCI/dementia was readjusted to reflect the average prevalence in most communities:

- Prevalence of MCI/dementia was artificially readjusted to 20% based on prior meta-analytic findings that community prevalence is ~15% for MCI [68-70] and ~5% for dementia [71-73]. Readjustment of prevalence was done by randomly selecting a subset of participants with MCI and normal cognition—for each participant with dementia, 3 participants with MCI and 16 participants with normal cognition were randomly selected (ie, so that the final dataset corresponded to 5% prevalence for dementia and 15% prevalence for MCI).
- Prevalence of MCI/dementia was artificially readjusted to 35% based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI [69,70] and ~10% for dementia [72,73]. Readjustment of prevalence was done by randomly selecting only a subset of participants with MCI and normal cognition—for each participant with dementia, 2.5 participants with MCI and 6.5 participants with normal cognition were randomly selected (ie, so that the final dataset corresponded to 10% prevalence for dementia and 25% prevalence for MCI).

The best-subset approach was performed with the “bestglm” [57] package in R (version 4.4.0; R Foundation for Statistical Computing). All remaining analyses were conducted in Stata (version 18; StataCorp LLC). No a priori sample size calculation was performed; the final sample size was determined pragmatically based on participants recruited between March 2022 and September 2024. Post hoc power analyses confirmed that the test sample provided robust power (90%) to distinguish MCI/dementia from normal cognition (α=.05, two-sided test). Power was also reasonably sufficient for participants with ≤10 years of education (72%) but was limited in those with >10 years of education (19%) due to fewer positive cases. The power calculations were conducted using PASS software (version 15.0.5; NCSS, LLC) and the Hanley and McNeil formula [74], with further details provided in MethodS4 in Multimedia Appendix 1.

Ethical Considerations

The study received ethical approval from the SingHealth Centralized IRB of Singapore (reference number: 2021/2590). Informed consent was obtained from all participants. Before obtaining informed consent, the mental capacity of participants was briefly assessed in accordance with the Mental Capacity Act of Singapore [75]. In the event of concerns about mental capacity, informed consent by proxy was then obtained from the legally authorized next-of-kin. Participants who completed the research assessments received Singapore Dollar $80 as compensation for their time, inconvenience and transportation costs.

Results

Overview

A total of 1856 participants were included, of whom 255 (13.7%) had MCI/dementia. Participant characteristics are presented in Table 2, with a median age of 72 years and a median education of 10 years (range 0-23 y). Corresponding informant characteristics are provided in Table S1 in Multimedia Appendix 1, with informants primarily comprising spouses (897/1856, 48.3%) and children (506/1856, 27.3%). The sample was randomly split into a *training sample* (1299/1856, 70%) and a *test sample* (557/1856, 30%). Table S2 in Multimedia Appendix 1 shows that the training and test samples had comparable demographic characteristics (P>.05 across covariates).

Table . Characteristics of the study participants.

Variable	Overall sample (n=1856)	Normal cognition (n=1601)	MCI[a] (n=207)	Dementia (n=48)	*P* value [b]
Age, median (IQR) [range]	72 (68 to 76) [65 to 101]	71 (68 to 75) [65 to 93]	74 (70 to 79) [65 to 91]	80 (76 to 82) [66 to 101]	<.001
Years of education, median (IQR) [range]	10 (9 to 13) [0 to 23]	10 (10 to 13) [0 to 23]	10 (6 to 12) [0 to 21.5]	10 (2 to 10) [0 to 17]	<.001
Sex (male), n (%)	688 (37.1)	572 (35.7)	101 (48.8)	15 (31.2)	<.001
Ethnicity, n (%)					.39
Chinese	1735 (93.5)	1504 (93.9)	188 (90.8)	43 (89.6)	
Malay/Indian	96 (5.2)	77 (4.8)	15 (7.2)	4 (8.3)	
Eurasian/others	25 (1.3)	20 (1.2)	4 (1.9)	1 (2.1)	
MoCA[c] total score, median (IQR)	26 (24 to 28)	27 (25 to 28)	21 (17 to 24)	14 (9 to 19)	<.001
NTB[d] Global Z-scores, median (IQR)	−0.2 (−0.6 to 0.1)	−0.1 (−0.4 to 0.2)	−1.0 (−1.3 to 0.7)	−1.6 (−2.1 to 1.2)	<.001
Global CDR[e], n (%)					<.001
0	1570 (84.6)	1557 (97.3)	13 (6.3)	0 (0.0)	
0.5	255 (13.7)	44 (2.7)	194 (93.7)	17 (35.4)	
1	22 (1.2)	0 (0.0)	0 (0.0)	22 (45.8)	
2	8 (0.4)	0 (0.0)	0 (0.0)	8 (16.7)	
3	1 (0.1)	0 (0.0)	0 (0.0)	1 (2.1)	

[a]MCI: mild cognitive impairment.

[b]Test of difference across diagnoses: chi-square test for categorical variables and Kruskal-Wallis test for continuous variables.

[c]MoCA: Montreal Cognitive Assessment.

[d]NTB: neuropsychological battery.

[e]CDR: clinical dementia rating.

Development of the Ultra-Brief Questionnaire

Table 3 presents the top models identified through the exhaustive search method in the *training sample* (n=1299). *iADL–medications* and *worry about cognition–informant* were among the most useful items in detecting MCI/dementia, whereas *iADL–meals* and *iADL–grocery* were among the least useful. Following 5-fold cross-validation, the model with 4 items was identified as the most parsimonious among the top models (Figure S1 in Multimedia Appendix 1). This model included 2 informant questions (*iADL–medications* and *worry about cognition–informant*), plus 2 demographic variables (*age* and *years of education*). These items were then selected to constitute the new ultra-brief questionnaire, henceforth denoted as IQ2+ (Informant Questionnaire for cognitive impairment–2 items plus demographics). Further model refinement considered the inclusion of quadratic terms for *age* and *years of education*, resulting in the addition of a quadratic term for years of education (*P*=.004) in the final model. Thus, the final IQ2+ model comprised five predictors: iADL–medications, worry about cognition–informant, age, years of education, and the quadratic term for years of education. Responses from the IQ2+ questionnaire can be converted to predicted probabilities using an interactive web-based calculator we have created [76]. A screenshot of the calculator is shown in Figure 1.

Table . The top models that best discriminate mild cognitive impairment and dementia from normal cognition (as identified by the best-subset approach) in the training sample (n=1299).

Question items, rearranged by their usefulness in detecting MCI[a]/dementia[b]	Number of items in the top models														
	1	2	3	4[c]	5	6	7	8	9	10	11	12	13	14	15
iADL[d]—medications	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Worry about cognition—informant	N/A[e]	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Age	N/A	N/A	●	●	●	●	●	●	●	●	●	●	●	●	●
Years of education	N/A	N/A	N/A	●	●	●	●	●	●	●	●	●	●	●	●
Sex	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●	●	●	●	●
iADL—money	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●	●	●	●
SCD[f] question—informant	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●	●	●
iADL—commute	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●	●
iADL—telephone	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●	●
Worry about cognition—participant	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●	●
iADL—laundry	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●	●
SCD question—participant	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●	●
iADL—housework	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●	●
iADL—grocery	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●	●
iADL—meals	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	●

[a]MCI/dementia: mild cognitive impairment or dementia.

[b]A full description of each item is available in Table 1. Briefly, items in the list were selected to assess the 2 key criteria in the diagnosis of mild cognitive impairment and dementia: (1) the presence of cognitive concerns as reported by the individual or a knowledgeable informant and (2) impairment in iADL [46,47]. The presence of cognitive concerns was evaluated using validated questions related to SCD, given prior literature on the usefulness of SCD to reflect early symptoms of cognitive decline [51-55]. Impairment in iADL was assessed using the locally validated modified Lawton scale [56], with informants asked about difficulties in various domains of iADL. Three basic demographic variables (ie, age, sex, and years of education) were also included, given their potential correlation with cognition.

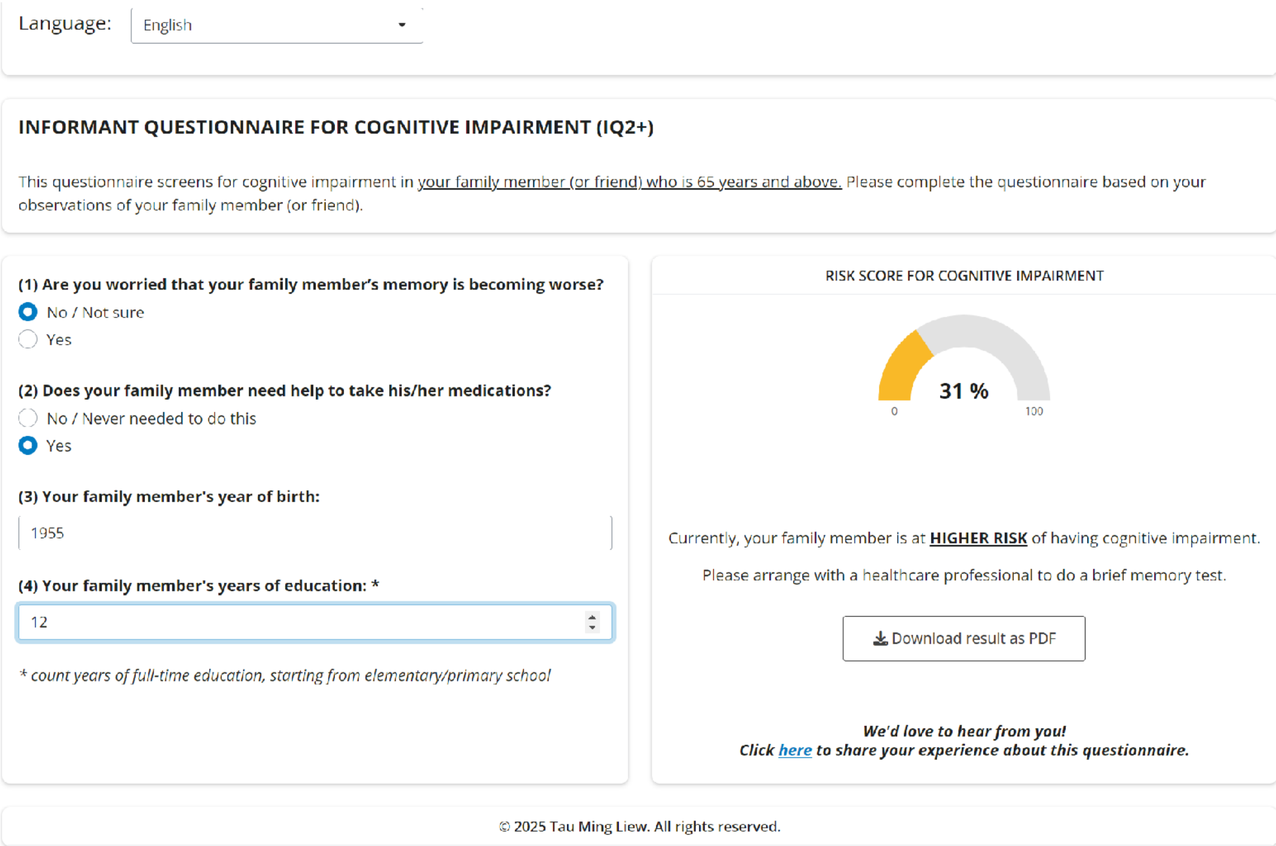
[c]The 4-item model was identified as the most parsimonious among the top models. Further details on model selection are available in Figure S1 in Multimedia Appendix 1.

[d]iADL: instrumental activities of daily living.

[e]N/A: not applicable.

[f]SCD: subjective cognitive decline

Figure 1. A sample screenshot of the interactive web-based calculator. The web-based calculator can be accessed at [76].



Performance of the Ultra-brief Questionnaire

Table 4 presents the AUC results for IQ2+ and AD8 in the *test sample* (n=557). Predicted probabilities from the new IQ2+ demonstrated excellent performance in distinguishing MCI/dementia from normal cognition (AUC 85.3%, 95% CI 80.4% - 90.2%), which was significantly better (*P*=.003) than a nested baseline model (comprising age, years of education, and the quadratic term of years of education; AUC 78.0%, 95% CI 72.6% - 83.4%). In contrast, AD8 had an AUC of 76.1% (95% CI 69.6% - 82.6%), which was not significantly different (*P*>.99) from that of the baseline model. For the detection of dementia, both IQ2+ and AD8 had AUCs >90%, which were significantly higher (*P*<.05) than the baseline model (83.2%). AUC results remained largely similar across education subgroups and in the 2 sensitivity analyses where the prevalence of MCI/dementia was increased to reflect average prevalence in most communities (ie, 20% [68-73] and 35% [69,70,72,73], respectively).

Table . Performance of IQ2+ for detecting cognitive impairment in the test sample (n=557) and a comparison with the performance of AD8.

Assessment tool	All education subgroups		≤10 y of education[a]		>10 y of education[a]	
	AUC[b], % (95% CI)	*P* value[c]	AUC, % (95% CI)	*P* value[c]	AUC, % (95% CI)	*P* value[c]
Detection of MCI[d]/dementia						
Baseline model (age and education)[e]	78.0 (72.6 - 83.4)	Ref[f]	76.7 (69.9 - 83.5)	Ref	74.8 (65.2 - 84.5)	Ref
IQ2+[g]	85.3 (80.4 - 90.2)	.003	84.5 (78.6 - 90.4)	.015	83.3 (72.5 - 94.0)	.095
AD8[h]	76.1 (69.6 - 82.6)	1.000	75.5 (68.0 - 83.0)	1.000	76.8 (62.7 - 91.0)	1.000
Detection of dementia						
Baseline model (age and education)[e]	83.2 (72.8 - 93.6)	Ref	77.8 (64.5 - 91.2)	Ref	94.1 (90.2 - 97.9)	Ref
IQ2+	96.7 (92.8 - 100)	.035	95.1 (89.6 - 100)	.034	99.8 (99.2 - 100)	.003
AD8	99.4 (98.8 - 99.9)	.005	99.6 (99.1 - 100)	.003	99.1 (97.8 - 100)	.059
Sensitivity analysis 1 (prevalence of MCI/dementia=20%)[i]						
Detection of MCI/dementia						
Baseline model (age and education)[e]	74.8 (68.3 - 81.3)	Ref	73.8 (65.7 - 81.9)	Ref	70.1 (58.4 - 81.7)	Ref
IQ2+	85.2 (79.6 - 90.8)	<.001	84.3 (77.4 - 91.2)	.005	83.6 (72.1 - 95.0)	.023
AD8	77.1 (70.1 - 84.2)	1.000	75.9 (67.6 - 84.2)	1.000	79.1 (64.7 - 93.5)	.525
Detection of dementia						
Baseline model (age and education)[e]	81.5 (70.7 - 92.3)	Ref	76.4 (62.9 - 90.0)	Ref	92.5 (87.7 - 97.4)	Ref
IQ2+	96.5 (92.1 - 100)	.023	94.8 (88.5 - 100)	.026	100 (100 - 100)	.005
AD8	99.1 (98.3 - 100)	.003	99.5 (98.9 - 100)	.002	98.6 (96.5 - 100)	.069
Sensitivity analysis 2 (prevalence of MCI/dementia=35%)[j]						
Detection of MCI/dementia						
Baseline model (age and education)[e]	77.0 (69.3 - 84.8)	Ref	75.6 (65.8 - 85.5)	Ref	73.2 (58.0 - 88.3)	Ref
IQ2+	85.1 (78.5 - 91.6)	.031	83.8 (75.7 - 91.9)	.135	82.4 (67.6 - 97.1)	.280
AD8	73.4 (65.0 - 81.9)	1.000	73.5 (63.5 - 83.4)	1.000	73.6 (56.3 - 91.0)	1.000
Detection of dementia						
Baseline model (age and education)[e]	80.1 (68.5 - 91.7)	Ref	73.6 (58.7 - 88.5)	Ref	95.8 (90.6 - 100)	Ref
IQ2+	94.4 (89.3 - 99.5)	.044	91.8 (84.4 - 99.2)	.046	99.2 (96.8 - 100)	.307
AD8	98.8 (97.5 - 100)	.004	99.2 (98.1 - 100)	.002	98.3 (95.2 - 100)	.893

[a]Education subgroups were stratified based on median split. This subgroup analysis has reduced statistical power (see Method S4 in Multimedia Appendix 1 for details) and is exploratory in nature.

[b]AUC: area under the receiver operating characteristic curve.

[c]*P* values were based on comparisons of AUC using the nonparametric method proposed by DeLong et al [61]. *P*<.05 indicates significant difference in AUC between the baseline model and the respective assessment tools. *P* values were Bonferroni-adjusted to minimize the risk of type 1 error in the

context of multiple testing.

[d]MCI/dementia: mild cognitive impairment or dementia.

[e]This baseline model was provided mainly for comparison purposes, by omitting IQ2+'s 2 core questions (ie, *assistance with medications* and *worry about cognition*) to examine the incremental utility of the 2 core questions beyond those provided by the demographic information of age and education. The baseline model was generated in the training sample using a logistic model with the dependent variable of MCI/dementia and with the independent variables of age, years of education, and the quadratic term of years of education. This baseline model was then applied to the test sample to generate predicted probabilities of MCI/dementia, with the predicted probabilities used for AUC comparisons.

[f]Ref: reference.

[g]IQ2+: the Informant Questionnaire for cognitive impairment–2 items (plus demographics).

[h]AD8: the 8-item Informant Interview to Differentiate Aging and Dementia.

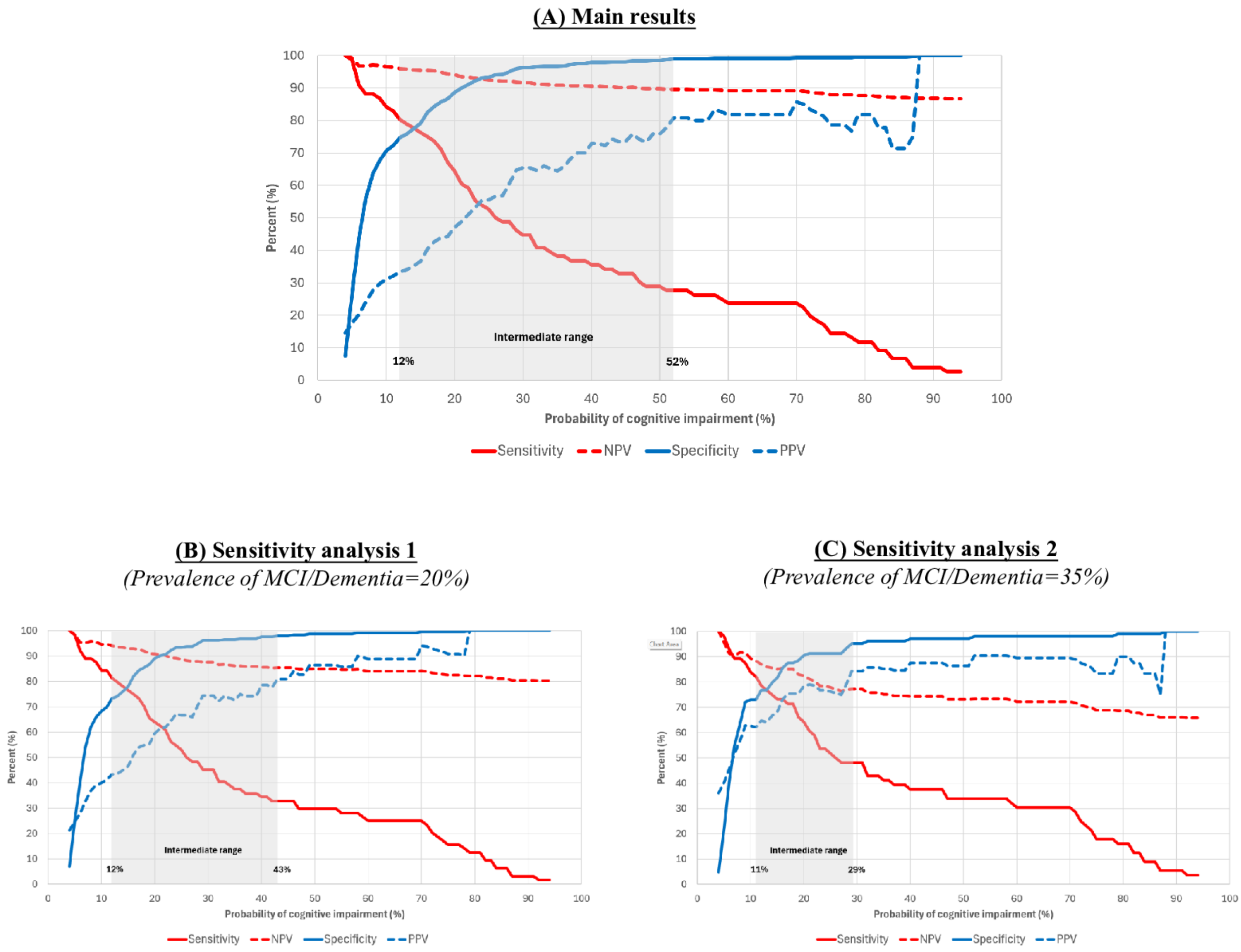
[i]Prevalence of MCI/dementia was readjusted to 20% in the test sample based on prior meta-analytic findings that community prevalence was ~15% for MCI and ~5% for dementia. In the test sample, a subset of participants with MCI and dementia was randomly selected to readjust the prevalence in the dataset (see Methods section for further details). The resulting dataset comprised 256 participants with normal cognition (80%), 48 participants with MCI (15%), and 16 participants with dementia (5%).

[j]Prevalence of MCI/dementia was readjusted to 35% in the test sample based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI and ~10% for dementia. In the test sample, a subset of participants with MCI and dementia was randomly selected to readjust the prevalence in the dataset (see Methods section for further details). The resulting dataset comprised 104 participants with normal cognition (65%), 40 participants with MCI (25%), and 16 participants with dementia (10%).

Optimal Cutoffs of the Ultra-Brief Questionnaire

Test statistics for IQ2+ are plotted in Figure 2A. Adopting a 2-cutoff approach, the lower cutoff (probability ≥12%) had 80.3% sensitivity and 96.0% negative predictive value and was used to rule out MCI/dementia (for individuals with probability scores below the cutoff), whereas the upper cutoff (probability ≥52%) had 99.0% specificity and 80.8% positive predictive value and identified those likely to have MCI/dementia (ie, to rule in MCI/dementia). These 2 cutoffs provide an intermediate range between them (grayed area in Figure 2A), identifying those who warrant further assessment. The optimal cutoffs varied slightly with changing prevalence of MCI/dementia, as shown in Figure 2B and 2C. Detailed results on test statistics for IQ2+ are also available in Tables S3–S5 in Multimedia Appendix 1.

Figure 2. Plot of sensitivity, specificity, NPV, and PPV based on probabilities of cognitive impairment generated from IQ2+ in the test sample (n=557). This plot is intended to demonstrate the 2-cutoff approach. The lower cutoff identifies sensitivity and NPV (red lines), which are >80% each and are used to rule out MCI/dementia (when probability scores fall below this threshold). The upper cutoff identifies specificity and PPV (blue lines), which are >80% each, and are used to rule in MCI/dementia (when probability scores exceed this threshold). The grayed area (demarcated by the lower and upper cutoffs) represents the intermediate range, identifying those who may warrant further assessment. Plot (A) was based on the main results from all the test samples (n=577). Plot (B) was based on results from sensitivity analysis 1, whereby the prevalence of MCI/dementia was readjusted to 20% in the test sample based on prior meta-analytic findings that community prevalence was ~15% for MCI and ~5% for dementia (in the test sample, a subset of participants with MCI and dementia were randomly selected in the test sample to readjust the prevalence in the dataset; see Methods section for further details). Plot (C) was based on results from sensitivity analysis 2, whereby the prevalence of MCI/dementia was readjusted to 35% in the test sample based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI and ~10% for dementia (in the test sample, a subset of participants with MCI and dementia were randomly selected in the test sample to readjust the prevalence in the dataset; see Methods section for further details). IQ2+: the Informant Questionnaire for Cognitive Impairment–2 Items (plus demographics); MCI/dementia: mild cognitive impairment or dementia; NPV: negative predictive value; PPV: positive predictive value.



Result Interpretation for the Ultra-Brief Questionnaire

The 2 cutoffs of IQ2+ effectively identified three risk categories for cognitive impairment: (1) less likely to have cognitive impairment, (2) higher risk of cognitive impairment, and (3) likely to have cognitive impairment. These 3 categories, along with their cross-tabulation with the final diagnoses, are presented in Table 5. In the first category (ie, less likely to have cognitive impairment), 88% to 96% of individuals had normal cognition. In the second category (ie, higher risk of cognitive impairment), 25% to 45% of individuals were diagnosed with MCI. In the third category (ie, likely to have cognitive impairment), 81% to 84% of the individuals had MCI/dementia, with a large proportion having dementia (47%-54%). Distinctions between these 3 risk categories of IQ2+ are also visible in the box plots in Figure 3A–3C. The first category (white region with probability scores below the lower cutoff) identified those with normal cognition, the third category (dark gray region with probability scores above the upper cutoff) identified almost all individuals with dementia, and the second category (light gray region between the lower and upper cutoffs) mostly captured those with MCI. In contrast, AD8 showed poorer discrimination between normal cognition and MCI, with discernible floor effects in the normal cognition and MCI groups as seen in the box plots in Figure 3D–3F. This is also reflected in the test statistics of AD8 (as presented in Tables S6–S8 in Multimedia Appendix 1), which demonstrated low sensitivity across its cutoff scores (with maximum sensitivity of 69.6% - 73.4%).

Table . Cross-tabulation between the output from IQ2+ and the final diagnosis in the test sample (n=557).

Result from IQ2+[ab]	Final diagnosis		
	Normal cognition	MCI[c]	Dementia
Less likely to have CI, n (%)[d]	360 (96.0)	14 (3.7)	1 (0.3)
Higher risk of CI, n (%)[d]	116 (74.4)	39 (25.0)	1 (0.6)
Likely to have CI, n (%) [d]	5 (19.2)	7 (26.9)	14 (53.8)
Sensitivity analysis 1 (prevalence of MCI/dementia=20%)[e]			
Less likely to have CI, n (%)[f]	188 (94.0)	11 (5.5)	1 (0.5)
Higher risk of CI, n (%)[f]	63 (67.0)	30 (31.9)	1 (1.1)
Likely to have CI, n (%)[f]	5 (19.2)	7 (26.9)	14 (53.8)
Sensitivity analysis 2 (Prevalence of MCI/dementia=35%)[g]			
Less likely to have CI, n (%)[h]	76 (88.0)	10 (12.0)	0 (0.0)
Higher risk of CI, n (%)[h]	23 (55.0)	18 (43.0)	1 (2.0)
Likely to have CI, n (%)[h]	5 (16.0)	12 (38.0)	15 (47.0)

[a]IQ2+, the Informant Questionnaire for cognitive impairment–2 items (plus demographics).

[b]Two-cutoff approach was adopted for IQ2+. The lower cutoff has high sensitivity and negative predictive value (>85% respectively) and is used to rule out MCI/dementia (for individuals with probability scores below the cutoff). The upper cutoff has high specificity and positive predictive value (>85% respectively) and identifies those who are likely to have MCI/dementia. These 2 cutoffs provide an intermediate range between them, identifying those who may be at higher risk and require further monitoring or assessment.

[c]MCI: mild cognitive impairment.

[d]Probability cutoff for the main results (ie, prevalence of MCI/dementia=14%):<12% (less likely to have CI), 12% to 51% (higher risk of CI), and ≥52% (likely to have CI).

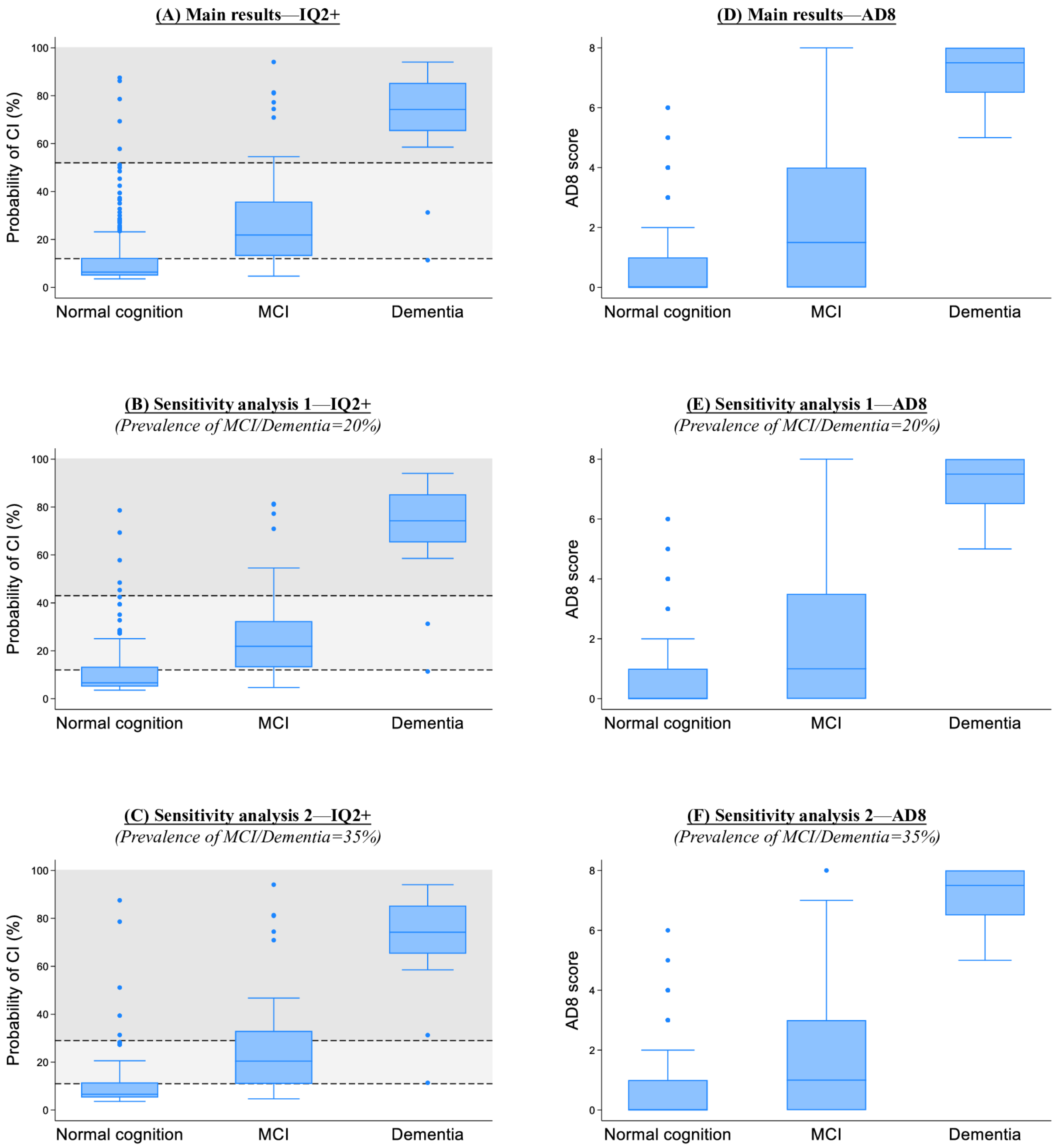
[e]Prevalence of MCI/dementia was readjusted to 20% in the test sample based on prior meta-analytic findings that community prevalence was ~15% for MCI and ~5 % for dementia. In the test sample, a subset of participants with MCI and dementia was randomly selected to readjust the prevalence in the dataset (see Methods section for further details).

[f]Probability cutoff for the first sensitivity analysis (ie, prevalence of MCI/dementia=20%):<12% (less likely to have CI), 12% to 42% (higher risk of CI), and ≥43% (likely to have CI).

[g]Prevalence of MCI/dementia was readjusted to 35% in the test sample based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI and ~10% for dementia. In the test sample, a subset of participants with MCI and dementia was randomly selected to readjust the prevalence in the dataset (see Methods section for further details).

[h]Probability cutoff for the second sensitivity analysis (ie, prevalence of MCI/dementia=35%):<11% (less likely to have CI), 11% to 28% (higher risk of CI), and ≥29% (likely to have CI).

Figure 3. Box plots showing the distribution of IQ2+ and AD8 in the test sample (n=577). Plots (A) and (B) were based on the main results from all the test samples (n=577). Plots (C) and (D) were based on results from sensitivity analysis 1, whereby the prevalence of MCI/dementia was readjusted to 20% in the test sample based on prior meta-analytic findings that community prevalence was ~15% for MCI and ~5% for dementia (in the test sample, a subset of participants with MCI and dementia were randomly selected in the test sample to readjust the prevalence in the dataset; see Methods section for further details). Plots (E) and (F) were based on results from sensitivity analysis 2, whereby the prevalence of MCI/dementia was readjusted to 35% in the test sample based on prior meta-analytic findings that community prevalence could be as high as ~25% for MCI and ~10% for dementia (in the test sample, a subset of participants with MCI and dementia were randomly selected in the test sample to readjust the prevalence in the dataset; see Methods section for further details). In plots (B), (D), and (F), the 2 horizontal dashed lines represent the 2 optimal cutoffs for IQ2+. The lower cutoff has high sensitivity and negative predictive value (>80% each) and is used to rule out MCI/dementia when probability scores fall below this threshold (as shown by the white region). The upper cutoff has high specificity and positive predictive value (>80% each) and identifies individuals likely to have MCI/dementia (when probability scores exceed this threshold, as shown by the dark gray region). The light gray region (demarcated by the lower and upper cutoffs) represents the intermediate range, identifying those who may warrant further assessment. AD8: the 8-item Informant Interview to Differentiate Aging and Dementia; IQ2+: the Informant Questionnaire for Cognitive Impairment–2 Items (plus demographics); CI: cognitive impairment; MCI/dementia: mild cognitive impairment or dementia.



Discussion

Principal Findings

In this study, we developed and validated an ultra-brief informant questionnaire (IQ2+) for the detection of cognitive impairment in literacy-diverse communities. Using a rigorous best-subset approach with cross-validation, we identified a parsimonious 4-item model that was most useful in detecting MCI/dementia, comprising 2 informant-based questions (*assistance with medications* and *worry about cognition*), along with *age* and *years of education*. In the independent test sample, IQ2+ achieved an AUC of 85.3%, outperforming the widely used AD8 (AUC=76.1%) and a baseline demographic model (AUC=78.0%). In contrast, AD8 was not significantly better than the baseline demographic model. The robust performance of IQ2+ was consistent across education subgroups and under varying prevalence scenarios, highlighting its potential utility for case finding in diverse community settings. Adopting 2 optimal cutoffs [63-67], IQ2+ demonstrated high sensitivity, specificity, negative predictive value, and positive predictive value.

Interpretation of Findings

Despite being widely recommended [26,39-42], the AD8 performed poorly in detecting MCI/dementia and was not substantially better than a basic demographic model comprising age and education. This finding aligns with recent literature, which demonstrates AD8's lower performance for detecting early cognitive impairment (AUC 61% - 69%) compared to its performance in clear-cut dementia cases (AUC 89% - 93%) [48,77]. AD8 was originally developed for the purpose of detecting dementia [30], and its items tend to focus on identifying conspicuous changes in iADL (ie, hallmarks of the onset of dementia) [46]. As a result, AD8 may possibly not be as sensitive or consistent in detecting subtle changes associated with early cognitive impairment [30,78]. In contrast, the 2 informant questions in IQ2+ may plausibly be more sensitive to early cognitive changes. Worry about cognition—especially when reported by a knowledgeable informant—has been shown to be an early symptom of cognitive decline [51]. This item requires informants to compare current abilities with previous premorbid abilities, making it more likely to detect subtle cognitive decline over time. It also prompts informants to apply a "threshold" to decide whether they are worried about the cognitive decline, which serves as a filter to identify individuals with meaningful changes. Similarly, the ability to manage one's medications is a complex task that possibly involves multiple cognitive domains [16], such as language (for understanding medication instructions), executive function (for scheduling medication intake and problem-solving when medications are not taken on schedule), and memory (for tracking medication intake). A subtle decline in any of these domains may potentially manifest as increasing difficulty in managing medications independently, serving as an early symptom of cognitive decline [79].

Implications

At the population level, IQ2+ can serve as a quick and efficient risk stratification tool, facilitating appropriate triage for further assessment. As summarized in Table 6, low-risk individuals (<12% probability of MCI/dementia) may be reassured and advised to repeat the test in 3 to 5 years or if circumstances change (eg, appearance of new symptoms). Intermediate-risk individuals (~25% - 45% probability of MCI/dementia) can be directed to further assessments using brief cognitive tests [4,32-34] to provide more conclusive evidence of cognitive impairment [27]. High-risk individuals (>80% probability of MCI/dementia) may benefit from direct referral to memory clinics for further clinical management. This approach, as summarized in Table 6, aligns with the 2-stage strategy for active case finding of cognitive impairment (ie, subjective reports, followed by brief cognitive tests) [27,28]. It offers a scalable approach to case finding in large populations by conserving resources for cognitive testing and reserving them for individuals who truly need further verification [27]. To ensure that IQ2+ remains useful in diverse populations, its optimal cutoffs can also be adjusted depending on the prevalence of MCI/dementia in different populations (as shown in Figure 1), thus providing a more tailored solution for identifying individuals at varying risk levels.

Table . Potential clinical implications based on the result from IQ2+.

Result from IQ2+[a]	Risk communication	Potential implications
Less likely to have CI[b]	<12% chance to have CI	• Repeat IQ2+ in 3 to 5 years or when circumstances change (eg, appearance of new symptoms)
Higher risk of CI	~25% to 45% chance to have CI	• Arrange with health care professionals to do brief cognitive tests
Likely to have CI	>80% chance to have CI	• Referral to memory clinic for further assessment and management

[a]IQ2+: the Informant Questionnaire for cognitive impairment–2 items (plus demographics).

[b]CI: cognitive impairment.

Essentially, IQ2+ may have 2 plausible use cases in the community. First, it allows family members who have concerns about a loved one's cognition to complete the questionnaire online (accessible at [76]), providing an immediate result with risk score and a brief interpretation of its implications. The results can also be downloaded as a PDF file for further discussion with health care providers. This approach mirrors a well-established practice in the field of diabetes mellitus, where the public is encouraged to complete an online risk test [80] as an initial screening to determine the need for further diagnostic tests [81]. By enabling self-appraisal, this use case of IQ2+ leverages family members for efficient case finding and empowers them to proactively support the brain health of older persons.

Second, IQ2+ can also be used by health care workers who routinely conduct comprehensive geriatric assessments in the community. Since 2019, the WHO has advocated for routine assessment of intrinsic capacity among community-dwelling older persons, which requires comprehensive evaluations of cognition, mobility, nutrition, hearing, vision, and mood [29]. Given the extensive range of components to cover, it can be challenging to include routine cognitive testing in all geriatric assessments. IQ2+, which can be completed in <1 minute, is well-suited to be embedded within initial geriatric assessments to prioritize individuals who require further cognitive testing. Given its excellent performance across education subgroups (Table 4), IQ2+ offers a practical tool that may potentially be broadly implemented in diverse populations, including in underserved populations in LMICs, which currently have the largest number of undiagnosed cognitive impairment [82]. This is also in line with the 2024 Lancet Commission's call to address the unmet need for cognitive screening tools that are also suited for individuals with lower literacy in LMICs [35].

Limitations

Several limitations are notable. First, IQ2+ was developed for older individuals aged ≥65 years. It is unclear whether IQ2+ would be useful for individuals aged <65 years. Second, IQ2+ would benefit from further validations in other cultures and languages, as cultural differences may affect how informants respond to its items and could potentially modify its performance in different settings. Third, IQ2+ requires the use of a web-based calculator to determine the probability of cognitive impairment. This approach leverages technology to automate test scoring, result visualization, and interpretation. However, it may limit accessibility compared to pen-and-paper tests that use raw scores as a cutoff (eg, AD8), particularly in settings with limited access to technology. Fourth, one of the core questions of IQ2+ involves difficulty managing medications. This item was identified as the most useful item in our exhaustive search (Table 3) and has some face validity (as it detects subtle changes in a complex task of managing medications [16,79], and the need to take regular medications may also be a proxy for higher risk due to the presence of chronic diseases). However, this question may be less useful among healthier older individuals who do not need to take medications regularly. Fifth, although IQ2+ can be a useful case finding tool, it is not intended to replace comprehensive clinical and neuropsychological assessments, which provide more definitive diagnoses as well as granular information on specific cognitive deficits [20,83,84].

Conclusions

Using rigorous methodology, this study developed an ultra-brief questionnaire that enables untrained laypersons to screen for cognitive impairment in <1 minute. Despite its brevity, the questionnaire demonstrated excellent performance in detecting MCI/dementia, outperforming the well-established AD8. The questionnaire can be completed by members of the public who have concerns about a family member's cognition (accessible at [76]) or embedded within community geriatric assessments to prioritize cognitive testing. Its excellent performance was consistent across education subgroups and varying prevalence scenarios, supporting its utility for case finding in diverse populations, including underserved communities and LMICs.

Acknowledgments

The authors would like to thank the following persons for their assistance in recruiting study participants: Yanling Tan, Alcantara Leicester Shawn, Kai Xin Choo, Megan Cheng Mun Choy, Xin Tong Tan, Lydia Jia En Cheong, Jane Mee Chin Liew, Xiao Hui Ng, and Spencer Peng Ming Yuen. The authors would also like to thank the following community organizations for their support in participant recruitment: Thye Hua Kwan Moral Charities (AMK 645, Bukit Merah View, Beo Crescent), Kreta Ayer Senior Activity Centres, NTUC Health Active Ageing Centre (Lengkok Bahru), Agency for Integrated Care (Caregiving & Community Mental Health Division), Silver Generation Office, People's Association, Precious Active Ageing Centre, Montfort Care, Yong-en Care Centre, Presbyterian Community Services, Lions Befrienders Service Association Singapore.

Funding

This research is funded by the Singapore Prime Minister Office's Smart Nation and Digital Government Office (grant I_20092346). Separately, TML is supported by the Singapore Ministry of Health's National Medical Research Council (grant HCSAINV23jul-0001, NMRC/CG2/005e/2022-SGH, and MOH-SEEDFD22apr-0001). The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Data Availability

The data used in this study contain sensitive information about the study participants, and they did not provide consent for public data sharing. The current approval by the SingHealth Centralized IRB of Singapore (reference number: 2021/2590) does not include data sharing. A minimal dataset could be shared by request from a qualified academic investigator for the sole purpose of replicating the present study, provided the data transfer is in agreement with Singapore's legislation on the general data protection regulation and a data sharing agreement has been signed. Contact person: corresponding author.

Authors' Contributions

Conceptualization – TML
Data curation – TML
Formal analysis – TML
Funding acquisition - TML
Investigation – consensus diagnosis – TML, KFY, KN, SKST, WL, SYT, WIK
Methodology – TML
Project administration – TML
Resources – training of research coordinators – TML, SYT, WIK
Software – TML
Supervision – TML
Validation – TML
Visualization – TML
Writing – original draft – TML
Writing – review & editing – TML, KFY, KN, SKST, WL, SYT, WIK

Conflicts of Interest

TML has provided consultation to Lundbeck and Eisai. KN has provided consultation to Takeda. The remaining authors declare no conflicts of interest.

Multimedia Appendix 1
Additional information on methodology and results.
[PDF File, 1207 KB - aging_v9i1e72963_app1.pdf]

References

1. Dementia: A Public Health Priority: World Health Organization; 2012. URL: https://www.who.int/publications/i/item/dementia-a-public-health-priority [accessed 2025-12-10]
2. Lang L, Clifford A, Wei L, et al. Prevalence and determinants of undetected dementia in the community: a systematic literature review and a meta-analysis. BMJ Open 2017 Feb 3;7(2):e011146. [doi: 10.1136/bmjopen-2016-011146] [Medline: 28159845]
3. Liu Y, Jun H, Becker A, Wallick C, Mattke S. Detection rates of mild cognitive impairment in primary care for the United States Medicare population. J Prev Alzheimers Dis 2024;11(1):7-12. [doi: 10.14283/jpad.2023.131] [Medline: 38230712]
4. Liew TM. A 4-item case-finding tool to detect dementia in older persons. J Am Med Dir Assoc 2019 Dec;20(12):1529-1534. [doi: 10.1016/j.jamda.2019.06.015] [Medline: 31399363]
5. Burns A, Iliffe S. Dementia. BMJ 2009 Feb 5;338:b75. [doi: 10.1136/bmj.b75] [Medline: 19196746]
6. Ying J, Yap P, Gandhi M, Liew TM. Iterating a framework for the prevention of caregiver depression in dementia: a multi-method approach. Int Psychogeriatr 2018 Aug;30(8):1119-1130. [doi: 10.1017/S1041610217002629] [Medline: 29223191]
7. Liew TM. Neuropsychiatric symptoms in early stage of Alzheimer's and non-Alzheimer's dementia, and the risk of progression to severe dementia. Age Ageing 2021 Sep 11;50(5):1709-1718. [doi: 10.1093/ageing/afab044] [Medline: 33770167]
8. Liew TM, Lee CS. Reappraising the efficacy and acceptability of multicomponent interventions for caregiver depression in dementia: the utility of network meta-analysis. Gerontologist 2019 Jul 16;59(4):e380-e392. [doi: 10.1093/geront/gny061] [Medline: 29860310]
9. Ang LC, Malhotra R, Roy Chowdhury A, Liew TM. Pre-and post-COVID-19 trends related to dementia caregiving on Twitter. Sci Rep 2025 Feb 12;15(1):5173. [doi: 10.1038/s41598-024-82405-8] [Medline: 39939632]
10. Thyrian JR, Hertel J, Wucherer D, et al. Effectiveness and safety of dementia care management in primary care: a randomized clinical trial. JAMA Psychiatry 2017 Oct 1;74(10):996-1004. [doi: 10.1001/jamapsychiatry.2017.2124] [Medline: 28746708]
11. Vickrey BG, Mittman BS, Connor KI, et al. The effect of a disease management intervention on quality and outcomes of dementia care: a randomized, controlled trial. Ann Intern Med 2006 Nov 21;145(10):713-726. [doi: 10.7326/0003-4819-145-10-200611210-00004] [Medline: 17116916]
12. Cepoiu-Martin M, Tam-Tham H, Patten S, Maxwell CJ, Hogan DB. Predictors of long-term care placement in persons with dementia: a systematic review and meta-analysis. Int J Geriatr Psychiatry 2016 Nov;31(11):1151-1171. [doi: 10.1002/gps.4449] [Medline: 27045271]
13. Jennings LA, Laffan AM, Schlissel AC, et al. Health care utilization and cost outcomes of a comprehensive dementia care program for Medicare beneficiaries. JAMA Intern Med 2019 Feb 1;179(2):161-166. [doi: 10.1001/jamainternmed.2018.5579] [Medline: 30575846]

14. Spijker A, Vernooij-Dassen M, Vasse E, et al. Effectiveness of nonpharmacological interventions in delaying the institutionalization of patients with dementia: a meta-analysis. J Am Geriatr Soc 2008 Jun;56(6):1116-1128. [doi: 10.1111/j.1532-5415.2008.01705.x] [Medline: 18410323]
15. Bott NT, Sheckter CC, Yang D, et al. Systems delivery innovation for Alzheimer disease. Am J Geriatr Psychiatry 2019 Feb;27(2):149-161. [doi: 10.1016/j.jagp.2018.09.015] [Medline: 30477913]
16. Elliott RA, Goeman D, Beanland C, Koch S. Ability of older people with dementia or cognitive impairment to manage medicine regimens: a narrative review. Curr Clin Pharmacol 2015;10(3):213-221. [doi: 10.2174/1574884710666150812141525] [Medline: 26265487]
17. Persson S, Saha S, Gerdtham UG, Toresson H, Trépel D, Jarl J. Healthcare costs of dementia diseases before, during and after diagnosis: longitudinal analysis of 17 years of Swedish register data. Alzheimers Dement 2022 Dec;18(12):2560-2569. [doi: 10.1002/alz.12619] [Medline: 35189039]
18. Chay J, Koh WP, Tan KB, Finkelstein EA. Healthcare burden of cognitive impairment: evidence from a Singapore Chinese health study. Ann Acad Med Singap 2024 Apr 29;53(4):233-240. [doi: 10.47102/annals-acadmedsg.2023253] [Medline: 38920180]
19. Liew TM. Distinct trajectories of subjective cognitive decline before diagnosis of neurocognitive disorders: longitudinal modelling over 18 years. J Prev Alzheimers Dis 2025 May;12(5):100123. [doi: 10.1016/j.tjpad.2025.100123] [Medline: 40057463]
20. Liew TM. Developing a brief neuropsychological battery for early diagnosis of cognitive impairment. J Am Med Dir Assoc 2019 Aug;20(8):1054-1e11. [doi: 10.1016/j.jamda.2019.02.028] [Medline: 30992186]
21. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. Lancet 2015 Jun 6;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5] [Medline: 25771249]
22. Baker LD, Espeland MA, Whitmer RA, et al. Structured vs self-guided multidomain lifestyle interventions for global cognitive function: the US POINTER randomized clinical trial. JAMA 2025 Aug 26;334(8):681-691. [doi: 10.1001/jama.2025.12923] [Medline: 40720610]
23. van Dyck CH, Swanson CJ, Aisen P, et al. Lecanemab in early Alzheimer's disease. N Engl J Med 2023 Jan 5;388(1):9-21. [doi: 10.1056/NEJMoa2212948] [Medline: 36449413]
24. Sims JR, Zimmer JA, Evans CD, et al. Donanemab in early symptomatic Alzheimer disease: the TRAILBLAZER-ALZ 2 randomized clinical trial. JAMA 2023 Aug 8;330(6):512-527. [doi: 10.1001/jama.2023.13239] [Medline: 37459141]
25. Prince M, Bryce R, Ferri C. World alzheimer report 2011: the benefits of early diagnosis and intervention. : Alzheimer's Disease International; 2011 URL: https://www.alzint.org/u/WorldAlzheimerReport2011.pdf [accessed 2025-12-10]
26. Morley JE, Morris JC, Berg-Weger M, et al. Brain health: the importance of recognizing cognitive impairment: an IAGG consensus conference. J Am Med Dir Assoc 2015 Sep 1;16(9):731-739. [doi: 10.1016/j.jamda.2015.06.017] [Medline: 26315321]
27. Liew TM. Active case finding of dementia in ambulatory care settings: a comparison of three strategies. Eur J Neurol 2020 Oct;27(10):1867-1878. [doi: 10.1111/ene.14353] [Medline: 32441837]
28. Hendry K, Hill E, Quinn TJ, Evans J, Stott DJ. Single screening questions for cognitive impairment in older people: a systematic review. Age Ageing 2015 Mar;44(2):322-326. [doi: 10.1093/ageing/afu167] [Medline: 25385272]
29. World Health Organization. Integrated Care for Older People (ICOPE): Guidance for Person-Centred Assessment and Pathways in Primary Care 2019. URL: https://www.who.int/publications/b/71300
30. Galvin JE, Roe CM, Powlishta KK, et al. The AD8: a brief informant interview to detect dementia. Neurology (ECronicon) 2005 Aug 23;65(4):559-564. [doi: 10.1212/01.wnl.0000172958.95282.2a] [Medline: 16116116]
31. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. J Gerontol 1982 May;37(3):323-329. [doi: 10.1093/geronj/37.3.323] [Medline: 7069156]
32. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The mini‐cog: a cognitive "vital signs" measure for dementia screening in multi‐lingual elderly. Int J Geriat Psychiatry 2000 Nov;15(11):1021-1027. [doi: 10.1002/1099-1166(200011)15:11<1021::AID-GPS234>3.0.CO;2-6] [Medline: 11113982]
33. Liew TM. The optimal short version of montreal cognitive assessment in diagnosing mild cognitive impairment and dementia. J Am Med Dir Assoc 2019 Aug;20(8):1055-10e1. [doi: 10.1016/j.jamda.2019.02.004] [Medline: 30910550]
34. McDicken JA, Elliott E, Blayney G, et al. Accuracy of the short‐form montreal cognitive assessment: systematic review and validation. Int J Geriat Psychiatry 2019 Oct;34(10):1515-1525. [doi: 10.1002/gps.5162]
35. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet Standing Commission. The Lancet 2024 Aug;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0]
36. Fayers PM, Machin D. Quality of Life: The Assessment, Analysis and Interpretation of Patient-Reported Outcomes, 2nd edition: John Wiley & Sons; 2007. [doi: 10.1002/9780470024522]
37. Tanwani R, Danquah MO, Butris N, et al. Diagnostic accuracy of ascertain dementia 8-item questionnaire by participant and informant-a systematic review and meta-analysis. PLoS ONE 2023;18(9):e0291291. [doi: 10.1371/journal.pone.0291291] [Medline: 37699028]

38. Taylor-Rowan M, Nafisi S, Owen R, et al. Informant-based screening tools for dementia: an overview of systematic reviews. Psychol Med 2023 Jan;53(2):580-589. [doi: 10.1017/S0033291721002002] [Medline: 34030753]
39. A 4-step process to detecting cognitive impairment and earlier diagnosis of dementia: approaches and tools for primary care providers. : The Gerontological Society of America; 2017 URL: https://issuu.com/gsastrategicalliances/docs/gsa_kaer_toolkit_4-step_process_to_detecting_cogni?fr=xKAE9_zU1NQ [accessed 2025-12-10]
40. Cordell CB, Borson S, Boustani M, et al. Alzheimer's Association recommendations for operationalizing the detection of cognitive impairment during the Medicare Annual Wellness Visit in a primary care setting. Alzheimers Dement 2013 Mar;9(2):141-150. [doi: 10.1016/j.jalz.2012.09.011] [Medline: 23265826]
41. Atri A, Dickerson BC, Clevenger C, et al. The Alzheimer's association clinical practice guideline for the diagnostic evaluation, testing, counseling, and disclosure of suspected Alzheimer's disease and related disorders (DETeCD-ADRD): validated clinical assessment instruments. Alzheimers Dement 2025 Jan;21(1):e14335. [doi: 10.1002/alz.14335] [Medline: 39713939]
42. Maslow K, Fortinsky RH. Nonphysician care providers can help to increase detection of cognitive impairment and encourage diagnostic evaluation for dementia in community and residential care settings. Gerontologist 2018 Jan 18;58(suppl_1):S20-S31. [doi: 10.1093/geront/gnx171] [Medline: 29361070]
43. The World Bank in Singapore. World Bank Group. 2024. URL: https://www.worldbank.org/en/country/singapore/overview [accessed 2025-12-02]
44. Liew TM, Foo JYH, Yang H, et al. PENSIEVE-AI a brief cognitive test to detect cognitive impairment across diverse literacy. Nat Commun 2025 Mar 23;16(1):2847. [doi: 10.1038/s41467-025-58201-x] [Medline: 40122854]
45. Patnode CD, Perdue LA, Rossom RC, et al. Screening for cognitive impairment in older adults: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2020 Feb 25;323(8):764-785. [doi: 10.1001/jama.2019.22258] [Medline: 32096857]
46. Diagnostic and Statistical Manual of Mental Disorders: DSM-5: American Psychiatric Association; 2013. [doi: 10.1176/appi.books.9780890425596]
47. Petersen RC, Morris JC. Mild cognitive impairment as a clinical entity and treatment target. Arch Neurol 2005 Jul;62(7):1160-1163. [doi: 10.1001/archneur.62.7.1160] [Medline: 16009779]
48. Kan CN, Zhang L, Cheng CY, et al. The informant AD8 can discriminate patients with dementia from healthy control participants in an Asian older cohort. J Am Med Dir Assoc 2019 Jun;20(6):775-779. [doi: 10.1016/j.jamda.2018.11.023] [Medline: 30661859]
49. Jessen F, Wiese B, Bachmann C, et al. Prediction of dementia by subjective memory impairment: effects of severity and temporal association with cognitive impairment. Arch Gen Psychiatry 2010 Apr;67(4):414-422. [doi: 10.1001/archgenpsychiatry.2010.30] [Medline: 20368517]
50. van Harten AC, Mielke MM, Swenson-Dravis DM, et al. Subjective cognitive decline and risk of MCI: the Mayo Clinic Study of Aging. Neurology (ECronicon) 2018 Jul 24;91(4):e300-e312. [doi: 10.1212/WNL.0000000000005863] [Medline: 29959257]
51. Jessen F, Amariglio RE, Buckley RF, et al. The characterisation of subjective cognitive decline. Lancet Neurol 2020 Mar;19(3):271-278. [doi: 10.1016/S1474-4422(19)30368-0] [Medline: 31958406]
52. Liew TM. Depression, subjective cognitive decline, and the risk of neurocognitive disorders. Alzheimers Res Ther 2019 Aug 9;11(1):70. [doi: 10.1186/s13195-019-0527-7] [Medline: 31399132]
53. Liew TM. Subjective cognitive decline, anxiety symptoms, and the risk of mild cognitive impairment and dementia. Alzheimers Res Ther 2020 Sep 11;12(1):107. [doi: 10.1186/s13195-020-00673-8] [Medline: 32917264]
54. Liew TM. Trajectories of subjective cognitive decline, and the risk of mild cognitive impairment and dementia. Alzheimers Res Ther 2020 Oct 27;12(1):135. [doi: 10.1186/s13195-020-00699-y] [Medline: 33109275]
55. Liew TM. Subjective cognitive decline, APOE e4 allele, and the risk of neurocognitive disorders: age- and sex-stratified cohort study. Aust N Z J Psychiatry 2022 Dec;56(12):1664-1675. [doi: 10.1177/00048674221079217] [Medline: 35229693]
56. Ng TP, Niti M, Chiam PC, Kua EH. Physical and cognitive domains of the instrumental activities of daily living: validation in a multiethnic population of Asian older adults. J Gerontol A Biol Sci Med Sci 2006 Jul;61(7):726-735. [doi: 10.1093/gerona/61.7.726] [Medline: 16870636]
57. Bestglm: best subset GLM. 2010. URL: https://cran.r-project.org/package=bestglm [accessed 2024-07-17]
58. Liew TM, Yap P. A brief, 6-item scale for caregiver grief in dementia caregiving. Gerontologist 2020;60:e1-e10. [doi: 10.1093/geront/gny161]
59. Liew TM, Yap P. A 3-item screening scale for caregiver burden in dementia caregiving: scale development and score mapping to the 22-item Zarit burden interview. J Am Med Dir Assoc 2019 May;20(5):629-633. [doi: 10.1016/j.jamda.2018.11.005] [Medline: 30591383]
60. Hosmer DW, Lemeshow S, Sturdivant RX. Applied Logistic Regression, 3rd edition: John Wiley & Sons; 2013. [doi: 10.1002/9781118548387]
61. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988 Sep;44(3):837-845. [Medline: 3203132]

62. Yu J, Yap P, Liew TM. The optimal short version of the Zarit Burden Interview for dementia caregivers: diagnostic utility and externally validated cutoffs. Aging Ment Health 2019 Jun;23(6):706-710. [doi: 10.1080/13607863.2018.1450841] [Medline: 29553806]
63. Dautzenberg GMFG, Lijmer JG, Beekman ATF. The Montreal Cognitive Assessment (MoCA) with a double threshold: improving the MoCA for triaging patients in need of a neuropsychological assessment. Int Psychogeriatr 2022 Jun;34(6):571-583. [doi: 10.1017/S1041610221000612] [Medline: 34463237]
64. Landsheer JA. Impact of the prevalence of cognitive impairment on the accuracy of the Montreal Cognitive Assessment: the advantage of using two MoCA thresholds to identify error-prone test scores. Alzheimer Dis Assoc Disord 2020;34(3):248-253. [doi: 10.1097/WAD.0000000000000365] [Medline: 31934880]
65. Thomann AE, Berres M, Goettel N, Steiner LA, Monsch AU. Enhanced diagnostic accuracy for neurocognitive disorders: a revised cut-off approach for the Montreal Cognitive Assessment. Alz Res Therapy 2020 Dec;12(1):39. [doi: 10.1186/s13195-020-00603-8]
66. Swartz RH, Cayley ML, Lanctôt KL, et al. Validating a pragmatic approach to cognitive screening in stroke prevention clinics using the Montreal Cognitive Assessment. Stroke 2016 Mar;47(3):807-813. [doi: 10.1161/STROKEAHA.115.011036] [Medline: 26903584]
67. Jack CR, Andrews JS, Beach TG, et al. Revised criteria for diagnosis and staging of Alzheimer’s disease: Alzheimer’s Association Workgroup. Alzheimer’s Dementia 2024 Aug;20(8):5143-5169. [doi: 10.1002/alz.13859]
68. Bai W, Chen P, Cai H, et al. Worldwide prevalence of mild cognitive impairment among community dwellers aged 50 years and older: a meta-analysis and systematic review of epidemiology studies. Age Ageing 2022 Aug 2;51(8):afac173. [doi: 10.1093/ageing/afac173] [Medline: 35977150]
69. Hu C, Yu D, Sun X, Zhang M, Wang L, Qin H. The prevalence and progression of mild cognitive impairment among clinic and community populations: a systematic review and meta-analysis. Int Psychogeriatr 2017 Oct;29(10):1595-1608. [doi: 10.1017/S1041610217000473] [Medline: 28884657]
70. Song WX, Wu WW, Zhao YY, et al. Evidence from a meta-analysis and systematic review reveals the global prevalence of mild cognitive impairment. Front Aging Neurosci 2023;15:1227112. [doi: 10.3389/fnagi.2023.1227112] [Medline: 37965493]
71. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP. The global prevalence of dementia: a systematic review and metaanalysis. Alzheimer's Dementia 2013 Jan;9(1):63-75. [doi: 10.1016/j.jalz.2012.11.007] [Medline: 23305823]
72. Fiest KM, Jetté N, Roberts JI, et al. The prevalence and incidence of dementia: a systematic review and meta-analysis. Can J Neurol Sci 2016 Apr;43:S3-S50. [doi: 10.1017/cjn.2016.18] [Medline: 27307127]
73. Cao Q, Tan CC, Xu W, et al. The prevalence of dementia: a systematic review and meta-analysis. J Alzheimers Dis 2020;73(3):1157-1166. [doi: 10.3233/JAD-191092] [Medline: 31884487]
74. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983 Sep;148(3):839-843. [doi: 10.1148/radiology.148.3.6878708] [Medline: 6878708]
75. Mental Capacity Act 2008. Singapore Statutes Online. URL: https://sso.agc.gov.sg/Act/MCA2008 [accessed 2024-07-17]
76. Informant questionnaire for cognitive impairment. URL: https://brainhealth-iq2plus.hf.space/ [accessed 2025-12-10]
77. Cai Y, Qiu P, Wan Y, et al. Establishing cut-off scores for the self-rating AD8 based on education level. Geriatr Nurs (Lond) 2021 Sep;42(5):1093-1098. [doi: 10.1016/j.gerinurse.2021.06.027]
78. Chen HH, Sun FJ, Yeh TL, et al. The diagnostic accuracy of the Ascertain Dementia 8 questionnaire for detecting cognitive impairment in primary care in the community, clinics and hospitals: a systematic review and meta-analysis. Fam Pract 2018 May 23;35(3):239-246. [doi: 10.1093/fampra/cmx098] [Medline: 29045636]
79. Barthold D, Marcum ZA, Chen S, et al. Difficulty with taking medications is associated with future diagnosis of Alzheimer’s disease and related dementias. J Gen Intern Med 2021 Apr;36(4):863-868. [doi: 10.1007/s11606-020-06279-y] [Medline: 33037589]
80. Take the type 2 risk test. American Diabetes Association. 2024. URL: https://diabetes.org/diabetes-risk-test [accessed 2024-07-14]
81. American Diabetes Association Professional Practice Committee. 2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes-2024. Diabetes Care 2024 Jan 1;47(Suppl 1):S20-S42. [doi: 10.2337/dc24-S002] [Medline: 38078589]
82. Kalaria R, Maestre G, Mahinrad S, et al. The 2022 symposium on dementia and brain aging in low- and middle-income countries: Highlights on research, diagnosis, care, and impact. Alzheimers Dement 2024 Jun;20(6):4290-4314. [doi: 10.1002/alz.13836] [Medline: 38696263]
83. Jacova C, Kertesz A, Blair M, Fisk JD, Feldman HH. Neuropsychological testing and assessment for dementia. Alzheimer's Dementia 2007 Oct;3(4):299-317. [doi: 10.1016/j.jalz.2007.07.011] [Medline: 19595951]
84. Ang LC, Yap P, Tay SY, Koay WI, Liew TM. Examining the validity and utility of Montreal Cognitive Assessment domain scores for early neurocognitive disorders. J Am Med Dir Assoc 2023 Mar;24(3):314-320. [doi: 10.1016/j.jamda.2022.12.028] [Medline: 36758620]

Abbreviations

AD8: 8-item Informant Interview to Differentiate Aging and Dementia
AUC: area under the receiver operating characteristic curve
iADL: instrumental activities of daily living
IQ2+: Informant Questionnaire for cognitive impairment–2 items plus demographics
LMIC: lower- and middle-income country
MCI/dementia: mild cognitive impairment or dementia
SCD: subjective cognitive decline
WHO: World Health Organization

Edited by G Seçkin; submitted 21.Feb.2025; peer-reviewed by D Urbanski, YJ Leow; revised version received 14.Aug.2025; accepted 14.Nov.2025; published 06.Jan.2026.

Please cite as:
Liew TM, Yip KF, Narasimhalu K, Ting SKS, Li W, Tay SY, Koay WI
An Ultra-Brief Informant Questionnaire for Case Finding of Cognitive Impairment Across Diverse Literacy: Diagnostic Accuracy Study
JMIR Aging 2026;9:e72963
URL: https://aging.jmir.org/2026/1/e72963
doi:10.2196/72963

From Digital Anxiety to Empowerment in Older Adults: Cross-Sectional Survey Study on Psychosocial Drivers of Digital Literacy

Han-Jen Niu[1*], PhD; Ming-Hsuan Li[1*], MS; Feng-Yu Hsieh[2], PhD; Chun-Chieh Yu[3], PhD; Chun-Ting Lin[1], MS

[1]Management Science, College of Business and Management, Tamkang University, No. 151 Ying-Chung Rd., New Taipei City, Taiwan

[2]Physical Education Instruction and Activities Section, Office of Physical Education, Tamkang University, New Taipei City, Taiwan

[3]AgriSense Agri-Food Education Association, New Taipei City, Taiwan

[*]these authors contributed equally

Corresponding Author:
Han-Jen Niu, PhD
Management Science, College of Business and Management, Tamkang University, No. 151 Ying-Chung Rd., New Taipei City, Taiwan

Abstract

Background: Amid the convergence of global population aging and accelerating digital transformation, older adults' digital adaptability has emerged as a critical indicator of their quality of life, autonomy, and capacity for successful aging. However, digital disparities, technology-related anxiety, and insufficient support systems continue to hinder older individuals from fully participating in digital society. Particularly in modern family structures—where children often live apart from aging parents—the diminishing role of family support further underscores the importance of broader social influences.

Objective: This study aims to examine how environmental factors (family support and social influence) and psychological factors (digital anxiety and sense of achievement) are associated with older adults' intention to use Assistive Digital Tools and Services (ADTS), and how these relationships contribute to the development of digital literacy. Drawing upon an integrative framework that combines constructs from the Technology Acceptance Model, the Unified Theory of Acceptance and Use of Technology, and social cognitive theory, the study also investigates the mediating and moderating mechanisms underlying these effects, offering strategic insights to support older adults in moving from social isolation to digital empowerment.

Methods: A structured questionnaire survey was conducted using a convenience sampling method among adults aged 55 years and older in Shenyang, Liaoning Province, China, yielding 480 valid responses. Structural equation modeling, bootstrapping, and moderation analysis were used to test the proposed integrative framework.

Results: For both family support and social influence, their associations with digital literacy were fully mediated by ADTS. Higher family support was associated with lower digital anxiety, which in turn correlated with greater intention, while stronger social influence was directly associated with higher intention. Digital anxiety showed a strong negative association with intention; however, this relationship was significantly weaker among those reporting a higher sense of achievement. These findings highlight usage intention as a central pathway through which environmental and psychological conditions are related to digital competence.

Conclusions: Digital literacy in later life is more than a technical skill set—it represents a vital form of psychological and social capital that empowers autonomy, well-being, and social integration. Strengthening older adults' intention to engage with digital tools through emotional reinforcement, achievement-oriented experiences, and supportive social environments is key to narrowing the digital divide. Beyond its personal benefits, fostering digital competence contributes to successful aging, which in turn brings profound advantages for families, strengthens community cohesion, and supports national goals in public health, economic participation, and social sustainability. Intergenerational learning initiatives, community-based engagement programs, and leveraging social influence to offset weakened family support can create a more inclusive, resilient, and age-friendly digital ecosystem—one that benefits not only older individuals but society at large.

(JMIR Aging 2026;9:e75245) doi:10.2196/75245

KEYWORDS

digital literacy; successful aging; family support; social influence; technology acceptance; older adults; psychological moderation; digital anxiety

Introduction

Global Aging and the Challenge of Digital Inclusion

The intersection of rapid global population aging and accelerated digital transformation has rendered digital inclusion among older adults a pressing concern for researchers and policymakers alike. According to the United Nations [1], by 2050, the global population aged 65 years and older is projected to surpass 1.6 billion, accounting for nearly 16% of the total population. This demographic shift not only places strain on health care systems and labor markets but also redefines the societal roles and needs of aging populations in an increasingly digital world.

Within this context, digital literacy—encompassing not only operational skills but also digital confidence and information navigation—has been widely recognized as a key enabler for bridging this divide and promoting the effective use of Assistive Digital Tools and Services (ADTS) [2].

Prior studies have linked digital literacy to improved psychological well-being, life satisfaction, and learning motivation among older adults [3]. Yet, how digital literacy develops under the influence of environmental and psychological factors—especially in the context of aging—remains underexplored.

Digital exclusion exacerbates social isolation and loneliness among older adults, increasing the risk of depression and cognitive decline—while digital literacy offers a potential buffer against these outcomes [4]. In digitally mediated societies, older adults may experience compounded vulnerability—excluded not only socially but technologically.

Crucially, digital literacy holds promise as a remedy to both social and digital isolation. Older adults with strong digital competencies are more likely to engage in video communication, access health services remotely, and maintain active social networks [5,6]. In this sense, digital literacy is not merely a technical skill—it is a bridge to successful aging.

COVID-19: A Double-Edged Catalyst for Digital Transformation

The COVID-19 pandemic served as a global stress test for digital readiness, rapidly shifting key aspects of daily life—health care, communication, commerce—into digital spaces. While digital services such as telemedicine and online grocery delivery became lifelines for many, they also exposed and widened the digital gap among older populations [7].

Older adults demonstrated significantly lower intention to adopt digital services during and after the pandemic, largely due to digital anxiety, low confidence, and lack of digital trust [8]. Digital anxiety encompasses fears about making mistakes, information overload, and concerns about fraud and data breaches—factors particularly salient for older first-time users [9].

While digital literacy training can alleviate anxiety and improve adoption, sustained engagement often depends on ongoing emotional and environmental support—particularly from family and community contexts [10]. However, the long-term sustainability of digital engagement often depends on external support structures such as family guidance or community-based learning. Without sustained environmental and emotional support, even trained users may regress into avoidance behaviors. This underscores the importance of examining how family and social influences shape not only technology adoption but also digital confidence and persistence among older adults.

Research Gaps and Objectives

While widely adopted models such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT) have contributed significantly to understanding digital behavior, they were primarily developed for younger, working-age users [11]. As such, these models often overlook emotional vulnerabilities and contextual factors that are particularly salient among older adults—such as digital anxiety, family support, and social influence.

At the same time, the Digital Literacy Framework highlights competencies essential for navigating the digital world but does not explicitly address how these skills develop within psychosocial environments. Moreover, existing studies tend to examine these psychological and environmental variables in isolation, without a cohesive explanatory structure [5,11].

To address this gap, we propose an integrative framework that draws on TAM, UTAUT, and social cognitive theory (SCT) [12]. In this model:

- Family support and social influence reflect environmental enablers of technology adoption (UTAUT);
- Digital anxiety and sense of achievement represent emotional and cognitive mechanisms (SCT);
- Usage intention (from TAM) serves as a behavioral mediator;
- Digital literacy is positioned as a dynamic outcome representing both skill and empowerment.

Furthermore, we incorporate a moderated mediation structure, positing that sense of achievement may buffer the negative impact of digital anxiety on usage intention, thereby shaping digital literacy development [12].

This integrative approach also aligns with Successful Aging Theory, framing digital literacy as a form of "aging capital" that enhances autonomy, social participation, and psychological well-being.

Based on this framework, our study addresses the following research questions:

1. To what extent are family support and social influence associated with older adults' intention to adopt ADTS?
2. Does digital anxiety negatively affect usage intention, and is this effect moderated by sense of achievement?
3. Does usage intention mediate the relationship between environmental/psychological factors and digital literacy?

By empirically validating this multilevel framework, we aim to offer new insights into how older adults move from digital exclusion to empowerment in an aging digital society.

In applying UTAUT as a guiding framework, we intentionally retained social influence and conceptualized family support as an environmental enabler, while excluding performance expectancy and effort expectancy. We do so for parsimony and contextual salience: (1) their predictive roles have been robustly established in prior work, so retesting them would add replication rather than novel insight; (2) prestudy interviews suggested many older adults lacked sufficient hands-on ADTS experience to meaningfully assess usefulness or ease of use; and (3) our theoretical focus is the age-specific psychosocial mechanisms (family support, social influence, anxiety, and achievement) that drive intention in later life. This selective adaptation contextualizes UTAUT rather than diluting it.

We also model digital literacy as an outcome—an accumulated form of "aging capital"—to trace how environmental and psychological conditions translate into competence via usage intention. While prior studies have treated literacy as an antecedent that can reduce anxiety and strengthen intention, our outcome-focused specification clarifies the developmental pathway we test here; it does not preclude reciprocal dynamics, which we note as an avenue for longitudinal research.

Literature Review

Family Support

Family support plays a critical role in facilitating older adults' digital engagement. Defined broadly, family support encompasses emotional encouragement, technical guidance, and intergenerational interaction that collectively enhance older adults' adaptation to digital life [13]. Caplan [14] viewed the family as a key provider of values and behavioral norms, offering essential mediation when individuals encounter challenges. Casper et al [15] further divided family support into instrumental, financial, and emotional dimensions, all of which contribute meaningfully to older adults' ability to learn and apply new technologies.

In the digital context, family members frequently act as facilitators in older adults' learning process, offering real-time help and reassurance [16]. Older adults are often influenced by their children's or grandchildren's attitudes toward technology, making family encouragement a significant determinant of their intention to engage with ADTS [17].

Empirical research supports this relationship. Xiong and Zuo [13] showed that emotional support and technical instruction from family members enhance digital literacy by reducing fear and uncertainty. Similarly, Meng et al [18] and Sosa Díaz [19] found that positive intergenerational communication increases digital confidence and learning motivation. Sosa Díaz [19] highlighted that assistance from younger family members strengthens adaptability, while Roman et al [20] emphasized how strong family bonds can boost learning confidence.

Thus, this study highlights family support as a foundational environmental factor influencing both psychological readiness and behavioral engagement in digital contexts. The focus on intergenerational interaction aims to deepen our understanding of how familial dynamics contribute to digital inclusion among older adults.

Social Influence

Social influence refers to the degree to which individuals' behaviors, attitudes, or decisions are shaped by those in their social environment—such as family, friends, or peers [11]. Rooted in the UTAUT, social influence is recognized as a key driver of technology adoption, particularly through normative pressure and perceived expectations from significant others [11,21].

From a broader psychological perspective, social influence encompasses beliefs, emotions, and behavioral patterns formed through interactions within social networks [21]. Morosan et al [22] highlighted that individuals' perceptions of others' expectations can significantly affect their own technology acceptance decisions. Karahanna et al [21] expanded this view by discussing how proximity, contact frequency, and interpersonal dynamics influence behavior, emphasizing that social structures can either facilitate or inhibit digital engagement.

Among older adults, social influence is particularly potent. Compared with younger individuals, older adults are more likely to rely on interpersonal cues and social norms when evaluating new technologies [11]. The impact of social influence is also more pronounced in collectivist cultures—such as in many East Asian societies—where conformity to group norms and maintaining social harmony are especially valued.

The role of social influence has become even more salient in the postpandemic digital era. During COVID-19, social distancing measures drove many older adults to adopt digital tools for health, communication, and daily tasks, often under the encouragement or guidance of their social circles [23]. Positive social reinforcement can lead to greater confidence and higher willingness to engage with digital services, whereas skepticism or lack of support may result in avoidance or anxiety [24-26].

Furthermore, recent studies show that social influence not only affects direct behavioral intention but also moderates psychological factors such as digital anxiety and self-efficacy [27]. These findings underscore the need to understand social influence as both an external motivator and a psychological buffer or amplifier in older adults' digital adaptation.

By focusing on the multifaceted nature of social influence—including social norms, peer encouragement, and perceived expectations—this study aims to clarify how social context shapes older adults' engagement with ADTS and the broader development of digital literacy.

Digital Anxiety: Concept and Development

Digital anxiety refers to the emotional discomfort, fear, or stress that individuals—particularly older adults—experience when interacting with digital technologies. Unlike simple unfamiliarity, digital anxiety reflects a deeper psychological resistance often rooted in low confidence, fear of failure, and perceptions of complexity or risk [26,28]. Wang and Zhang [29] defined it as a form of state anxiety, varying with task demands and context. Among older adults, this anxiety is especially pronounced due to the widening digital divide and their limited prior exposure to emerging technologies.

Rook [30] highlighted that older individuals frequently report elevated anxiety when using ADTSs such as telemedicine platforms, smartphones, or mobile payments. This is compounded when they lack the digital self-efficacy or training needed to operate these tools confidently. Straub [26] found that prior experience and perceived self-efficacy can significantly reduce anxiety and increase willingness to adopt digital tools. Similarly, the TAM posits that perceived ease of use and usefulness directly shape user attitudes and behaviors.

From the perspective of Rogers' Diffusion of Innovations Theory, perceived complexity and compatibility with past experiences are key barriers for older adults. Even simple digital interfaces can feel cognitively taxing, potentially triggering avoidance behaviors rather than active engagement [31,32].

One of the most consistently supported buffers against digital anxiety is family support. Beyond emotional reassurance, family members often provide practical technical guidance and encouragement, helping older adults feel less overwhelmed and more motivated to engage with digital tools. Studies have shown that such support fosters emotional security, strengthens digital self-confidence, and decreases feelings of uncertainty in digital environments [33]. For instance, when family members offer direct assistance—such as walking through digital tasks step-by-step—older adults are more likely to persist in their learning process and overcome initial apprehension.

However, overly reliant support can have mixed effects. While moderate, empowering support improves outcomes, excessive dependence may inadvertently signal incompetence or fuel learned helplessness, thereby reinforcing anxiety [7,34]. Nevertheless, the preponderance of evidence suggests a net protective role of family involvement in alleviating digital stress, leading to the following hypothesis:

H1: Family support negatively influences digital anxiety.

In addition to family, broader social influence plays a crucial role in shaping digital anxiety. When older adults perceive positive expectations or encouragement from peers, neighbors, or community members, it can reduce fear and promote digital exploration. On the contrary, negative social feedback—such as expressions of doubt, impatience, or age-related stereotypes—can heighten self-doubt and anxiety. Recent studies have demonstrated that social environments that are judgmental or unsupportive intensify older adults' digital apprehension, especially when they fear being seen as "incompetent" or "too old to learn" [35,36].

Conversely, positive social modeling and group-based learning environments—such as community digital workshops—can boost older adults' digital confidence and reduce anxiety levels. Venkatesh et al [11] and Cambre and Cook [27] have also pointed out that perceived social expectations (a key dimension of UTAUT) can influence both emotional responses and behavioral intentions related to technology.

H2: Social influence positively influences digital anxiety.

In summary, digital anxiety among older adults is shaped not only by individual cognitive appraisals but also by the presence—or absence—of emotional and environmental support. Understanding how family support and social influence interact with psychological states like anxiety is essential for building more inclusive and effective digital interventions.

Usage Intention of Assistive Digital Tools and Services

The intention to use ADTS among older adults is a central construct in models of technology adoption. Usage intention refers to an individual's motivational readiness and willingness to adopt or engage with a specific technology [37]. In the context of aging populations, intention plays a critical intermediary role, linking psychological and environmental factors with actual digital behavior and literacy development.

Family Support and Usage Intention

Family support has been widely documented as a catalyst for technology acceptance among older adults. Beyond alleviating anxiety, supportive family environments contribute to a more proactive stance toward digital learning. Emotional encouragement, technical guidance, and shared experiences with family members strengthen older adults' belief in their own capability, thereby enhancing their readiness to use ADTS.

Studies suggest that family members who model digital behavior or provide hands-on help not only increase access but also shape attitudes of usefulness and ease of use—2 key predictors in the TAM [38]. For instance, Zhang [17] found that the opinions of close family members significantly influenced older adults' willingness to try new technologies. Likewise, Selwyn [16] and Park [39] noted that active family involvement enhances digital confidence and curiosity, which translates into a stronger intention to use tools like mobile banking or telemedicine services.

While over-dependence on family for digital engagement may hinder independent learning, moderate and empowering support appears to foster technology acceptance. Therefore, based on the strong empirical and theoretical link between family support and technology adoption, the following hypothesis is proposed:

H3: Family support positively influences older adults' intention to use ADTS.

Social Influence and Usage Intention

Social influence extends beyond the family unit to include peers, community members, and broader societal norms that shape older adults' willingness to engage with digital technologies. In UTAUT, social influence is considered a direct antecedent of usage intention, reflecting how individuals perceive the expectations and behaviors of others as relevant to their own decision-making [11].

Research demonstrates that when older adults observe peers successfully using ADTS, they are more likely to emulate those behaviors, especially in collectivist cultures where conformity and social approval are emphasized [40,41]. Additionally, informational and normative forms of social influence—such as digital learning groups, community workshops, or word-of-mouth recommendations—play a crucial role in shaping perceived accessibility and relevance of digital services [42,43].

Moreover, positive social reinforcement enhances self-efficacy, reduces uncertainty, and increases perceived behavioral control—factors that strongly predict technology use. Conversely, negative perceptions within one's social network may lead to skepticism or hesitation. Given this evidence, we propose:

H4: Social influence positively influences older adults' intention to use ADTS.

Digital Anxiety and Usage Intention

Digital anxiety, as established in prior sections, serves as a psychological inhibitor that can undermine willingness to engage with ADTS. Fear of making mistakes, concerns over privacy, and low digital confidence create a barrier that suppresses both motivation and perceived ability.

Drawing from TAM and SCT, when anxiety increases, perceived ease of use diminishes, leading to lower adoption intention [3,9]. In real-world contexts, this means older adults who feel overwhelmed or unsupported are more likely to avoid digital tools, even when such tools offer substantial benefits.

Studies in health care technology and fintech show consistent negative correlations between digital anxiety and usage intention [44,45]. Lee et al [9] emphasized that enhancing digital self-efficacy can reduce anxiety and indirectly boost usage intention, but when anxiety remains unaddressed, intention significantly declines. Given this robust evidence, we hypothesize:

H5: Digital anxiety negatively influences older adults' intention to use ADTS.

Digital Literacy

Digital literacy is a multidimensional competency that encompasses not only the ability to use digital tools, but also to evaluate, create, and communicate digital content effectively [46,47]. It includes technical, cognitive, and socioemotional skills, all of which are crucial for navigating digital environments [48,49]. In aging populations, digital literacy has emerged as a key determinant of autonomy, participation, and well-being [50].

Research has shown that the intention to use digital tools often precedes and fosters the development of digital literacy. When older adults are motivated to use digital tools—such as smartphones, e-payment systems, or telehealth platforms—they are more likely to develop the necessary skills and confidence through experiential learning [51,52]. This iterative process suggests that usage intention is not merely an outcome but also a catalyst in digital literacy acquisition.

H6: Usage intention of ADTS positively influences digital literacy.

Beyond direct effects, usage intention also serves as a mediating mechanism through which environmental support influences digital literacy. For instance, when family members provide technical guidance and emotional encouragement, they boost older adults' confidence and motivation. This, in turn, leads to increased digital usage, which fosters skill development and digital empowerment [38,53].

H7: Usage intention of ADTS mediates the relationship between family support and digital literacy.

Similarly, social influence plays a significant role in digital skill development. Older adults who are embedded in digitally engaged social networks are more likely to receive informal training, tips, and positive reinforcement, which enhance both their usage intention and their digital competence [35,41,42]. This cascading effect suggests that social influence indirectly contributes to digital literacy via behavioral intention.

H8: Usage intention of ADTS mediates the relationship between social influence and digital literacy.

Sense of Achievement

Sense of achievement refers to the intrinsic satisfaction and psychological reward individuals experience upon successfully accomplishing a task, overcoming a challenge, or attaining a self-defined goal [54,55]. This construct is closely tied to self-efficacy, motivation, and emotional resilience—factors that are especially relevant when individuals are navigating unfamiliar or cognitively demanding environments.

For older adults adapting to digital life, a sense of achievement plays a pivotal role. Unlike younger users who often grow up immersed in digital environments, older adults frequently approach technology as late adopters. When they successfully use digital tools—such as conducting a video call, managing a mobile payment, or accessing web-based health services—this sense of accomplishment becomes a powerful motivator for further learning and sustained engagement [56]. It also fosters autonomy and contributes to their psychological well-being, aligning with key principles of successful aging.

However, the digital environment presents numerous barriers that can threaten or diminish this feeling of accomplishment. These include perceived complexity, lack of intuitive design, cybersecurity concerns, and insufficient support systems [57]. When older adults encounter repeated failure or confusion, their confidence may erode, reinforcing negative beliefs about their ability to master technology. In this context, a diminished sense of achievement can magnify digital anxiety and lead to avoidance behaviors.

Recent studies suggest that a strong sense of achievement can buffer the negative effects of digital anxiety. Older adults who perceive themselves as capable learners are more likely to persist through technological challenges and less likely to internalize failure as a reflection of their overall competence [9]. In contrast, those who lack this sense of accomplishment are more prone to technostress and self-doubt—factors that reduce usage intention [58].

Given these dynamics, this study positions sense of achievement as a moderator in the relationship between digital anxiety and the intention to use ADTS. Specifically, we propose that achievement-oriented individuals will experience a weaker negative effect of digital anxiety on usage intention, as their internal sense of progress and mastery helps to offset fear and hesitation.

H9: Sense of achievement moderates the relationship between digital anxiety and usage intention of ADTS, such that the negative effect is weaker at higher levels of achievement.

Methods

Participants

This study focuses on older adults in Shenyang, Liaoning Province, China, considering the region's demographic aging trends and digital adaptation status. As one of the major cities in Northeast China, Shenyang has a rapidly aging population, with individuals aged 60 years and older accounting for over 20% of its residents, meeting the United Nations' definition of a "deeply aging society"[96]. With the advancement of digital technologies, the internet behaviors and technological adaptation of older adults have become critical research topics. However, despite Shenyang's well-developed urban infrastructure, older adults in the region still face significant challenges in using ADTS (smartphones and e-services, etc).

Previous research has highlighted the existence of a digital divide among China's aging population, which is particularly pronounced in the northeastern region. Chu (2010) further pointed out that older individuals' digital learning and technology adaptation are influenced by educational background, social support, and technology anxiety, all of which contribute to lower adoption rates of e-services. These factors may limit older adults' ability to access web-based resources, subsequently affecting their digital quality of life [13].

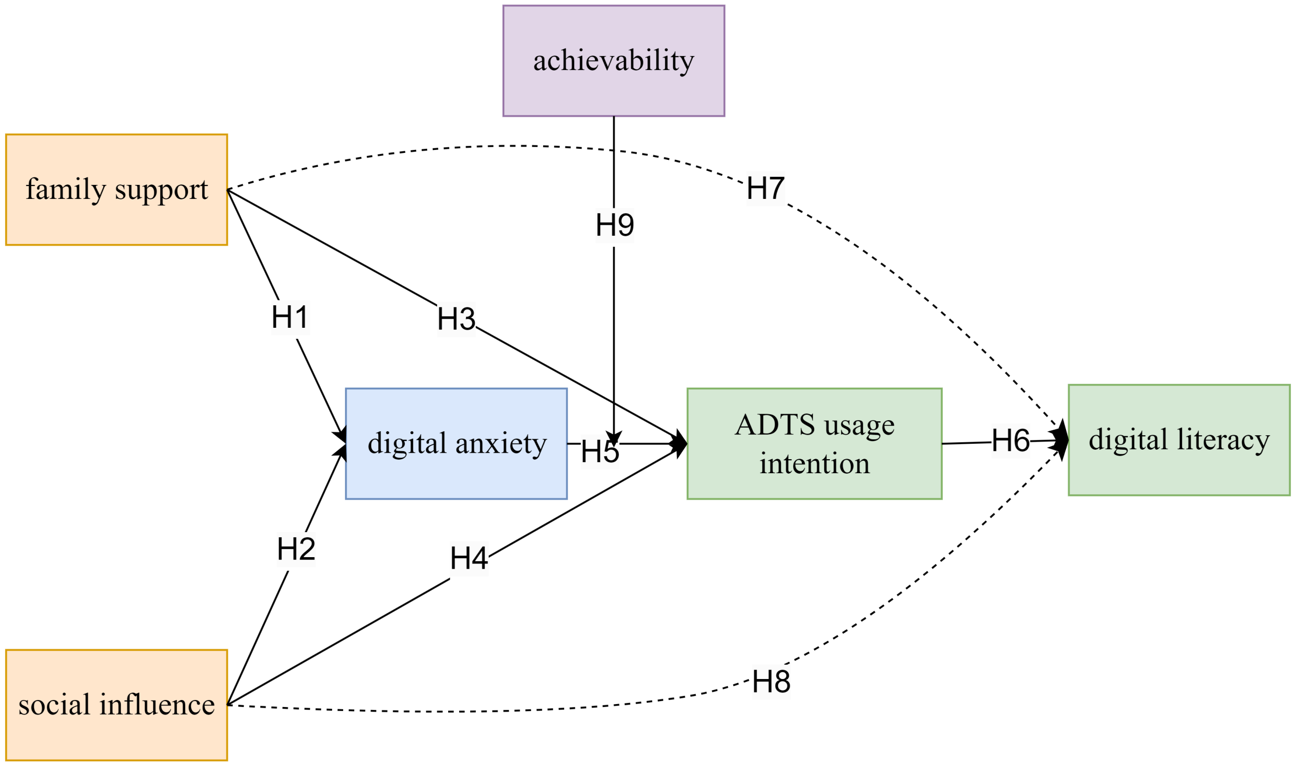
The socioeconomic structure of Shenyang also plays a crucial role in shaping older adults' digital behaviors. As a traditional industrial city, Shenyang has a significant population of retirees from state-owned enterprises, whose social support networks and digital behavior patterns differ from those in coastal cities or rural areas [5]. Research by Chen et al [5] has further demonstrated a correlation between older adults' socioeconomic status, online activity types, and susceptibility to digital fraud, emphasizing the importance of digital literacy in ensuring their safety on the web. Given these factors, this study selects Shenyang as its research site to gain deeper insights into older adults' behavioral patterns and psychological influences in digital environments.

Additionally, this study includes individuals aged 55 years and older rather than limiting the sample to traditional older adult groups. There are 2 key reasons for this decision. First, individuals in this age group are rapidly adopting internet technologies, making them a highly relevant demographic for this study. Second, early retirement is common in China, with many individuals experiencing retirement before the official retirement age. By incorporating participants aged 55 years and older, this study effectively captures the early stages of aging and the transition into digital adaptation [5].

This study used a convenience sampling method, with surveys distributed across community centers, senior activity venues, and elderly care institutions in Shenyang. Individuals aged 55 years and older were invited to voluntarily complete the questionnaire. A total of 541 responses were collected. After excluding 61 incomplete questionnaires, 480 valid responses were retained, resulting in an effective valid completion rate among respondents of 88.7%. All participants were residents of Shenyang, Liaoning Province, China.

Demographic data were collected on participants' gender, age, education level, marital status, family structure, place of residence, and smartphone usage behavior. To ensure accurate comprehension and response quality, the research team provided detailed instructions and on-site assistance during the survey process. The conceptual framework guiding this study is presented in Figure 1.

Figure 1. Research Framework. ADTS: Assistive Digital Tools and Services.



Measures

This study primarily investigates the psychological states and digital behaviors of older adults. A 6-point Likert scale was used, ranging from "strongly disagree" to "strongly agree," to measure key variables, including social influence, family support, digital anxiety, digital literacy, adoption intention of ADTS, and sense of achievement. The measurement instruments were derived from well-established scales in the existing literature:

Family support was adapted from Wang and Wu [28], originally developed in the context of older adults' digital adaptation in urban China. The scale assesses 3 core dimensions of support provided by family members: emotional encouragement, technical assistance, and informational guidance. While Wang and Wu's [28] study focused on digital literacy outcomes, the psychometric properties of the family support scale were independently validated through confirmatory factor analysis (CFA) and internal consistency measures (Cronbach α=0.82). This scale was chosen for its cultural contextual relevance and conceptual alignment with our study focus. In our sample, Cronbach α was 0.802.

Social influence was adapted from Venkatesh et al [11] in the UTAUT. The scale evaluates the extent to which individuals perceive that important others—such as family members, peers, or community figures—believe they should use digital technologies. Each item was rated on a 6-point Likert scale. This construct has been validated in numerous technology acceptance studies across diverse populations, including older adults. In this study, the scale demonstrated good internal consistency (Cronbach α=0.798).

Digital anxiety was also adapted from Venkatesh et al [11]. The items reflect the UTAUT construct of anxiety, which captures individuals' apprehension, fear, or discomfort when interacting with digital technology. Although originally tested in younger user populations, the scale has been applied and adapted in aging research. In this study, internal consistency was acceptable (Cronbach α=0.762), supporting its reliability among older adult users.

Intention to use ADTS was adapted from the UTAUT model (Venkatesh et al [11]), assessing respondents' willingness to engage with digital services such as e-health platforms, digital payments, and digital communication tools. Each item was rated on a 6-point Likert scale. The original scale has been validated in both general and older adult populations. In this study, Cronbach α was .789, indicating acceptable reliability.

Digital literacy was evaluated using the eHealth Literacy Scale developed by Norman and Skinner [59], consisting of 8 items that measure perceived ability to find, understand, evaluate, and apply digital health information. This scale has been widely adopted in aging, health, and digital divide studies and has demonstrated strong psychometric properties across diverse populations. In our study, the scale yielded a Cronbach α of 0.871, indicating high internal consistency.

Sense of achievement was measured using a 3-item subscale from the achievement motivation framework developed by Janke and Lüftenegger [50]. The items reflect individuals' feelings of satisfaction and competence when overcoming digital challenges. The original instrument was validated in academic and motivational contexts and has since been adapted in studies of digital learning among older adults. The internal consistency for this scale in our sample was strong (Cronbach α=0.851).

To ensure the statistical adequacy of structural equation modeling (SEM), this study adhered to the sample size recommendations proposed by Bentler and Chou [60], which suggest that the sample size should be at least 5 times the number of estimated parameters. Additionally, data screening procedures, including missing data analysis and normality tests, were conducted to ensure the suitability of the dataset for subsequent analysis. Ultimately, a total of 480 valid responses were collected, meeting the theoretical requirements for robust statistical analysis.

Reliability (Cronbach α), CFA, and SEM were conducted using SPSS (version 26.0; IBM Corp) and AMOS (version 24.0; IBM Corp). Bootstrapping (n=5000) and PROCESS Macro (Model 1) were used to test mediation and moderation effects, ensuring robust statistical validity.

Ethical Considerations

This study utilized fully anonymized survey data that contained no personally identifiable information. As such, ethical approval and informed consent were not required under institutional and national research ethics regulations. All research procedures adhered to internationally recognized ethical standards, ensuring transparency, data integrity, and the responsible handling of anonymized information.

Results

Sample Profile

Regarding gender distribution, 69% (n=331) of respondents were male, and 31% (n=149) were female. In terms of age, the majority were aged between 60 and 65 years (263/480, 54.8%), followed by those aged 65 to 70 years (167/480, 34.8%), while only 1.0% (n=5) were aged 75 years and older. Residential distribution indicated that 60.4% (n=290) of participants lived in urban areas, whereas 39.6% (n=190) resided in rural regions. In terms of educational attainment, 57.1% (n=274) of respondents had completed middle school or below, 23.5% (n=111) held a university degree, and 1.5% (n=7) had attained a postgraduate degree or higher. Regarding marital status, 87.5% (n=420) of respondents were married, 4.8% (n=23) were unmarried, 2.9% (n=14) were divorced, and 4.8% (n=23) were widowed. As for family structure, 74.8% (n=359) lived with their spouse, 12.5% (n=56) lived with their children, and 12.7% (n=61) lived alone.

Regarding the usage of ADTS, 85.0% (n=408) of respondents primarily used smartphones for voice calls, followed by text messaging (381/480, 79.4%) and video calls (357/480, 74.4%). In contrast, mobile shopping (261/480, 54.3%) and digital payments (271/480, 56.5%) were less frequently used. Notably, only 53.4% (n=256) of participants used their smartphones for digital information retrieval, suggesting that a considerable portion of older adults remains unfamiliar with internet-based search functions.

Reliability Analysis

To assess the internal consistency of the measurement scales, this study calculated Cronbach α for each variable. The results, presented in Table 1, indicate that all Cronbach α values exceeded the recommended threshold of 0.7, demonstrating good reliability across the constructs.

Table . The results of reliability analysis (N=480). All constructs were measured on a 6-point Likert scale (1=strongly disagree, 6=strongly agree).

Variable	Values, mean (SD)	Values, Cronbach α
Social influence	2.850 (1.023)	0.798
Family support	4.960 (0.803)	0.802
Digital anxiety	4.980 (0.784)	0.762
ADTS[a] usage intention	5.030 (0.844)	0.789
Digital literacy	4.510 (0.919)	0.871
Achievability	5.380 (0.817)	0.851

[a]ADTS: Assistive Digital Tools and Services.

Among the variables, Digital Literacy exhibited the highest internal consistency, with a Cronbach α of 0.871, suggesting strong reliability. Meanwhile, Digital Anxiety had a Cronbach α of 0.762, which remains within an acceptable range, indicating sufficient internal consistency for further analysis.

Confirmatory Factor Analysis

To ensure convergent validity and discriminant validity of the measurement model, this study conducted CFA. Following the recommendations of Hair et al [61], composite reliability should exceed 0.7, and average variance extracted should be above 0.5 to confirm the internal consistency and validity of the measurement instruments.

As shown in Table 2, all constructs in this study met these criteria, with composite reliability values exceeding 0.7 and average variance extracted values above 0.5, indicating strong convergent validity. Additionally, all standardized factor loadings were statistically significant (>0.60), demonstrating that the observed variables effectively captured their respective latent constructs.

Table . Results of convergent validity analysis (N=480).

Concept and item	Standard factor loading	Composite reliability	Average variance extracted
Social influence (SI)		0.802	0.506
SI1	0.692		
SI3	0.818		
SI4	0.631		
F1	0.670		
Family support (F)		0.803	0.504
F3	0.705		
F4	0.748		
F5	0.716		
D2	0.586		
Digital literacy (D)		0.872	0.535
D4	0.690		
D5	0.696		
D6	0.808		
D7	0.798		
D8	0.785		
ADTS1[a]	0.765		
ADTS usage intention (ADTS)		0.802	0.578
ADTS2	0.865		
ADTS3	0.633		
DA1	0.704		
Digital anxiety (DA)		0.803	0.506
DA2	0.732		
DA3	0.729		
DA4	0.678		
A1	0.804		
Achievability (A)		0.880	0.711
A2	0.950		
A3	0.765		

[a]ADTS: Assistive Digital Tools and Services.

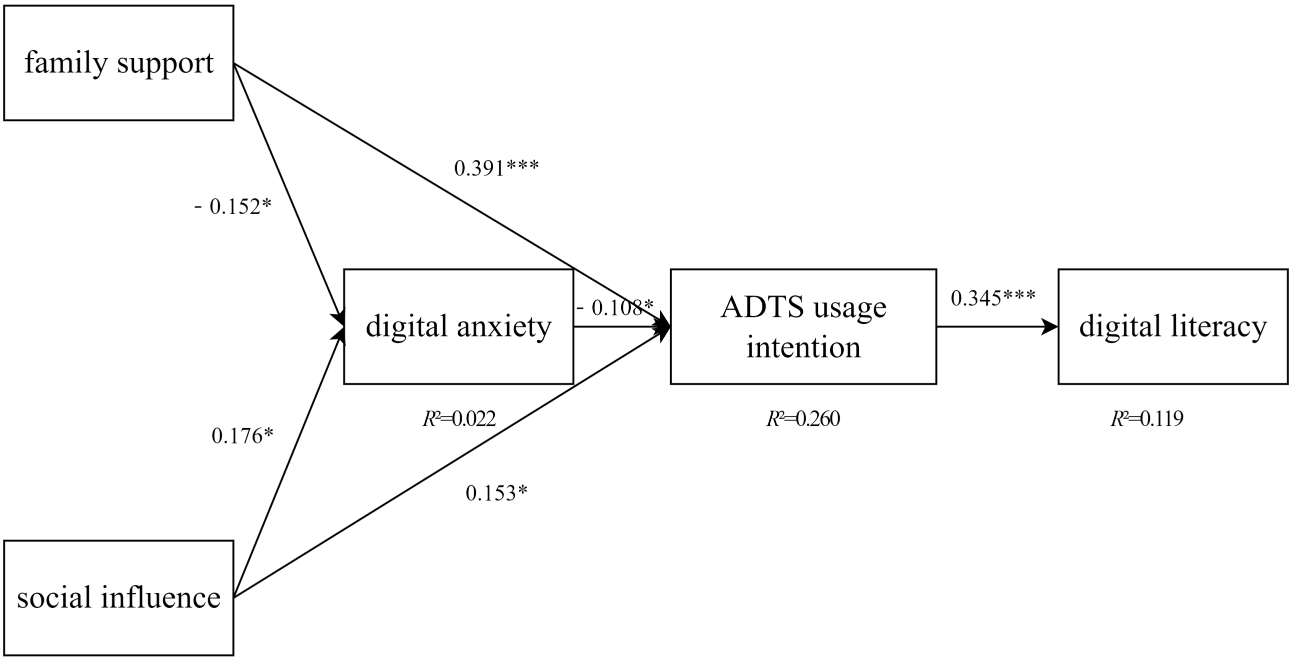
Overall, the results confirm that the measurement tools used in this study exhibit high reliability and validity, ensuring their suitability for further SEM analysis.

Structural Equation Modeling (SEM) Analysis

Model Fit Assessment

This study used SEM to validate the proposed research framework, as illustrated in Figure 2. Prior to conducting SEM analysis, we assessed the model fit indices to ensure the suitability of the model.

Figure 2. The model of structural equation modeling. ADTS: Assistive Digital Tools and Services. *P<.001, **P<.01, ***P<.05.



By examining the Modification Index, we identified and removed items with large residual discrepancies, including 2 items from the Family Support scale and items 1 and 3 from the Digital Literacy scale. After these modifications, the revised model exhibited improved goodness-of-fit indices, confirming its appropriateness for further SEM analysis (Table 3).

Table . Goodness of model fit indices (N=480).

Goodness of Fit Index	Scholars	Goodness fit range	Modified model (Delete SI2, D1, D3)
CMIN/DF[a]	Joreskog and Sorbom [52]	<3.00	2.499
GFI[b]	Doll, Xia, and Torkzadeh [53]	>0.8	0.925
AGFI[c]	MacCallum and Hong [62]	>0.8	0.902
SRMR[d]	Hu and Bentler [63]	<0.08	0.057
RMSEA[e]	MacCallum, Browne, and Sugawara [54]	<0.08	0.053
NFI[f]	Hair et al [61]	>0.8	0.902
NNFI[g]	Bentler and Bonett [55]	>0.8	0.884
CFI[h]	Hair et al [61]	>0.8	0.941
PNFI[i]	Mulaik et al [64]	>0.5	0.751
PGFI[j]	Mulaik et al [64]	>0.5	0.705
CN[k]	Mulaik et al [64]	>200	259

[a]CMIN/DF: chi-square minimum/degrees of freedom.
[b]GFI: Goodness-of-Fit Index.
[c]AGFI: Adjusted Goodness-of-Fit Index.
[d]SRMR: Standardized Root Mean Square R.
[e]RMSEA: Root Mean Square Error of Approximation.
[f]NFI: Normed Fit Index.
[g]NNFI: Non-Normed Fit Index.
[h]CFI: Comparative Fit Index.
[i]PNFI: Parsimony Normed Fit Index.
[j]PGFI: Parsimony Goodness-of-Fit Index.
[k]CN: Hoelter's Critical N.

Structural Equation Modeling (SEM) Analysis

The SEM analysis in this study was conducted in 2 stages. First, we examined the overall model fit indices to ensure the model's suitability. Second, we tested the path coefficients to validate the research hypotheses.

By performing path analysis on the latent variables, we assessed the relationships among key constructs. The results are presented in Table 4 and Figure 2.

Table . The results of structural equation modeling path analysis.

Hypothesis	Independent variable	Dependent variable	β	SE	Composite reliability	*P* value
H1	F[a]	A[b]	−.152	0.080	–2.087	<.037
H2	SI[c]	A	.176	0.081	2.362	<.018
H3	F	ADTS[d]	.391	0.077	5.356	<.001
H4	SI	ADTS	.153	0.071	2.235	<.025
H5	A	ADTS	−.108	0.049	–2.112	<.035
H6	ADTS	D[e]	.345	0.041	6.031	<.001

[a]F: Family support.
[b]A: Achievability.
[c]SI: Social influence.
[d]ADTS: Assistive Digital Tools and Services.
[e]D: Digital literacy.

Mediating Effect

This study used the PROCESS macro in SPSS (version 26.0; IBM Corp) with the bootstrap method to examine mediating effects. The analysis tested whether usage intention of ADTS mediated the relationships between (1) family support and digital literacy, and (2) social influence and digital literacy.

For the family support → digital literacy pathway (Table 5), the indirect effect was 0.071, with a 95% CI of (0.095-0.262), indicating statistical significance as the interval did not include zero. In contrast, the direct effect was 0.397, with a 95% CI of (–0.130 to 0.082), which was nonsignificant. This pattern reflects full mediation, meaning that family support influences digital literacy entirely through its impact on usage intention.

Table . Mediating effect analysis (family support and digital literacy).

F→ADTS→D[a,b,c]	Effect	Bootstrapping (BC[d] 95% CI)
Indirect effect	0.071	0.095 to 0.262
Direct effect	0.397	–0.130 to 0.082
Total effect	0.468	0.003 to 0.052

[a]F: family support.

[b]ADTS: Assistive Digital Tools and Services.

[c]D: digital literacy.

[d]BC: bias-corrected percentile method.

For the social influence → digital literacy pathway (Table 6), the indirect effect was also significant (95% CI excluding zero), whereas the direct effect was −0.418 with a 95% CI of (−0.141 to 0.186), which was nonsignificant. This likewise indicates full mediation, suggesting that social influence affects digital literacy only indirectly via usage intention.

Table . Mediating effect analysis (social influence and digital literacy).

SI→ADTS→D[a,b,c]	Effect	Bootstrapping (BC[d] 95% CI)
Indirect effect	0.128	0.069 to 0.194
Direct effect	–0.418	–0.141 to 0.186
Total effect	0.086	0.014 to 0.186

[a]SI: social influence.

[b]ADTS: Assistive Digital Tools and Services.

[c]D: digital literacy.

[d]BC: bias-corrected percentile method.

Overall, these findings demonstrate that both family support and social influence enhance older adults' digital literacy not by directly improving their skills, but by strengthening their intention to use ADTS—thereby facilitating subsequent skill acquisition and competence development.

Moderation Effect

The moderating effect of sense of achievement on the relationship between digital anxiety and ADTS usage intention was examined using PROCESS Model 1 with 5000 bootstrap samples. The overall model was significant (R^2=0.389, $F_{1, 476}$=101.222, P=.003), and the interaction term (Digital Anxiety×Sense of Achievement) was also significant, indicating the presence of a moderation effect.

Simple-slopes analysis revealed that the negative association between digital anxiety and ADTS usage intention was weaker at higher levels of sense of achievement and stronger at lower levels. This suggests that a greater sense of achievement can buffer the detrimental impact of digital anxiety on the intention to use ADTS. Figures 3 and 4 depict the slope differences across varying levels of sense of achievement, providing visual confirmation of the moderation effect.

In line with Self-Determination Theory, these results indicate that perceived competence and mastery help sustain motivation when engaging in tasks that may provoke anxiety. Accordingly, H9 is supported: sense of achievement attenuates the negative effect of digital anxiety on ADTS usage intention.

Figure 3. Moderation path coefficients. ADTS: Assistive Digital Tools and Services. *P<.05.

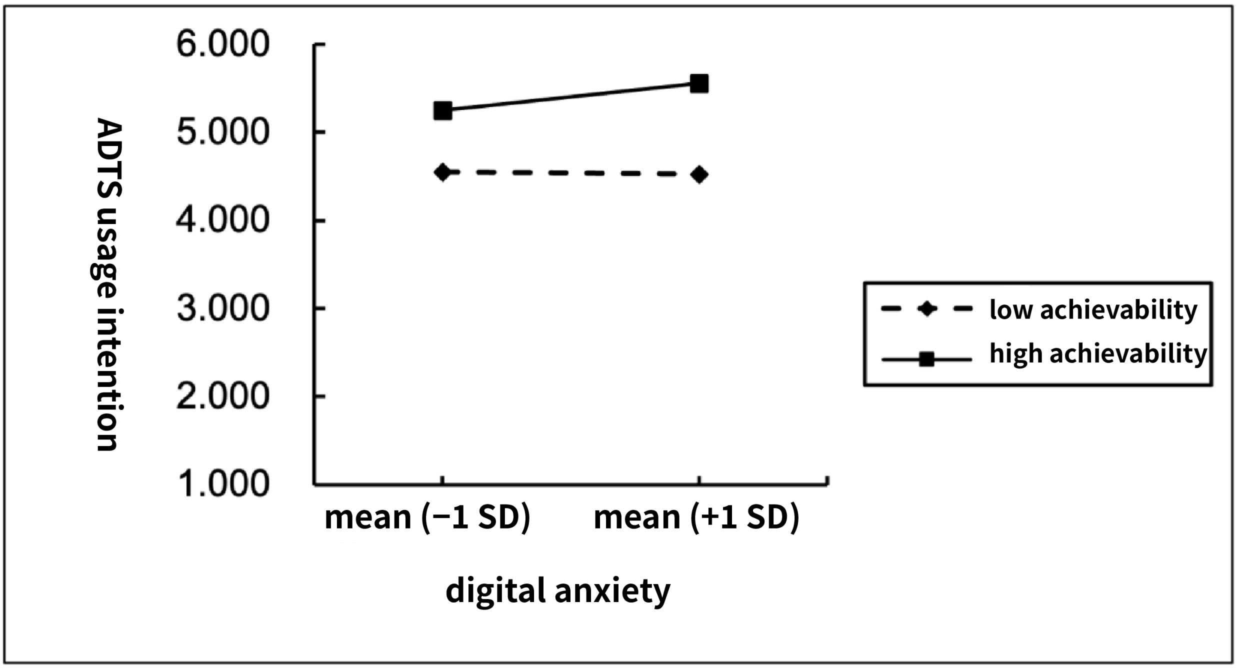
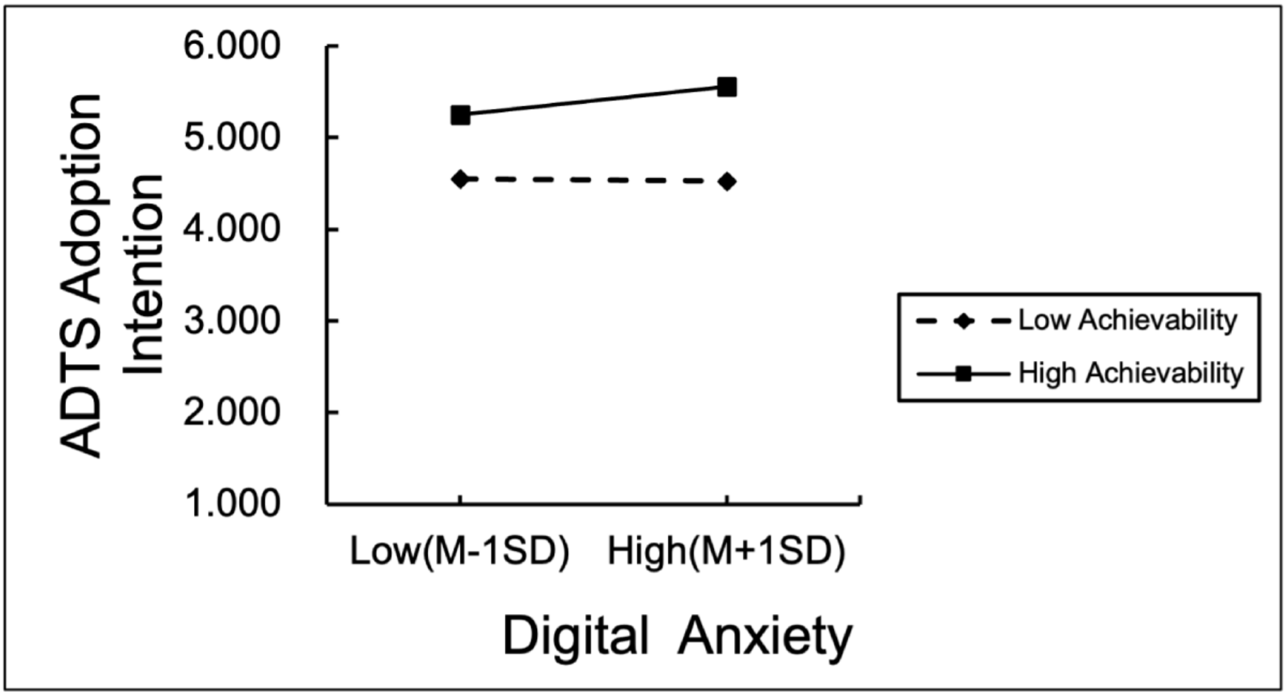


Figure 4. Moderation slope analysis. ADTS: Assistive Digital Tools and Services.



Discussion

Summary of Key Findings

In an era marked by rapid digital transformation and global population aging, older adults' ability to adapt to digital technologies has become a critical determinant of quality of life, autonomy, and the trajectory of successful aging [2,3]. Grounded in the UTAUT and the Digital Literacy Framework, this cross-sectional study tested an integrated model that incorporated environmental, psychological, and behavioral factors. Specifically, it examined how family support, social influence, digital anxiety, and sense of achievement are associated with older adults' intention to use ADTS, and how these factors, in turn, relate to the development of digital literacy. The findings provide new insights into how digital engagement can be transformed into real-life competence and social connection.

Family Support and Social Influence Operate Through Intention

Both factors were positively associated with ADTS usage intention (β=.413 and β=.160, respectively, $P<.05$). Mediation analysis showed that their effects on digital literacy were fully mediated by usage intention, with nonsignificant direct paths after accounting for the mediator.

Family Support Was Associated With Lower Digital Anxiety

Family support was negatively associated with digital anxiety (β=−0.168, $P<.05$), which in turn was negatively associated with usage intention (β=−.103, $P<.05$). This suggests that family support may help older adults overcome apprehension toward technology, indirectly promoting adoption.

Sense of Achievement Attenuates the Association of the Negative Effect of Digital Anxiety

The interaction term (Digital Anxiety × Sense of Achievement) was significant, indicating that a higher sense of achievement attenuates the negative relationship between digital anxiety and usage intention.

Usage Intention Is a Central Behavioral Pathway

Across environmental and psychological predictors, usage intention emerged as the pivotal link to higher digital literacy scores (β=.345, $P<.001$).

Theoretical Implications

This study advances the literature by showing how psychosocial factors operate within an integrative, later-life adaptation of UTAUT. Specifically, it contributes three distinctive insights—each grounded in the SEM results:

1. Extending UTAUT with age-specific psychosocial mechanisms: Beyond utilitarian appraisals, digital anxiety exerts a negative effect on intention (β=−0.103, $P<.05$), while sense of achievement attenuates this link (significant interaction), indicating that affective constraints and motivational buffers are integral to older adults' adoption processes.
2. Unpacking the pathway from environmental enablers to competence: The effects of family support (β=.413) and social influence (β=.160) on digital literacy are fully mediated by usage intention (indirect effects with 95% CIs excluding zero), positioning intention as a necessary conduit from context to competence.
3. Reframing successful aging for the digital era: Digital literacy functions as aging capital—a capability accumulated through intentional engagement and shaped by psychosocial conditions—thereby warranting its incorporation into the theoretical core of successful aging in technology-mediated societies.

Practical / Managerial Implications

The findings translate into the following action priorities for policymakers, community organizations, and technology designers:

1. Leverage family and peer networks as complementary pillars: Given the stronger association of family support with intention (β=.413) relative to social influence (β=.160), intergenerational pairings can serve as high-impact entry points. To avoid over-dependence, pair these with peer-led workshops, senior tech clubs, and neighborhood "digital ambassadors" for continuity.
2. Design for progressive achievement to counter anxiety: Because achievement buffers the anxiety-intention link, use milestone-based modules, micro-goals (eg, a first mobile payment), and instant positive feedback to convert apprehension into mastery and sustained engagement.
3. Prioritize sustained, habitual use over one-off adoption: With literacy gains operating through intention, build reinforcement sessions, periodic "digital check-ins," and adaptive nudges (reminders, personalized suggestions, calibrated difficulty) to routinize use and consolidate skills. These recommendations move beyond generic calls for support, advancing age-sensitive, culturally grounded strategies that directly target the psychosocial dynamics revealed by our model.

Future Research and Practice Directions

Limitations and Future Directions

The cross-sectional design precludes causal inference; results reflect associations within one metropolitan sample, limiting generalizability. Demographic covariates (eg, age, education, marital status) were collected but excluded from the structural model to preserve parsimony; future work can incorporate them to refine estimates.

Modeling Choices

We intentionally excluded performance expectancy and effort expectancy to foreground later-life psychosocial mechanisms and avoid redundancy with well-established findings; future studies should reincorporate these UTAUT constructs to test robustness and generalizability. We also modeled digital literacy as an outcome to trace competence formation; longitudinal or cross-lagged designs should examine possible bidirectional effects whereby literacy reduces anxiety, enhances self-efficacy, and strengthens intention.

Next Steps

We recommend longitudinal/experimental designs, cross-cultural comparisons, and the inclusion of digital trust/privacy and emerging AI/IoT tools to evaluate scalable, age-friendly pathways to digital inclusion.

Conflicts of Interest

None declared.

References

1. Department of Economic and Social Affairs. World Population Prospects 2024: Summary of Results: New York: United Nations; 2024.
2. Radovanović D, Holst C, Belur SB, et al. Digital literacy key performance indicators for sustainable development. SI 2020;8(2):151-167. [doi: 10.17645/si.v8i2.2587]
3. Lee H. Analysis of the impact of digital literacy on life satisfaction (2019–2022) for older adults in South Korea: a national community-based panel study. Sci Rep 2024;14(1):20399. [doi: 10.1038/s41598-024-71397-0]
4. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. Ann Behav Med 2010 Oct;40(2):218-227. [doi: 10.1007/s12160-010-9210-8] [Medline: 20652462]
5. Chen L, Jia J, Xiao M, Wu C, Zhang L. A study on the influence of digital literacy on elderly user's intention to identify social media false information. EL 2024 Sep 23;42(5):701-721. [doi: 10.1108/EL-10-2023-0257]
6. Wang D, Duan Y, Jin Y. Navigating online perils: socioeconomic status, online activity lifestyles, and online fraud targeting and victimization of old adults in China. Comput Human Behav 2025 Jan;162:108458. [doi: 10.1016/j.chb.2024.108458]
7. Majeed A, Zhang X. On the adoption of modern technologies to fight the COVID-19 pandemic: a technical synthesis of latest developments. COVID 2023;3(1):90-123. [doi: 10.3390/covid3010006]
8. Alassaf P, Szalay ZG. The impact of 'Compulsory' shifting to use e-services during COVID-19 pandemic restrictions period on e-services users' future attitude and intention "Case Study of Central European Countries/Visegrád Group (V4)". Sustainability 2022;14(16):9935. [doi: 10.3390/su14169935]
9. Lee K, Kim S, Kim SH, et al. Digital health interventions for adult patients with cancer evaluated in randomized controlled trials: scoping review. J Med Internet Res 2023 Jan 6;25:e38333. [doi: 10.2196/38333] [Medline: 36607712]
10. Miller LMS, Callegari RA, Abah T, Fann H. Digital literacy training for low-income older adults through undergraduate community-engaged learning: single-group pretest-posttest study. JMIR Aging 2024 May 14;7:e51675. [doi: 10.2196/51675] [Medline: 38599620]
11. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
12. Crowther MR, Parker MW, Achenbaum WA, Larimore WL, Koenig HG. Rowe and Kahn's model of successful aging revisited: positive spirituality--the forgotten factor. Gerontologist 2002 Oct;42(5):613-620. [doi: 10.1093/geront/42.5.613] [Medline: 12351796]
13. Xiong J, Zuo M. How does family support work when older adults obtain information from mobile internet? ITP 2019 Dec 2;32(6):1496-1516. [doi: 10.1108/ITP-02-2018-0060]
14. Caplan N. Social research and national policy: what gets used, by whom, for what purposes, and with what effects. Int Soc Sci J 1976;28(1):187-194 [FREE Full text]
15. Chopdar PK, Korfiatis N, Sivakumar VJ, Lytras MD. Mobile shopping apps adoption and perceived risks: a cross-country perspective utilizing the unified theory of acceptance and use of technology. Comput Human Behav 2018 Sep;86:109-128. [doi: 10.1016/j.chb.2018.04.017]
16. Selwyn N. Reconsidering political and popular understandings of the digital divide. New Media Soc 2004 Jun;6(3):341-362. [doi: 10.1177/1461444804042519]
17. Zhang M. Older people's attitudes towards emerging technologies: a systematic literature review. Public Underst Sci 2023 Nov;32(8):948-968. [doi: 10.1177/09636625231171677]
18. Meng Q, Yan Z, Abbas J, Shankar A, Subramanian M. Human–computer interaction and digital literacy promote educational learning in pre-school children: mediating role of psychological resilience for kids' mental well-being and school readiness. International Journal of Human–Computer Interaction 2025 Jan 2;41(1):16-30. [doi: 10.1080/10447318.2023.2248432]
19. Sosa Díaz MJ. Emergency remote education, family support and the digital divide in the context of the COVID-19 lockdown. Int J Environ Res Public Health 2021 Jul 28;18(15):7956. [doi: 10.3390/ijerph18157956] [Medline: 34360248]
20. Roman NV, Balogun TV, Butler-Kruger L, et al. Strengthening family bonds: a systematic review of factors and interventions that enhance family cohesion. Soc Sci (Basel) 2025;14(6):371. [doi: 10.3390/socsci14060371]
21. Karahanna E, Straub DW, Chervany NL. Information technology adoption across time: a cross-sectional comparison of pre-adoption and post-adoption beliefs. MIS Q 1999 Jun;23(2):183. [doi: 10.2307/249751]
22. Morosan C, DeFranco A. It's about time: revisiting UTAUT2 to examine consumers' intentions to use NFC mobile payments in hotels. Int J Hosp Manag 2016 Feb;53:17-29. [doi: 10.1016/j.ijhm.2015.11.003]
23. Peral-Peral B, Villarejo-Ramos Á, Arenas-Gaitán J. Self-efficacy and anxiety as determinants of older adults' use of internet banking services. Univ Access Inf Soc 2020 Nov;19(4):825-840. [doi: 10.1007/s10209-019-00691-w]
24. Arenas-Gaitán J, Villarejo Ramos AF, Peral-Peral B. A posteriori segmentation of elderly internet users: applying PLS-POS. MIP 2020 May 13;38(3):340-353. [doi: 10.1108/MIP-01-2019-0057]
25. Igbaria M, Iivari J. The effects of self-efficacy on computer usage. Omega (Westport) 1995 Dec;23(6):587-605. [doi: 10.1016/0305-0483(95)00035-6]
26. Straub ET. Understanding technology adoption: theory and future directions for informal learning. Rev Educ Res 2009 Jun;79(2):625-649. [doi: 10.3102/0034654308325896]
27. Cambre MA, Cook DL. Computer anxiety: definition, measurement, and correlates. Journal of Educational Computing Research 1985 Feb;1(1):37-54. [doi: 10.2190/FK5L-092H-T6YB-PYBA]

28. Wang CH, Wu CL. Bridging the digital divide: the smart TV as a platform for digital literacy among the elderly. Behav Inf Technol 2022 Sep 10;41(12):2546-2559. [doi: 10.1080/0144929X.2021.1934732]
29. Wang W, Zhang Y, Zhao J. Technological or social? influencing factors and mechanisms of the psychological digital divide in rural Chinese elderly. Technol Soc 2023 Aug;74:102307. [doi: 10.1016/j.techsoc.2023.102307]
30. Rook KS. Parallels in the study of social support and social strain. J Soc Clin Psychol 1990 Mar;9(1):118-132. [doi: 10.1521/jscp.1990.9.1.118]
31. Carlo S, Nanetti S. Understanding ICTS in older life: a scope review of 'The Gerontologist' and 'Research on Aging'. 2023 Jul Presented at: Human Aspects of IT for the Aged Population: 9th International Conference, ITAP 2023, Held as Part of the 25th HCI International Conference, HCII 2023; Jul 23-28, 2023; Copenhagen, Denmark p. 426-442. [doi: 10.1007/978-3-031-34917-1_29]
32. Valta M, Hildebrandt Y, Maier C. Fostering the digital mindset to mitigate technostress: an empirical study of empowering individuals for using digital technologies. INTR 2024 Nov 25;34(6):2341-2369. [doi: 10.1108/INTR-09-2022-0766]
33. Mariano J, Marques S, Ramos MR, et al. Too old for technology? Stereotype threat and technology use by older adults. Behav Inf Technol 2022 May 19;41(7):1503-1514. [doi: 10.1080/0144929X.2021.1882577]
34. Helsper EJ. The social relativity of digital exclusion: applying relative deprivation theory to digital inequalities. Commun Theor 2017 Aug;27(3):223-242. [doi: 10.1111/comt.12110]
35. Pejić Bach M, Ivančić L, Bosilj Vukšić V, Stjepić AM, Milanović Glavan L. Internet usage among senior citizens: self-efficacy and social influence are more important than social support. JTAER 2023;18(3):1463-1483. [doi: 10.3390/jtaer18030074]
36. Jang HW, Moon C, Jung HS, Cho M, Bonn MA. Normative and informational social influence affecting digital technology acceptance of senior restaurant diners: a technology learning perspective. Int J Hosp Manag 2024 Jan;116:103626. [doi: 10.1016/j.ijhm.2023.103626]
37. Balki E, Hayes N, Holland C. Loneliness and older adults: psychological resilience and technology use during the COVID-19 pandemic-a cross sectional study. Front Aging 2023;4:1184386. [doi: 10.3389/fragi.2023.1184386] [Medline: 37434741]
38. Eshet Y. Digital literacy: a new terminology framework and its application to the design of meaningful technology-based learning environments. Assoc Adv Comput Educ 2002:493-498 [FREE Full text]
39. Park H, Kim HS, Park HW. A scientometric study of digital literacy, ICT literacy, information literacy, and media literacy. Journal of Data and Information Science 2021 Apr 4;6(2):116-138. [doi: 10.2478/jdis-2021-0001]
40. Fu J. ICT in education: a critical literature review and its implications. Int J Educ Dev Using Inf Commun Technol 2013;9(1):112-125 [FREE Full text]
41. Rogers SE. Bridging the 21st century digital divide. TechTrends 2016 May;60(3):197-199. [doi: 10.1007/s11528-016-0057-0]
42. Tang X, Sun Y, Zhang B, et al. I never imagined grandma could do so well with technology. Proc ACM Hum-Comput Interact 2022 Nov 7;6(CSCW2):1-29. [doi: 10.1145/3555579]
43. Zhu R, Yu X, Krever R. The double burden: the digital exclusion and identity crisis of elderly patients in rural China. MaC ;12. [doi: 10.17645/mac.8106]
44. Estrela M, Semedo G, Roque F, Ferreira PL, Herdeiro MT. Sociodemographic determinants of digital health literacy: a systematic review and meta-analysis. Int J Med Inform 2023 Sep;177:105124. [doi: 10.1016/j.ijmedinf.2023.105124] [Medline: 37329766]
45. Magen-Nagar N, Cohen L. Learning strategies as a mediator for motivation and a sense of achievement among students who study in MOOCs. Educ Inf Technol 2017 May;22(3):1271-1290. [doi: 10.1007/s10639-016-9492-y]
46. Martin A. DigEuLit–a European framework for digital literacy: a progress report. J eLiteracy 2005;2(2):130-136 [FREE Full text]
47. Demirdis B. Integrating digital literacy to enhance emotional and social skills in education. In: In Innovative Educational Frameworks for Future Skills and Competencies: IGI Global Scientific Publishing; 2025:1-38. [doi: 10.4018/979-8-3693-7555-6.ch001]
48. Bianchi C. Exploring how internet services can enhance elderly well-being. JSM 2021 Sep 1;35(5):585-603. [doi: 10.1108/JSM-05-2020-0177]
49. Safarov N. Personal experiences of digital public services access and use: older migrants' digital choices. Technol Soc 2021 Aug;66:101627. [doi: 10.1016/j.techsoc.2021.101627]
50. Janke S, Bardach L, Oczlon S, Lüftenegger M. Enhancing feasibility when measuring teachers' motivation: a brief scale for teachers' achievement goal orientations. Teaching and Teacher Education 2019 Jul;83:1-11. [doi: 10.1016/j.tate.2019.04.003]
51. National Bureau of Statistics of China. China Statistical Yearbook 2023: Aging and Demographic Trends: China Statistics Press; 2023.
52. Jöreskog KG, Sörbom D. LISREL 8.80. : Scientific Software International Inc; 2006 URL: https://www.ssicentral.com/lisrel [accessed 2025-12-11]
53. Doll WJ, Xia W, Torkzadeh G. A confirmatory factor analysis of the end-user computing satisfaction instrument. MIS Q 1994 Dec;18(4):453. [doi: 10.2307/249524]

54. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996;1(2):130-149. [doi: 10.1037//1082-989X.1.2.130]
55. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull 1980;88(3):588-606. [doi: 10.1037//0033-2909.88.3.588]
56. Deci EL, Ryan RM. The general causality orientations scale: self-determination in personality. J Res Pers 1985 Jun;19(2):109-134. [doi: 10.1016/0092-6566(85)90023-6]
57. Badr J, Motulsky A, Denis JL. Digital health technologies and inequalities: a scoping review of potential impacts and policy recommendations. Health Policy 2024 Aug;146:105122. [doi: 10.1016/j.healthpol.2024.105122] [Medline: 38986333]
58. Farmania A, Elsyah RD, Fortunisa A. The phenomenon of technostress during the COVID-19 pandemic due to work from home in Indonesia. Sustainability ;14(14):8669. [doi: 10.3390/su14148669]
59. Norman CD, Skinner HA. eHEALS: the eHealth literacy scale. J Med Internet Res 2006;8(4):e27. [doi: 10.2196/jmir.8.4.e27]
60. Bentler PM, Chou CP. Practical issues in structural modeling. Sociol Methods Res 1987 Aug;16(1):78-117. [doi: 10.1177/0049124187016001004]
61. Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis, 7th edition: Pearson; 2009.
62. MacCallum RC, Hong S. Power analysis in covariance structure modeling using GFI and AGFI. Multivariate Behav Res 1997 Apr;32(2):193-210. [doi: 10.1207/s15327906mbr3202_5]
63. Hu LT, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
64. Mulaik SA, James LR, Van Alstine J, Bennett N, Lind S, Stilwell CD. Evaluation of goodness-of-fit indices for structural equation models. Psychol Bull 1989;105(3):430-445. [doi: 10.1037/0033-2909.105.3.430]

Abbreviations

ADTS: Assistive Digital Tools and Services
CFA: confirmatory factor analysis
SCT: social cognitive theory
SEM: structural equation modeling
TAM: Technology Acceptance Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by R Bjarnadottir; submitted 31.Mar.2025; peer-reviewed by GHY Wong, S Shapira; revised version received 03.Sep.2025; accepted 20.Oct.2025; published 21.Jan.2026.

Please cite as:
Niu HJ, Li MH, Hsieh FY, Yu CC, Lin CT
From Digital Anxiety to Empowerment in Older Adults: Cross-Sectional Survey Study on Psychosocial Drivers of Digital Literacy
JMIR Aging 2026;9:e75245
URL: https://aging.jmir.org/2026/1/e75245
doi:10.2196/75245

Level of eHealth Literacy and Its Associations With Health Behaviors and Outcomes in Chinese Older Adults: Cross-Sectional Analysis of Baseline Data From a Large-Scale Community Project

Siu Long Chau[1*], PhD; Wanjia He[1*], PhD; Tzu Tsun Luk[2], PhD; Sophia Siu Chee Chan[1*], PhD

[1]School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Metro South Tower 1, 39 Wong Chuk Hang Road, Wong Chuk Hang, Hong Kong, China (Hong Kong)

[2]Alice Lee Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

[*]these authors contributed equally

Corresponding Author:
Sophia Siu Chee Chan, PhD
School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Metro South Tower 1, 39 Wong Chuk Hang Road, Wong Chuk Hang, Hong Kong, China (Hong Kong)

Abstract

Background: eHealth literacy is important for older adults to be able to seek and evaluate online health information. However, there is a scarcity of large-scale data on their eHealth literacy levels, particularly among the oldest older individuals (aged >75 years) in unique, high-income Asian regions such as Hong Kong. A comprehensive understanding of how eHealth literacy is associated with specific health behaviors, mental well-being, and physical health outcomes in this population is lacking.

Objective: This study aims to assess the level of eHealth literacy and its associations with health behaviors and health-related outcomes among older adults in Hong Kong.

Methods: We conducted a cross-sectional analysis of baseline data from the Generations Connect Project. This is an ongoing large-scale community-based project, where we trained university students to conduct home visits and assess the health status of older adults (N=6704) in Hong Kong. eHealth literacy was measured using the eHealth Literacy Scale (eHEALS; score: 8‐40). Health behavior measurements included physical activity levels (metabolic equivalent of task minutes per week) and smoking, drinking, and eating habits. Mental well-being was measured using the World Health Organization-Five Well-Being Index (percentage score: 0‐100) and UCLA 3-item Loneliness Scale (score: 3‐9). Physical health was assessed on the basis of self-reported medical diagnosis of noncommunicable diseases (yes/no), including hypertension, diabetes, cardiovascular disease, and stroke. Adjusted unstandardized coefficients (b) and odds ratios (ORs) were calculated to determine the associations between variables.

Results: Among the 6704 participants (mean age 77.8, SD 7.0 years), the mean eHEALS score was 18.2 (SD 10.2), and 44.1% (2897/6566) of the participants had inadequate eHealth literacy (score: 8‐15.99). Increasing age (adjusted b −0.32, 95% CI −0.35 to −0.28; *P*<.001), support from the Comprehensive Social Security Assistance Scheme (adjusted b −1.49, 95% CI −2.04 to −0.95; *P*<.001), and living in public housing (adjusted b −1.60, 95% CI −2.69 to −0.50; *P*=.004) were associated with a lower eHEALS score. Participants with moderate eHealth literacy (score: 24‐31.99) were less likely to be current smokers (adjusted OR 0.60, 95% CI 0.38-0.95; *P*=.04), more physically active (adjusted b 39.83, 95% CI 2.04-77.62; *P*=.04), more likely to be community health center members (adjusted OR 1.52, 95% CI 1.30-1.77; *P*<.001) and to have healthy diets (adjusted b 0.034, 95% CI 0.006-0.063; *P*=.04), and less likely to have a medical diagnosis of diabetes (adjusted OR 0.73, 95% CI 0.62-0.85; *P*<.001). Moreover, they had a higher score on the World Health Organization-Five Well-being Index (adjusted b 2.89, 95% CI 1.42-4.36; *P*<.001) and a lower score on the UCLA 3-item Loneliness Scale (adjusted b −0.26, 95% CI −0.37 to −0.15; *P*<.001).

Conclusions: The level of eHealth literacy was low among older adults in Hong Kong. eHealth literacy was associated with positive health behaviors and health-related outcomes. Interventions are warranted to boost their eHealth literacy in the future.

(JMIR Aging 2026;9:e74110) doi:10.2196/74110

KEYWORDS

eHealth literacy; older adults; health behaviors; health-related outcomes; healthy aging

Introduction

Older adults (aged >60 years) accounted for 5% of the global population in the last few decades, and this estimate is projected to rise to 16% by 2050 [1]. In China, the population aged ≥80 years reached 36 million in 2020 and will increase to 115 million in 2050 [2]. This global demographic shift toward an older population is accompanied by an increased burden of noncommunicable diseases (NCDs) [3] and mental illnesses [4], which places immense strain on public health care systems. More than 20% of older adults in the world have 1 or more mental illnesses [4], and the prevalence increases with population size. In 2017, 8.6% of older adults in Hong Kong (N=4368) had depression and anxiety [5]. Mental illnesses among older people remained underrecognized and undertreated, which accounted for 10.6% of the total loss of disability-adjusted life years in older adults [6]. In response, and in line with the World Health Organization's call for primary health care reform [7], there is an urgent need for innovative strategies to support healthy aging. The rapid development of digital technologies, accelerated by the COVID-19 pandemic, offers a key opportunity. Digital tools have the potential to support older adults' intrinsic capacity, promote active and healthy aging, and alleviate the economic burden on health care sectors, making them an increasingly recognized priority in geriatric care [8].

The emergence of digital health offers a new avenue for improving individuals' quality of life. Rapid advances in information and communications technology contribute to the increasing innovation and upgrading of health service modes [9]. Digital health enables low-cost, timely health care services for older adults and supports their physical and mental health by providing access to reliable health information, self-monitoring tools for chronic conditions, and social connection platforms [9]. Although accessing online medical services and information is convenient, the quality of online information varies [10]. Older adults need eHealth literacy to navigate and evaluate the accuracy of medical information, especially in the context of primary health care reform in Hong Kong, where the government is committed to enhancing digital primary health care services to shift the emphasis from tertiary to primary care.

eHealth literacy is defined as the ability of an individual to find and evaluate health information from online platforms and apply health-related knowledge to address health problems [11]. Although smartphone ownership has increased among older adults [12], they often possess lower technological skills than the younger generations and are consequently less likely to use these devices to search for health information [12]. eHealth literacy impacts how older adults search for health information online, which might influence health-related decisions [12]. With the rapid and continuous development of eHealth resources, it is crucial to assess older adults' eHealth literacy and examine its association with health-related outcomes.

Previous studies suggested that higher eHealth literacy was associated with some healthy aging components (eg, less cognitive impairment and functional limitation) in older adults [13]. However, database searches in PubMed and CINAHL (using keywords such as "eHealth literacy," "older adults," "aging population," "health behavior," "smoking," and "physical activity") suggest that studies on the association of eHealth literacy with specific health behaviors (eg, smoking, drinking, and physical exercise) and health status are scarce. Only 1 recent meta-analysis showed a positive association between eHealth literacy and general health-related behaviors (overall estimate of the correlation: 0.31, 95% CI: 0.25 - 0.34), with a similar effect size observed in older adult populations [14]. Another systematic review showed that higher eHealth literacy was associated with better health knowledge and attitude in older adults, but the associations with physical and psychosocial outcomes were still inconsistent [15]. The relationships between eHealth literacy and specific health-related outcomes (eg, mental and physical well-being) need to be assessed to further explore their underlying mechanisms. In addition, studies on digital health often recruit participants from more accessible, community-dwelling populations [16], potentially underrepresenting the most vulnerable, such as homebound or socioeconomically disadvantaged older adults, who may face the greatest digital divide.

While existing research provides a foundational understanding of eHealth literacy, several critical gaps remain, limiting the development of effective, evidence-based interventions for older adults. First, eHealth literacy levels are known to vary significantly across different sociocultural and demographic contexts, with studies in the United States [16], South Korea [17], and Norway [18] reporting higher average scores than those reported in mainland China [19]. However, there is a scarcity of large-scale, representative data from Hong Kong, a unique high-income region characterized by extreme longevity, high digital penetration, and a government-led push for digital primary care reform. Second, much of the existing literature focuses on younger older populations (eg, those aged 60-65 years); however, the challenges and capabilities of the oldest older individuals (ie, those aged ≥75 years) may differ substantially. Third, while a study has confirmed a general positive association between eHealth literacy and composite health-related behaviors [14], there is a lack of research examining the associations between eHealth literacy and a comprehensive suite of specific, modifiable behaviors (eg, smoking, physical activity, and diet) and health service use. Finally, outcomes are often studied in silos. A holistic understanding requires a simultaneous investigation of the links between eHealth literacy and a broad spectrum of outcomes, including health behaviors, mental well-being, social connectedness, and physical health status.

To address these gaps, this study aimed to (1) assess the level of eHealth literacy in a large, community-based sample of underprivileged older adults in Hong Kong, with a particular focus on the oldest older population and (2) comprehensively examine the associations of eHealth literacy with specific health behaviors, mental health outcomes (well-being and loneliness), and physical health outcomes (prevalence of multiple NCDs).

Methods

Study Design

This study used a cross-sectional design, using baseline data from the Generations Connect Project. The parent project is an ongoing, large-scale, community-based, quasi-experimental pre-post study that trained over 1000 nursing students to provide home visits and perform a baseline health assessment and to deliver a health intervention (face-to-face and digital) to older adults in Hong Kong. The present study focused on the analysis of the baseline data collected before any intervention took place.

Older adults were recruited through collaborations with 20 local nongovernmental organizations covering 18 districts across Hong Kong. Between November 2022 and December 2024, we recruited participants aged ≥65 years who could read and communicate in Chinese and were cognitively and physically capable of understanding and answering survey questions. Older adults who were bedbound and had a history of mental illnesses were excluded. The exclusion of older adults with self-reported mental illness was to minimize potential confounding effects on the mental health outcomes measured in this study. In contrast, older adults with physical limitations (eg, visual or hearing impairments) were not excluded, as these conditions are prevalent in the target population, and their inclusion is crucial for the generalizability of our findings regarding the digital divide among the oldest older population. We recruited University of Hong Kong students to be student ambassadors through open recruitment via internal university emails and on-campus promotional materials (eg, posters and flyers) as well as selected curriculum (medical, nursing, dental, and common core courses) integration. Before making a home visit, all student ambassadors received a 3-hour training to ensure the intervention's fidelity and data quality during the data collection process. Details of the training session and intervention are reported elsewhere [20].

Ethical Considerations

Ethical approval was sought from the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (IRB/REC reference: UW 22 - 693). The institutional review board also allowed secondary analysis without additional consent. The reporting of this study strictly followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines (Checklist 1). The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Informed consent was obtained using comprehensive information sheets and verbal explanations, and signed consent forms were obtained from participants before any procedures began. The consent form included information about the study's purpose and procedures and the participant's right to withdraw at any time without penalty.

Participants enrolled in the study received gift bags worth HK $50 (US $6.44) as an incentive during the initial student visits.

All data collected during the study were deidentified to protect the participants' privacy, with all identifying information removed. Access to the data was restricted to the research team, and all data were stored securely.

Measurements

We collected and analyzed the baseline data on the sociodemographic characteristics (eg, sex, age, highest educational attainment, marital status, and family monthly income), health behaviors, mental health, physical health, and eHealth literacy of participants.

Health behaviors included physical activity levels in metabolic equivalent of task minutes (MET min) per week (measured using the International Physical Activity Questionnaire Short Form) [21], number of days of vegetable and fruit consumption in the past week, smoking status (never smoker, ex-smoker, and current smoker), intention to quit within 30 days (yes/no) for current smokers, and drinking frequency in the past year (never, monthly or less, 2-4 times per month, 2-3 times per week, and ≥4 times per week).

Mental health and perceived level of loneliness were assessed using the World Health Organization-Five (WHO-5) Well-Being Index [22] and UCLA 3-item Loneliness Scale, respectively [23]. The raw score on the WHO-5 ranges from 0 to 25, and the percentage score ranges from 0 to 100 (raw score multiplied by 4). A percentage score of 0 represents the worst possible mental health status, and 100 represents the best possible mental health status [22]. The WHO-5 has been validated for the Chinese population [24]. In our sample, the scale demonstrated good internal consistency, with a Cronbach α of 0.89. The UCLA 3-item Loneliness Scale measures 3 dimensions of loneliness (relational and social connectedness and self-perceived isolation) with a score ranging from 3 to 9, where a higher score indicates a higher level of loneliness [23]. The UCLA 3-item Loneliness Scale was validated for older Chinese adults [25]. The scale showed high reliability in the present study, with a Cronbach α of 0.88. Physical health was assessed by asking whether the participants had been medically diagnosed with each NCD (yes/no), including hypertension, diabetes, cardiovascular disease, and stroke. A multimorbidity variable was constructed (defined as having been medically diagnosed with ≥2 of the following 4 conditions: hypertension, diabetes, cardiovascular disease, and stroke).

eHealth literacy was measured using the eHealth Literacy Scale (eHEALS) [11]. The scale measures participants' perceived skills, knowledge, and comfort toward eHealth [11]. The scale consists of 8 items, which are rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The total score ranges from 8 to 40, with higher scores indicating higher self-perceived eHealth literacy [11]. On the basis of the score, the respondents can be categorized into 4 groups: inadequate eHealth literacy (score: 8 - 15.99), low eHealth literacy (score: 16 - 23.99), moderate eHealth literacy (score: 24 - 31.99), and high eHealth literacy (score: 32 - 40) [26]. The scale has high internal consistency, with a Cronbach α of 0.96 [26], and has been validated for Chinese older adults, with a Cronbach α of 0.88 [26].

Statistical Analysis

Inverse probability weighting based on the sex distribution of Hong Kong older adults in 2023 (from the census) was conducted to make the sample more representative of Hong Kong's older adult population. Descriptive statistics were used to describe the participants' sociodemographic characteristics, eHealth literacy, health behaviors, mental health, and physical health at baseline.

The associations of sociodemographic characteristics (sex, age, highest educational attainment, family monthly income, Comprehensive Social Security Assistance [CSSA] Scheme status, type of housing, and living with family members) with the eHEALS score were analyzed using multivariable linear regression (adjusted unstandardized coefficient [b]) with mutual adjustment for baseline sociodemographic characteristics.

Multiple logistic regression (adjusted odds ratio [AOR]) was used to analyze the associations of the level of eHealth literacy with smoking and drinking status (yes/no), District Health Center member status (yes/no), multimorbidity of NCDs (yes/no), and medical diagnosis of NCDs (yes/no). Multivariable linear regression (adjusted unstandardized coefficient [b]) was used to examine the associations of the level of eHealth literacy with total physical activity (MET min/week), the WHO-5 Well-Being Index score, and the UCLA 3-item Loneliness Scale score. The associations of the level of eHealth literacy with the number of days of vegetable and fruit consumption in the past week were analyzed using Poisson regression. The models were adjusted for baseline sociodemographic characteristics. All analyses were conducted in Stata (version 15.1; StataCorp LLC). A 2-sided *P* value <.05 was considered statistically significant.

Results

Between November 2022 and December 2024, a total of 7087 potential participants were screened and found eligible, and 6704 agreed to participate and were included in the baseline analysis. The mean time to conduct baseline assessments and deliver interventions was 102 (SD 7.3) minutes.

Table 1 shows that the mean age of the participants was 77.8 (SD 7.0) years, and 26.9% (1805/6704) of the participants were men. Most participants had primary education or below (4003/6676, 60.0%), were married (3173/6626, 47.9%), and had a family monthly income of HK $25,000 (US $3205.9) or less (5127/6653, 77.1%). Moreover, 25.4% (1691/6650) of participants were supported by the CSSA Scheme. Most participants were retired (6460/6694, 96.5%), living in public housing (4829/6696, 72.1%), and living with family members (3479/6696, 52.0%). In total, 71.2% (4774/6704) of participants were not District Health Center members. The mean eHEALS score of participants was 18.2 (SD 10.2), and most participants (2897/6566, 44.1%) had inadequate eHealth literacy (score: 8 - 15.99).

Table . Sociodemographic characteristics of participants (N=6704)[a].

Characteristic	Participants, n (%)
Sex	
Male	1805 (26.9)
Female	4894 (73.1)
Age[b] (years)	77.8 (7.0)
Highest educational attainment	
Primary education or below	4003 (60)
Secondary education	2288 (34.2)
Tertiary education	385 (5.8)
Marital status	
Single	397 (6)
Married	3173 (47.9)
Divorced	454 (6.8)
Widowed	2602 (39.3)
Family monthly income	
None	204 (3.1)
HK $25,000 (US $ 3205.9) or less	5127 (77.1)
More than HK $25,000 (US $ 3205.9)	1322 (19.9)
Support from the Comprehensive Social Security Assistance Scheme	
No	4959 (74.6)
Yes	1691 (25.4)
Occupational status	
Full-time (employed)	35 (0.5)
Part-time (employed)	90 (1.3)
Unemployed	13 (0.2)
Retired	6460 (96.5)
Caregiver/housewife	96 (1.4)
Type of housing	
Public	4829 (72.1)
Private	1546 (23.1)
Other	321 (4.8)
Living with family members	
No	3046 (45.5)
Yes	3479 (52)
Other	171 (2.5)
District Health Center member status	
No	4774 (71.2)
Yes	1930 (28.8)
eHealth literacy (eHEALS score[b]: 8 - 40)[c]	18.2 (10.2)
Inadequate (score: 8 - 15.99)	2897 (44.1)
Low (score:16 - 23.99)	1322 (20.1)
Moderate (score: 24 - 31.99)	1182 (18)

Characteristic	Participants, n (%)
High (score: 32 - 40)	1165 (17.7)

[a]The proportions were weighted by sex distribution of older adults in Hong Kong 2023. The observations (n) were unweighted.

[b]Reported as mean (SD).

[c]eHealth literacy is measured on a 5-point eHealth Literacy Scale (eHEALS), with ratings ranging from 1 (strongly disagree) to 5 (strongly agree). The overall score ranges from 8 to 40, with a higher score indicating more perceived skills in finding, evaluating, and using electronic information to make health decisions.

Table 2 shows that the participants' mean total physical activity was 628.7 (SD 542.9) MET min per week. The mean number of days on which vegetables and fruits were consumed in the past week was 6.6 (SD 1.3) and 5.9 (SD 2.0), respectively. In total, 3% (198/6686) of participants were current smokers, and 91% (163/179) had no intention to quit. Moreover, 8.5% (570/6704) of participants were current drinkers. The mean percentage score on the WHO-5 Well-being Index was 68.1 (SD 21.3), and the mean score on the UCLA 3-item Loneliness Scale was 4.0 (SD 1.6). Notably, 45.2% (3020/6674) of participants had multimorbidity of NCDs (medically diagnosed with ≥2 NCDs). Of the 6704 participants, 60.5% (n=4059) were medically diagnosed with hypertension, 27.1% (n=1816) had diabetes, 13.3% (n=894) had cardiovascular disease, and 4.6% (n=309) had a medical diagnosis of stroke.

Table . Health behaviors, mental health, and physical health of participants (N=6704)[a].

Parameter	Outcome
Physical activity, IPAQ-SF continuous score (MET[b] min/week), mean (SD)[c]	
Walking	463.5 (443.8)
Moderate	139.9 (222.6)
Vigorous	25.2 (108.6)
Total	628.7 (542.9)
Number of days of vegetable consumption in the past week, mean (SD)	6.6 (1.3)
Number of days of fruit consumption in the past week, mean (SD)	5.9 (2.0)
Smoking status, n (%)	
Never smoker	5812 (86.9)
Ex-smoker	676 (10.1)
Current smoker	198 (3)
Intention to quit for current smokers, n (%)	
No	163 (91)
Yes	16 (9)
Drinking frequency in the past year, n (%)	
Never	6104 (91.5)
Monthly or less	289 (4.3)
2 - 4 times per month	129 (1.9)
2 - 3 times per week	46 (0.7)
>4 times per week	106 (1.6)
WHO-5[d] Well-Being Index percentage score (0 - 100)[e], mean (SD)	68.1 (21.3)
UCLA 3-item Loneliness Scale score: 3 - 9)[f], mean (SD)	4.0 (1.6)
Multimorbidity[g] of NCDs[h], n (%)	
No	3654 (54.7)
Yes	3020 (45.2)
Medical diagnosis of hypertension, n (%)	
No	2645 (39.5)
Yes	4059 (60.5)
Medical diagnosis of diabetes, n (%)	
No	4888 (72.9)
Yes	1816 (27.1)
Medical diagnosis of cardiovascular disease, n (%)	
No	5810 (86.7)
Yes	894 (13.3)
Medical diagnosis of stroke, n (%)	
No	6395 (95.4)
Yes	309 (4.6)

[a]The proportions were weighted by sex distribution of older adults in Hong Kong 2023. The observations (n) were unweighted.

[b]MET: metabolic equivalent of task.

[c]The International Physical Activity Questionnaire Short Form (IPAQ-SF) continuous scores are expressed in MET min/week: MET level × minutes of activity × events per week. The total physical activity is computed as the sum of walking, moderate, and vigorous scores (MET min/week).

[d]WHO-5: World Health Organization-Five Well-Being Index.

[e]The raw score, ranging from 0 to 25, is multiplied by 4 to obtain a percentage score. A percentage score of 0 represents the worst possible quality of life, whereas a score of 100 represents the best possible quality of life.

[f]The 3-point response scale for each item ranges from “hardly ever or never” (1 point) to “often” (3 points), and the total score is the sum of all items, which ranges from 3 to 9, with higher scores indicating a higher level of perceived loneliness.

[g]Multimorbidity was defined as having been medically diagnosed with ≥2 of the following 4 conditions: hypertension, diabetes, cardiovascular disease, and stroke.

[h]NCD: noncommunicable disease.

Table 3 shows that increasing age (adjusted b −0.32, 95% CI −0.35 to −0.28; *P*<.001), support from the CSSA Scheme (adjusted b −1.49, 95% CI −2.04 to −0.95; *P*<.001), and living in public housing (adjusted b −1.60, 95% CI −2.69 to −0.50; *P*=.004) were associated with a lower eHEALS score, after mutual adjustment for baseline sociodemographic characteristics. Secondary education (adjusted b 4.31, 95% CI 3.81-4.82; *P*<.001) and tertiary education (adjusted b 9.04, 95% CI 8.04-10.06; *P*<.001) were associated with higher eHEALS scores.

Table . Associations of sociodemographic characteristics with eHealth literacy of participants (N=6704).

Characteristic	Association with eHealth Literacy Scale score (8 - 40)			
	Crude b (95% CI)	*P* value	Adjusted b (95% CI)[a]	*P* value
Sex (reference: male)				
Female	0.05 (−0.50 to 0.61)	.85	0.17 (−0.38 to 0.71)	.89
Age	−0.40 (−0.43 to −0.36)	<.001	−0.32 (−0.35 to −0.28)	<.001
Highest educational attainment (reference: primary education or below)				
Secondary education	5.37 (4.86 to 5.87)	<.001	4.31 (3.81 to 4.82)	<.001
Tertiary education	9.05 (8.02 to 10.08)	<.001	9.04 (8.04 to 10.06)	<.001
Family monthly income (reference: none)				
HK $25,000 (US $3205.9) or less	0.50 (−0.12 to 1.12)	.12	0.24 (−0.34 to 0.82)	.16
More than HK $25,000 (US $3205.9)	2.27 (0.81 to 3.73)	.002	0.91 (−0.46 to 2.29)	.06
Support from the Comprehensive Social Security Assistance Scheme (reference: no)				
Yes	−2.39 (−2.96 to −1.82)	<.001	−1.49 (−2.04 to −0.95)	<.001
Type of housing (reference: other)				
Public	−2.77 (−3.93 to −1.62)	<.001	−1.60 (−2.69 to −0.50)	.004
Private	0.24 (−0.98 to 1.47)	.70	0.13 (−0.73 to 1.41)	.69
Living with family members (reference: other)				
No	0.06 (−1.72 to 1.84)	.94	−0.79 (−2.46 to 0.87)	.87
Yes	0.99 (−0.78 to 2.76)	.95	−0.43 (−2.10 to 1.23)	.89

[a]The model was mutually adjusted for all sociodemographic characteristics listed in the table, including sex, age, highest educational attainment, family monthly income, Comprehensive Social Security Assistance status, type of housing, and living with family members.

Descriptive analysis showed that 93% (6227/6704) of participants were smartphone users. In total, 57.5%(3549/6173) used a smartphone for 1 hour or more daily, 91.5% (6135/6704) had instant messaging apps installed on the smartphone, and 60.2% (3992/6634) had not searched for health-related information from online sources using the smartphone (Multimedia Appendix 1). Table 4 shows that participants with moderate eHealth literacy (score: 24 - 31.99) were less likely to be current smokers (AOR 0.60, 95% CI 0.38-0.95; *P*=.04) and more likely to be District Health Center members (AOR 1.52, 95% CI 1.30-1.77; *P*<.001), and moderate eHealth literacy was associated with higher total physical activity levels (adjusted b 39.83, 95% CI 2.04-77.62; *P*=.04) and a higher number of days of fruit consumption in the past week (adjusted b 0.034, 95% CI 0.006-0.063; *P*=.04). Participants with high eHealth literacy (score 32‐40) were less likely to be current smokers (AOR 0.57, 95% CI 0.36-0.92; *P*=.03) and more likely to be District Health Center members (AOR 1.76, 95% CI 1.51-2.05; *P*<.001), and high eHealth literacy was associated with a higher number of days of fruit consumption in the past week (adjusted b 0.046, 95% CI 0.017-0.075; *P*=.003).

Table . Associations of eHealth literacy with health behaviors of participants (N=6704).

Health behavior	eHealth literacy (eHEALS[a] score: 8‐40) (reference: inadequate eHealth literacy [score: 8-15.99], n=2897)					
	Low (score:16‐23.99) n=1322	*P* value	Moderate (score: 24‐31.99) n=1182	*P* value	High (score: 32‐40) n=1165	*P* value
Current smoker						
Crude OR (95% CI)	0.94 (0.65 to 1.35)	.73	0.69 (0.45 to 1.06)	.09	0.66 (0.42 to 1.02)	.06
Adjusted OR (95% CI)[b]	0.90 (0.61 to 1.34)	.67	0.60 (0.38 to 0.95)	.04	0.57 (0.36 to 0.92)	.03
Current drinker						
Crude OR (95% CI)	1.23 (0.98 to 1.55)	.07	1.33 (1.06 to 1.68)	.02	1.01 (0.79 to 1.30)	.91
AOR (95% CI)[b]	1.16 (0.92 to 1.48)	.10	1.19 (0.93 to 1.53)	.06	0.85 (0.65 to 1.11)	.93
Physical activity (total MET min/week)						
Crude b (95% CI)	5.85 (−29.77 to 41.47)	.75	74.47 (37.54 to 111.39)	<.001	80.12 (42.96 to 117.28)	<.001
Adjusted b (95% CI)[b]	−10.07 (−45.96 to 25.82)	.80	39.83 (2.04 to 77.62)	.04	36.51 (−2.16 to 75.18)	.06
District Health Center member status						
Crude OR (95% CI)	1.25 (1.07 to 1.45)	.003	1.87 (1.61 to 2.16)	<.001	2.27 (1.97 to 2.63)	<.001
AOR (95% CI)[c]	1.13 (0.97 to 1.32)	.06	1.52 (1.30 to 1.77)	<.001	1.76 (1.51 to 2.05)	<.001
Number of days of vegetable consumption in the past week[d]						
Crude b (95% CI)	0.01 (−0.02 to 0.03)	.68	0.01 (−0.01 to 0.04)	.72	0.02 (−0.01 to 0.05)	.44
Adjusted b (95% CI)[b]	0.01 (−0.02 to 0.03)	.70	0.02 (−0.01 to 0.04)	.83	0.02 (−0.01 to 0.05)	.45
Number of days of fruit consumption in the past week[d]						
Crude b (95% CI)	−0.003 (−0.030 to 0.024)	.56	0.034 (0.006 to 0.062)	.04	0.047 (0.019 to 0.074)	.003
Adjusted b (95% CI)[b]	−0.006 (−0.033 to 0.022)	.57	0.034 (0.006 to 0.063)	.04	0.046 (0.017 to 0.075)	.003

[a]eHEALS: eHealth Literacy Scale.

[b]Adjusted for sex, age, highest educational attainment, family monthly income, living with family members, and being a District Health Center member.

[c]Adjusted for sex, age, highest educational attainment, family monthly income, and living with family members.

[d]Poisson regression was used to calculate the unstandardized coefficient.

Table 5 shows that participants with moderate eHealth literacy (score 24 - 31.99) were less likely to have a medical diagnosis of diabetes (AOR 0.73, 95% CI 0.62-0.85; *P*<.001) and stroke (AOR 0.58, 95% CI 0.40-0.85; *P*=.004), and moderate eHealth literacy was associated with a higher score on the WHO-5 Well-being Index (adjusted b 2.89, 95% CI 1.42-4.36; *P*<.001) and a lower score on the UCLA 3-item Loneliness Scale (adjusted b −0.26, 95% CI −0.37 to −0.15; *P*<.001). Participants with high eHealth literacy (score: 32 - 40) were less likely to have a medical diagnosis of hypertension (AOR 0.77, 95% CI 0.67-0.90; *P*=.004) and diabetes (AOR 0.69, 95% CI 0.58-0.81; *P*<.001), and high eHealth literacy was associated with a higher score on the WHO-5 Well-being Index (adjusted b 5.37, 95% CI 3.86-6.87; *P*<.001) and lower score on the UCLA 3-item Loneliness Scale (adjusted b −0.34, 95% CI −0.45 to −0.23; *P*<.001).

Table . Associations of eHealth literacy with the physical health and mental health of participants (N=6704).

Parameter	eHealth literacy (eHEALS[a] score: 8 - 40) (reference: inadequate eHealth literacy [score: 8-15.99], n=2897)					
	Low (score:16 - 23.99) n=1322	*P* value	Moderate (score: 24 - 31.99) n=1182	*P* value	High (score: 32 - 40) n=1165	*P* value
Medical diagnosis of hypertension						
Crude OR[b] (95% CI)	1.01 (0.88 to 1.15)	.92	0.77 (0.67 to 0.89)	<.001	0.65 (0.56 to 0.74)	<.001
Adjusted OR (95% CI)[c]	1.06 (0.93 to 1.22)	.95	0.89 (0.77 to 1.02)	.06	0.77 (0.67 to 0.90)	.004
Medical diagnosis of diabetes						
Crude OR (95% CI)	0.84 (0.73 to 0.97)	.02	0.73 (0.63 to 0.85)	<.001	0.69 (0.59 to 0.80)	<.001
Adjusted OR (95% CI)[c]	0.83 (0.71 to 0.96)	.03	0.73 (0.62 to 0.85)	<.001	0.69 (0.58, 0.81)	<.001
Medical diagnosis of cardiovascular disease						
Crude OR (95% CI)	0.99 (0.83 to 1.20)	.98	0.81 (0.66 to 1.00)	.06	0.93 (0.76 to 1.14)	.49
Adjusted OR (95% CI)[c]	1.10 (0.91 to 1.34)	.98	0.97 (0.79 to 1.21)	.07	1.12 (0.91 to 1.40)	.52
Medical diagnosis of stroke						
Crude OR (95% CI)	0.76 (0.55 to 1.04)	.08	0.58 (0.40 to 0.83)	.003	0.77 (0.55 to 1.06)	.11
Adjusted OR (95% CI)[c]	0.74 (0.54 to 1.03)	.08	0.58 (0.40 to 0.85)	.004	0.80 (0.56 to 1.12)	.13
Multimorbidity[f] of NCDs[g]						
Crude b (95% CI)	0.83 (0.57 to 1.09)	.12	0.67 (0.44 to 1.07)	.10	0.65 (0.50 to 1.12)	.14
Adjusted b (95% CI)[c]	0.86 (0.62 to 1.12)	.14	0.69 (0.48 to 1.10)	.10	0.66 (0.52 to 1.14)	.14
WHO-5 Well-Being Index score[d]						
Crude b (95% CI)	0.96 (−0.42 to 2.34)	.17	2.75 (1.32 to 4.19)	<.001	5.65 (4.21 to 7.09)	<.001
Adjusted b (95% CI)[c]	0.87 (−0.52 to 2.27)	.15	2.89 (1.42 to 4.36)	<.001	5.37 (3.86 to 6.87)	<.001
UCLA 3-item Loneliness Scale score[e]						
Crude b (95% CI)	−0.19 (−0.29 to −0.08)	.003	−0.27 (−0.38 to −0.15)	<.001	−0.40 (−0.52 to −0.29)	<.001
Adjusted b (95% CI)[c]	−0.17 (−0.28 to −0.07)	.003	−0.26 (−0.37 to −0.15)	<.001	−0.34 (−0.45 to −0.23)	<.001

[a]eHEALS: eHealth Literacy Scale.

[b]OR: odds ratio.

[c]Adjusted for sex, age, highest educational attainment, family monthly income, living with family members, and District Health Center member.

[d]WHO-5: World Health Organization-Five Well-Being Index. The score ranges from 0 to 25, with 0 representing the worst possible and 25 representing the best possible quality of life. The score is multiplied by 4. A percentage score of 0 represents the worst possible quality of life, whereas a score of 100 represents the best possible quality of life.

[e]The 3-point response scale for each item ranges from “hardly ever or never” (1 point) to “often” (3 points), and the total score is the sum of all items, which ranges from 3 to 9, with higher scores indicating a higher level of perceived loneliness.

[f]Multimorbidity was defined as having been medically diagnosed with ≥2 of the following 4 conditions: hypertension, diabetes, cardiovascular disease, and stroke.

[g]NCD: noncommunicable disease.

Discussion

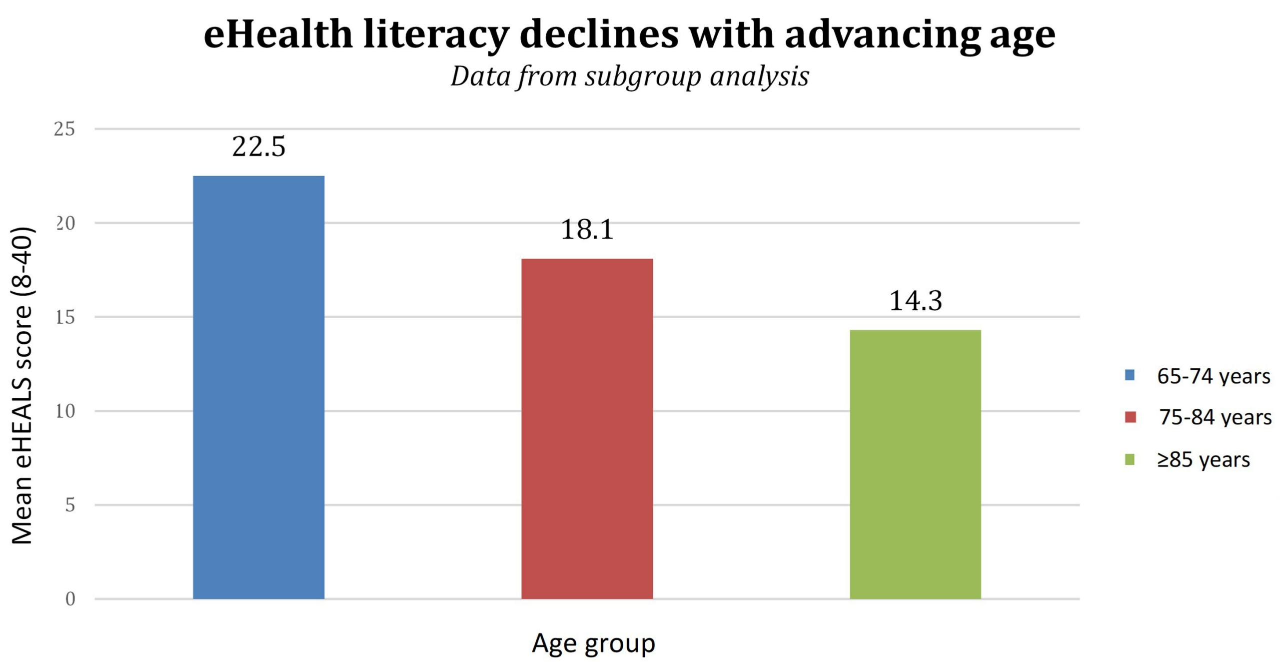
Principal Findings

We analyzed the baseline data from an ongoing, large-scale, quasi-experimental, pre-post study to investigate the level of eHealth literacy and its associations with health behaviors, primary care service use, mental health, and physical health among Hong Kong Chinese older adults (N=6704). The results suggested that their eHealth literacy level was generally low. In addition, our findings showed that older adults with higher eHealth literacy were associated with positive health behaviors and better health-related outcomes at baseline. A key strength of this study is its examination of specific, modifiable health behaviors (eg, smoking, physical activity, and diet) and health outcomes (eg, mental health and NCD risks). This granular approach provides more actionable insights for public health interventions. For example, identifying a link between eHealth literacy and smoking status allows for more targeted health promotion campaigns. This specificity moves the field beyond broad associations toward evidence that can directly inform the content of programs that leverage digital tools for health improvement in aging populations.

The mean eHealth literacy score of the study participants was 18.2 (SD 10.2), which was lower than the scores reported in the United States (30.9) [16], South Korea (30.5) [17], Norway (25.7) [18], and China (21.4) [19]. The notably higher age (mean 77.8, SD 7.0 years) of our sample is likely a key contributor to this low average score, a conclusion supported by our finding that increasing age was the most significant demographic predictor of lower eHealth literacy. This suggests that interventions to enhance older adults’ eHealth literacy are much needed. Our study indicated that 44.1% of older adults (mean age 77.8, SD 7.0 years) lacked eHealth literacy despite having a smartphone with internet access. This can be attributed to older adults’ attitudes toward the internet, including self-efficacy in managing digital technology and a preference for in-person interaction with health practitioners [27]. eHealth literacy may also be lower due to age-related problems and cognitive decline [27]. These findings are relevant to delivering eHealth resources to older adults [27]. Overcoming these sociodemographic and psychological barriers may increase eHealth literacy and help older adults improve their general health. In addition, despite 93% (6227/6697) of participants owning a smartphone, 60.2% (3992/6634) reported never having used it to search for health information (Multimedia Appendix 1). This highlights a critical gap between device ownership and meaningful health-related use, underscoring that access to technology alone is insufficient. This “know-do” gap underscores the urgent need for interventions that build skills and confidence in searching for health information with technological devices among older adults.

Furthermore, we performed subgroup analyses of eHealth literacy by age group (65 - 74 years, 75 - 84 years, and ≥85 years) in our study sample (Figure 1). The results showed a stepwise decline in eHealth literacy across age strata. Older adults with increasing age had the highest burden of chronic disease and the greatest need for health care services, yet they are the most likely to be excluded from an increasingly digitized health care system. Our study’s findings can serve as a data point for policymakers and health care providers, highlighting the urgent need for tailored, age-appropriate support systems to prevent the digital divide from becoming a health equity crisis.

Figure 1. Mean eHealth literacy Scale (eHEALS) score by age group in the study sample (N=6704).



Our findings showed that older adults with higher education were associated with higher eHealth literacy, and older adults with indicators of lower socioeconomic status (receiving support from the CSSA Scheme and living in public housing) were associated with lower eHealth literacy. Lower-income older adults may have poorer access to digital technology, such as limited computer and internet access, leading to skill disparity [28]. The relationship between socioeconomic status and eHealth literacy needs to be investigated in depth to prevent the potential loss of the silver economy in a digitalized society. It is important to consider the social disparities and the digital divide factors when promoting eHealth services to ensure that those needing these services are not left out [28]. Primary health care is the cornerstone of the health care system, and we should optimize the primary health care services to enhance older adults' eHealth literacy. In addition, training should be provided to carers and the younger generation (eg, university students) to guide underprivileged older adults in using electronic information technology and meeting their health needs in the digital age.

We found that older adults with moderate and high eHealth literacy were less likely to be smokers and more likely to engage in positive health behaviors. Higher eHealth literacy may enhance an individual's self-efficacy in managing their health. With an enhanced ability to find, evaluate, and apply online health information, older adults may feel more confident in making informed decisions [29], leading to better health behaviors such as improved diet or increased physical activity and more effective self-management of chronic conditions. In addition, greater access to eHealth resources increases the awareness of health maintenance and encourages older adults to adopt healthy lifestyles [30]. Older adults skilled in using digital health services will be able to communicate with health care professionals in advance and manage their health more efficiently, enabling them to receive more comprehensive NCD prevention services.

Our findings revealed a strong association between higher eHealth literacy and better mental health outcomes, specifically higher levels of well-being and lower levels of loneliness. This finding aligns with a growing body of literature demonstrating the positive role of digital engagement in the psychosocial health of older adults [31]. For instance, a study in Hong Kong during the COVID-19 pandemic found that the use of instant messaging apps was a key factor in mitigating loneliness [32]. Our results expand on this by suggesting that eHealth literacy, as a core component of digital competency, is a crucial enabler of such beneficial social connections. The mechanism is likely 2-fold. First, digitally literate older adults can more easily use technology to maintain contact with family and friends, thereby strengthening their social support networks. Second, those with low eHealth literacy may struggle to differentiate credible information from misinformation online [33], leading to increased anxiety and negative emotional states. The implication for practice is significant: interventions aimed at improving mental health in older adults should consider incorporating digital literacy training not merely as a technical skill but as a fundamental tool for fostering social connectedness and resilience in a digital world.

Implications and Future Directions

Based on the findings identified in this cross-sectional study, several avenues for future research are warranted to deepen our understanding and inform effective interventions. First, there is a critical need for longitudinal studies to establish temporal relationships between eHealth literacy and health outcomes. For instance, a recent 3-wave longitudinal study demonstrated that higher baseline eHealth literacy predicted a better health-promoting lifestyle over time among Chinese older adults [34]. Future research should build on this finding by following up older adults to determine whether improving eHealth literacy leads to sustained health behavior change and better long-term health outcomes and to explore the mediating pathways (eg, self-efficacy) in these relationships.

Second, to move beyond association and test for causality, future work should include intervention studies, such as randomized controlled trials. These studies could evaluate the effectiveness of tailored eHealth literacy training programs on specific outcomes, including health behavior change, chronic disease self-management, and mental well-being in older adults. Such interventions should be multifaceted and co-designed with older adults, moving beyond basic technology access to build practical skills and confidence. Strategies could include (1) tailored, hands-on training workshops; (2) intergenerational peer-tutoring models; (3) the development of age-friendly user interfaces for health apps; and (4) the integration of eHealth literacy support within primary care settings and community centers.

Third, the measurement of eHealth literacy itself requires advancement. This study, similar to many others, relied on a self-report scale. Future research would benefit from incorporating objective, performance-based assessments. This would provide a more accurate measure of actual competency and help bridge the gap between perceived and demonstrated skills. Additionally, research is needed to validate eHealth literacy assessment tools and their cut-off scores for the older and oldest older populations.

Finally, to enhance generalizability, future research should include older adults with a history of mental illness. While they were excluded from our analysis to avoid confounding the mental well-being outcomes, research focus on this vulnerable group is essential for developing tailored interventions that address their specific barriers to digital engagement and health management [35], thus further narrowing the digital divide.

Limitations

This study has several limitations. First, the findings were based on an analysis of the baseline data, and causal relationships between variables could not be confirmed. Given the initial observational data, intervention studies (eg, randomized controlled trials) should be conducted to determine causal relationships between eHealth literacy, health behaviors, and health-related outcomes. Follow-up data and qualitative research may provide more insights into these mechanisms in the future. Second, our study excluded older adults with a self-reported history of mental illness (eg, depression and anxiety), and this exclusion may limit the generalizability of our findings. Mental health status may be associated with an individual's motivation and ability to engage with digital technology [31]. Further research is needed to investigate the relationships between eHealth literacy and health outcomes, specifically within this vulnerable subgroup of older adults. Third, individual eHealth literacy and health status are closely related to the local economic development level [36], and the study findings may not be generalizable to regions with different socioeconomic contexts, cultural backgrounds, or levels of digital infrastructure. In addition, the study was conducted in Hong Kong, a high-income region with advanced digital infrastructure. The observed levels of eHealth literacy and their associations with health outcomes may not be directly generalizable to older adults in low-income or middle-income settings, where access to technology and digital skills may be substantially different. Fourth, this study used the eHEALS to assess eHealth literacy. While widely used and validated, this scale has significant limitations. Developed in 2006 [11], eHEALS may not fully capture the complex skills required to navigate today's ecosystem of mobile apps and interactive health technologies [37]. Furthermore, newer instruments such as the Digital Health Literacy Instrument (DHLI) have been developed and show good utility in Chinese older adults [38]. Recent systematic reviews also highlight the need for more robust, performance-based tools, as eHEALS assesses self-perceived skills rather than actual competence, which may be subject to self-report bias [37,39]. Finally, the generalizability of our findings to the entire older adult population in Hong Kong may be limited. While our sample was large and weighted by sex distribution to match the Hong Kong census, the mean age of our participants was 77.8 (SD 7.0) years, which is higher than that of the general older adult population in Hong Kong [40].

Conclusions

eHealth literacy is an essential skill in a rapidly digitalizing world. The findings showed that the eHealth literacy of older adults in Hong Kong was low and needs improvement, especially in the context of global primary health care reform. We also observed that higher levels of eHealth literacy were associated with health-promoting behaviors, primary care service use, and better physical and mental health outcomes. With the application of the results from this study, tailored interventions should be implemented to improve eHealth literacy and narrow the digital divide among older adults.

Acknowledgments

We would like to thank the participants who enrolled in this study.

Funding

This research was supported by the Sino Group, Ng Teng Fong Charitable Foundation, China.

Data Availability

Data from this study will be made available upon reasonable request. For data access, please contact the corresponding author, Sophia Siu Chee CHAN (University of Hong Kong, Hong Kong).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Characteristics of smartphone usage of participants (N=6704).
[DOCX File, 22 KB - aging_v9i1e74110_app1.docx]

Checklist 1
STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.
[PDF File, 91 KB - aging_v9i1e74110_app2.pdf]

References

1. Kassebaum NJ. Global, regional, and national burden of diseases and injuries for adults 70 years and older: systematic analysis for the Global Burden of Disease 2019 Study. BMJ 2022;376:e068208. [doi: 10.1136/bmj-2021-068208]
2. Lv J, Zhang ZF. Progress and challenges in NCD prevention and control in China. BMJ 2024 Oct 18;387:q2098. [doi: 10.1136/bmj.q2098] [Medline: 39424318]
3. Bai J, Cui J, Shi F, Yu C. Global epidemiological patterns in the burden of main non-communicable diseases, 1990-2019: relationships with socio-demographic index. Int J Public Health 2023;68:1605502. [doi: 10.3389/ijph.2023.1605502] [Medline: 36726528]
4. Petrova NN, Khvostikova DA. Prevalence, structure, and risk factors for mental disorders in older adults [Article in Russian]. Adv Gerontol 2021;34(1):152-159. [doi: 10.1134/S2079057021040093] [Medline: 33993676]
5. Fung AWT, Chan WC, Wong CSM, et al. Prevalence of anxiety disorders in community dwelling older adults in Hong Kong. Int Psychogeriatr 2017 Feb;29(2):259-267. [doi: 10.1017/S1041610216001617] [Medline: 27766997]
6. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3]
7. Global momentum on primary health care: time to unite. World Health Organization. 2025 Jun 4. URL: https://www.who.int/news/item/04-06-2025-global-momentum-on-primary-health-care-time-to-unite [accessed 2025-10-07]
8. Stara V, Soraci L, Takano E, et al. Intrinsic capacity and active and healthy aging domains supported by personalized digital coaching: survey study among geriatricians in Europe and Japan on eHealth opportunities for older adults. J Med Internet Res 2023 Oct 12;25:e41035. [doi: 10.2196/41035] [Medline: 37824183]
9. Junaid SB, Imam AA, Balogun AO, et al. Recent advancements in emerging technologies for healthcare management systems: a survey. Healthcare (Basel) 2022 Oct 3;10(10):1940. [doi: 10.3390/healthcare10101940] [Medline: 36292387]
10. Fahy E, Hardikar R, Fox A, Mackay S. Quality of patient health information on the Internet: reviewing a complex and evolving landscape. Australas Med J 2014;7(1):24-28. [doi: 10.4066/AMJ.2014.1900] [Medline: 24567763]
11. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
12. Ye W. The impact of internet health information usage habits on older adults' e-health literacy. Digit Health 2024;10:20552076241253473. [doi: 10.1177/20552076241253473] [Medline: 38726215]
13. Tsai YIP, Beh J, Ganderton C, Pranata A. Digital interventions for healthy ageing and cognitive health in older adults: a systematic review of mixed method studies and meta-analysis. BMC Geriatr 2024 Mar 4;24(1):217. [doi: 10.1186/s12877-023-04617-3] [Medline: 38438870]
14. Kim K, Shin S, Kim S, Lee E. The relation between eHealth literacy and health-related behaviors: systematic review and meta-analysis. J Med Internet Res 2023 Jan 30;25:e40778. [doi: 10.2196/40778] [Medline: 36716080]
15. Xie L, Zhang S, Xin M, Zhu M, Lu W, Mo PKH. Electronic health literacy and health-related outcomes among older adults: a systematic review. Prev Med 2022 Apr;157:106997. [doi: 10.1016/j.ypmed.2022.106997] [Medline: 35189203]
16. Sanchez R, Kay H, Srikanth P, Sandow L, Zhang M. Health behaviors and eHealth literacy among older adults, HINTS 2019. Innov Aging 2020 Dec 16;4(Supplement_1):226-226. [doi: 10.1093/geroni/igaa057.730]
17. Yang E, Chang SJ, Ryu H, Kim HJ, Jang SJ. Comparing factors associated with eHealth literacy between young and older adults. J Gerontol Nurs 2020 Aug 1;46(8):46-56. [doi: 10.3928/00989134-20200707-02] [Medline: 32936926]
18. Brørs G, Wentzel-Larsen T, Dalen H, et al. Psychometric properties of the Norwegian version of the Electronic Health Literacy Scale (eHEALS) among patients after percutaneous coronary intervention: cross-sectional validation study. J Med Internet Res 2020 Jul 28;22(7):e17312. [doi: 10.2196/17312] [Medline: 32720900]
19. Jiang X, Wang L, Leng Y, et al. The level of electronic health literacy among older adults: a systematic review and meta-analysis. Arch Public Health 2024 Nov 7;82(1):204. [doi: 10.1186/s13690-024-01428-9] [Medline: 39511667]
20. He AWJ, Yuan R, Luk TT, Wang KMP, Chan SSC. Boosting digital health engagement among older adults in Hong Kong: pilot pre-post study of the Generations Connect Project. JMIR Form Res 2025 May 8;9:e69611. [doi: 10.2196/69611] [Medline: 40340793]

21. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act 2011 Oct 21;8:115. [doi: 10.1186/1479-5868-8-115] [Medline: 22018588]
22. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176. [doi: 10.1159/000376585] [Medline: 25831962]
23. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
24. Du J, Jiang Y, Lloyd C, et al. Validation of Chinese version of the 5-item WHO well-being index in type 2 diabetes mellitus patients. BMC Psychiatry 2023 Nov 29;23(1):890. [doi: 10.1186/s12888-023-05381-9] [Medline: 38031007]
25. Liu T, Lu S, Leung DKY, et al. Adapting the UCLA 3-item loneliness scale for community-based depressive symptoms screening interview among older Chinese: a cross-sectional study. BMJ Open 2020 Dec 10;10(12):e041921. [doi: 10.1136/bmjopen-2020-041921] [Medline: 33303463]
26. Ghazi SN, Berner J, Anderberg P, Sanmartin Berglund J. The prevalence of eHealth literacy and its relationship with perceived health status and psychological distress during Covid-19: a cross-sectional study of older adults in Blekinge, Sweden. BMC Geriatr 2023 Jan 4;23(1):5. [doi: 10.1186/s12877-022-03723-y] [Medline: 36597040]
27. Jung SO, Son YH, Choi E. E-health literacy in older adults: an evolutionary concept analysis. BMC Med Inform Decis Mak 2022 Jan 31;22(1):28. [doi: 10.1186/s12911-022-01761-5] [Medline: 35101005]
28. Fung KK, Hung SSL, Lai DWL, Shum MHY, Fung HW, He L. Access to information and communication technology, digital skills, and perceived well-being among older adults in Hong Kong. Int J Environ Res Public Health 2023 Jun 23;20(13):6208. [doi: 10.3390/ijerph20136208] [Medline: 37444058]
29. Wang Y, Song Y, Zhu Y, Ji H, Wang A. Association of eHealth literacy with health promotion behaviors of community-dwelling older people: the chain mediating role of self-efficacy and self-care ability. Int J Environ Res Public Health 2022;19(10):6092. [doi: 10.3390/ijerph19106092]
30. Zhao YC, Zhao M, Song S. Online health information seeking among patients with chronic conditions: integrating the health belief model and social support theory. J Med Internet Res 2022 Nov 2;24(11):e42447. [doi: 10.2196/42447] [Medline: 36322124]
31. Ao SH, Zhu Y, Wang J, Zhao X. ehealth use and psychological health improvement among older adults: the sequential mediating roles of social support and self-esteem. Digit Health 2025;11:20552076251346659. [doi: 10.1177/20552076251346659] [Medline: 40487887]
32. Yang C, Lai DWL, Sun Y, Ma CY, Chau AKC. Mobile application use and loneliness among older adults in the digital age: insights from a survey in Hong Kong during the COVID-19 pandemic. Int J Environ Res Public Health 2022;19(13):7656. [doi: 10.3390/ijerph19137656]
33. Paige SR, Krieger JL, Stellefson ML. The influence of eHealth literacy on perceived trust in online health communication channels and sources. J Health Commun 2017 Jan;22(1):53-65. [doi: 10.1080/10810730.2016.1250846] [Medline: 28001489]
34. Xie L, Mo PKH. Prospective relationship between electronic health literacy and health-promoting lifestyle among Chinese older adults: a three-wave longitudinal study. Soc Sci Med 2025 Jan;364:117166. [doi: 10.1016/j.socscimed.2024.117166] [Medline: 39586136]
35. Tan G, Chao J, Jin S, et al. The association between digital health literacy and health inequalities among Chinese older adults: a multicenter cross-sectional study. Digit Health 2025;11:20552076251349901. [doi: 10.1177/20552076251349901] [Medline: 40547434]
36. Milanti A, Chan DNS, Parut AA, So WKW. Determinants and outcomes of eHealth literacy in healthy adults: a systematic review. PLoS One 2023;18(10):e0291229. [doi: 10.1371/journal.pone.0291229] [Medline: 37792773]
37. Lee J, Lee EH, Chae D. eHealth literacy instruments: systematic review of measurement properties. J Med Internet Res 2021;23(11):e30644. [doi: 10.2196/30644]
38. Xie L, Mo PKH. Comparison of eHealth Literacy Scale (eHEALS) and Digital Health Literacy Instrument (DHLI) in assessing electronic health literacy in Chinese older adults: a mixed-methods approach. Int J Environ Res Public Health 2023 Feb 13;20(4):3293. [doi: 10.3390/ijerph20043293] [Medline: 36833987]
39. Huang YQ, Liu L, Goodarzi Z, Watt JA. Diagnostic accuracy of eHealth literacy measurement tools in older adults: a systematic review. BMC Geriatr 2023 Mar 29;23(1):181. [doi: 10.1186/s12877-023-03899-x] [Medline: 36978033]
40. Population estimates. Census and Statistics Department.: The Government of the Hong Kong Special Administrative Region; 2020. URL: https://www.censtatd.gov.hk/en/scode150.html [accessed 2025-10-20]

Abbreviations

AOR: adjusted odds ratio
CSSA: Comprehensive Social Security Assistance
DHLI: Digital Health Literacy Instrument
eHEALS: eHealth Literacy Scale
MET: metabolic equivalent of task

NCD: noncommunicable disease
OR: odds ratio
WHO-5: World Health Organization-Five Well-Being Index

Edited by H Köttl; submitted 18.Mar.2025; peer-reviewed by E Maranesi, J Chen, L Xie; accepted 29.Dec.2025; published 30.Jan.2026.

Please cite as:
Chau SL, He W, Luk TT, Chan SSC
Level of eHealth Literacy and Its Associations With Health Behaviors and Outcomes in Chinese Older Adults: Cross-Sectional Analysis of Baseline Data From a Large-Scale Community Project
JMIR Aging 2026;9:e74110
URL: https://aging.jmir.org/2026/1/e74110
doi:10.2196/74110

Experiences of Ageism in mHealth App Usage Among Older Adults: Interview Study Among Older Adults Based on Extended Unified Theory of Acceptance and Use of Technology and Risks of Ageism Models

Jiayi Sun[1*], BSc; Yawen Liu[1*], BSc; Chengrui Zhang[2], BSc; Ying Xing[3], BSc; Wanqiong Zhou[2], BSc; Wei Luan[2], MD

[1]Shuguang Clinical Medical College, Shanghai University of TCM, Shanghai, China
[2]Nursing Department, Shuguang Hospital Affiliated to Shanghai University of TCM, No. 528, Zhangheng Road, Pudong New Area, Shanghai, China
[3]Shanghai Jiao Tong University School of Nursing, Shanghai, China
[*]these authors contributed equally

Corresponding Author:
Wei Luan, MD
Nursing Department, Shuguang Hospital Affiliated to Shanghai University of TCM, No. 528, Zhangheng Road, Pudong New Area, Shanghai, China

Abstract

Background: As the global aging population accelerates, mobile health (mHealth) apps have emerged as critical tools in the health management of older people. However, the promotion of mHealth apps has faced multiple obstacles, including insufficient technological adaptation to aging, digital resistance, and ageism. The impact of ageism on technology usage experiences among older adults is influenced by mechanisms such as stereotypes and biases. Notably, extant research has not adequately explored the subjective experiences of older adults in the context of mHealth app usage scenarios.

Objectives: The present study was predicated on the extended unified theory of acceptance and use of technology model and the risks of ageism model to systematically explore and understand older adults' ageism experiences in mHealth app usage. Our objectives were to provide a reference for optimizing age-friendly design and enhancing digital health management capabilities for older adults.

Methods: This qualitative study utilized an interpretive phenomenological design and was conducted between February and April 2025. Purposive sampling was employed to select older adults with experience using mHealth apps in a Shanghai community for semistructured interviews. This study used Colaizzi's phenomenological method to analyze and summarize older adults' experiences and perceptions of ageism and to extract themes.

Results: The study identified 3 core themes: (1) internalized age stereotypes, which manifest as technological uselessness and learning barriers; (2) anxiety and avoidance behaviors caused by stereotype threat; and (3) external unfair treatment (such as age-friendly design flaws and inadequate support systems), which inhibits usage. These experiences significantly impact older adults' intention to use mHealth apps.

Conclusions: Ageism profoundly affects the engagement of older adults with mHealth apps. It is advisable to execute systematic interventions to improve digital inclusion and health self-management capabilities, including strategies to challenge age stereotypes, optimize intergenerational support, refine age-friendly design, and establish strong social support networks.

(JMIR Aging 2026;9:e79457) doi:10.2196/79457

KEYWORDS

mHealth apps; ageism; technology use; older adults; healthy aging

Introduction

Background

The aging process of China's population is accelerating and has reached an advanced stage of development. Recent statistics indicate that by the conclusion of 2024, China's population aged 60 years and older will approximate 310 million, with those aged 65 years and older reaching around 220 million. These groups represent 22% and 15.6% of the total population, respectively [1]. As the aging process advances, the population of older internet users continues to expand. The 55th Statistical Report on China's Internet Development indicates that the population of internet users aged 60 years and above has increased from 7.3 million in 2009 to 157.25 million in 2024 [2]. The internet has progressively emerged as a significant platform for older adults to obtain health information and services. Additionally, the incidence of chronic diseases in older adults has attained 75%, necessitating increasingly individualized and varied health care requirements [3]. This requires transcending conventional health care approaches to address the personalized health needs of the digital era [4]. The State Council of China has released a medium- to long-term plan (2017 - 2025) on the prevention and treatment of chronic diseases, proposing the use of information technology to promote health management [5]. Mobile health (mHealth) apps enable remote diagnostics and personalized therapies [6], particularly aiding older adults by enhancing health care access and fostering health autonomy [7]. Nevertheless, older adults exhibit low awareness and utilization rates of mobile health care [8], attributable to inadequate age-appropriate design, technological resistance, and insufficient social support [8-10]. To resolve this issue, it is necessary to improve nationwide access to digital health services. This necessitates the systematic elimination of obstacles to digital health adoption among older adults through the refinement of technical standards, optimization of service systems, and enhancement of social support.

Notably, ageism has been demonstrated to intensify the digital divide [11,12]. The World Health Organization [13] defines ageism as stereotyping, prejudice, or discrimination based on actual age [14]. This includes self-directed ageism, which refers to negative internalized beliefs about aging [15], as well as benevolent forms, such as overprotection, and hostile forms, including neglect and judgment [16,17]. Ageism impacts older adults via 3 mechanisms: internalized stereotypes, avoidance strategies, and direct discrimination experiences [18]. These mechanisms undermine the psychological well-being of older adults and exert various negative impacts on their physical and mental health [19]. Research demonstrates that ageism markedly diminishes life satisfaction among older adults, intensifies social isolation, hinders chronic disease management outcomes, and elevates the risk of depression and cognitive decline [20]. Moreover, ageism may instigate apprehension regarding operational mistakes and reduce self-efficacy by leading older adults to internalize adverse stereotypes such as "techno-phobia." This markedly diminishes their willingness to engage with mHealth programs, thereby hindering their capacity to access health resources through mHealth apps [21]. Previous research has mostly concentrated on quantitative analyses of influencing factors [22-24]; however, it has insufficiently addressed the subjective experiences of older adults. Research on mHealth apps has identified issues such as inadequate aging [25,26], yet it seldom explores the subjective psychological effects resulting from ageism.

Objective

In order to investigate how older persons actually perceive ageism in the use of mHealth apps, this study takes an interpretive phenomenological method. It provides a reference and foundation for enhancing the age-friendly mHealth app design and creating an inclusive social support network, which will encourage older individuals' digital integration and health management capabilities.

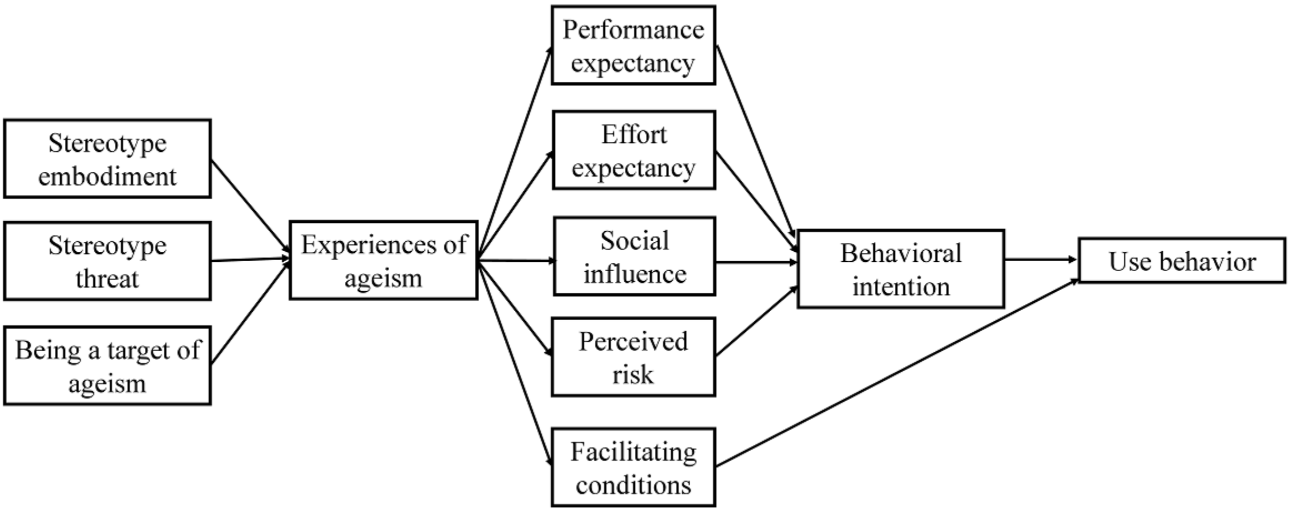
Theoretical Framework

In 2003, Venkatesh et al [27] proposed the unified theory of acceptance and use of technology (UTAUT) model, which posits that user behavior in adopting technology is explained by four core dimensions: performance expectation (PE), effort expectation (EE), social influence (SI), and facilitation conditions (FC). The model demonstrated an explanatory power of 70%. It has been extensively used in research examining the adoption of new technologies among older adults [28]. This study examines privacy risks in mobile health care services by including a perceived risk (PR) dimension into the UTAUT model to create an enhanced framework. That improvement has been substantiated in the domain of medical technology [29,30].

The risks of ageism model (RAM) proposed by Swift et al explains the obstacles to positive aging through 3 pathways [18]: stereotype embodiment (negative labels imposed on older people by society), stereotype threat (self-doubt caused by older people internalizing negative labels), and being a target of ageism. The RAM model's psychosocial mechanisms are not covered by the extended UTAUT model, although it can evaluate older adults' desire to use mHealth apps due to its five dimensions: PEs, EEs, SI, enabling factors, and PR. The RAM model focuses on the manifestation of ageism and is not directly related to technology acceptance behavior.

The extended UTAUT model, primarily utilized in quantitative research, encompasses 5 core dimensions that systematically address essential factors influencing technology acceptance behavior, thereby offering a thorough theoretical framework for examining the digital health usage behavior of older adults. This study applies qualitative research to examine the influence of ageism on older adults' perceptions and experiences within the UTAUT dimensions, aiming to reveal underlying mechanisms that quantitative research may not address. To establish a theoretical framework, as seen in Figure 1, this study combines the 2 models and maps the 3 routes of RAM to the 5 dimensions of UTAUT. This reveals how ageism affects the mechanism of interaction between the intention to use mHealth apps and behavior through these dimensions.

Figure 1. Extended unified theory of acceptance and use of technology and risks of ageism models.



Methods

Design

We conducted face-to-face semistructured interviews from February to April 2025. This study employed a qualitative research design based on interpretive phenomenology. This approach was chosen to deeply explore the subjective experiences and psychological sentiments of older adults concerning ageism in their usage of mHealth apps from their own viewpoints. This study used Colaizzi's method to understand individuals' real experiences and experiences [31]. We followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting guideline to guarantee rigor and transparency.

Participant Selection

We employed purposive sampling to select older adults in a community in Shanghai based on gender, age, and mHealth app usage experience. The selection criteria for interviewees were as follows: (1) individuals aged 60 years or older, who have resided in the community for a minimum of 5 years (or at least 10 mo/y); (2) individuals with normal cognitive function, adequate physical strength to participate in the interview, and the ability to communicate in Mandarin; (3) regular use of mHealth apps for at least 3 months; and (4) voluntary consent obtained through the signing of an informed consent form. Exclusion criteria included the presence of mental illness or cognitive impairment, severe hearing or language impairment, and withdrawal from the study. We adopted the principle of maximum differentiation sampling, with sample size determined by information saturation (ie, no new themes appeared in the interview content).

Eligible interviewees were recruited through community hospital nurses. The study conducted face-to-face semistructured interviews from February to April 2025, with the interview outline designed based on the RAM and UTAUT theoretical frameworks. The research design emphasized ethical approval to ensure the protection of participants' privacy.

Ethical Considerations

This study was approved by the hospital ethics committee (RA-2021 - 465). All participants were informed that their participation was voluntary and anonymous and that no adverse consequences would result from the interview. All participants signed written informed consent forms after being informed of the study's purpose, procedures, risks, and benefits, and received guidance on mHealth apps usage as compensation. We refer to all interviewees with a number (N) and a letter to ensure their anonymity.

Setting

To facilitate comprehension among older adults regarding the study, information leaflets were handed out, and interviews were conducted in soundproof rooms to ensure privacy and minimize interference from unrelated individuals. The interviewer initially articulated the study's purpose and significance to the interviewee, ensuring that their personal information would remain confidential. Following that, after establishing trust through a 30-minute warm-up, the interviewer conducted a semistructured interview lasting 25 to 35 minutes, employing the "dual recording method" (audio recording along with written notes). The interviewer dynamically adjusted the questioning strategy and documented nonverbal cues, such as facial expressions and body language, using information saturation as the criterion for termination. The interviewer (first author) transcribed the recorded interviews verbatim within 24 hours and invited the interviewees to verify the transcripts. Data were encrypted and stored, accessible only by the project team within an ethical framework, in full compliance with the Declaration of Helsinki requirements.

Data Collection

After obtaining approval from the community hospital, we contacted the hospital staff to determine the visit time. Before inviting eligible residents to sign the informed consents, the researcher (first author) explained the study to them. After signing the informed consents, data were collected through semistructured face-to-face interviews and observations. The first author conducted individual semistructured face-to-face interviews from February to April 2025.

Following the identification of participants, researchers engaged with them to observe their daily utilization of mHealth apps, including the types of apps and their level of proficiency in usage. Using the extended UTAUT and RAM theoretical framework, we employed an interpretive phenomenological approach to develop semistructured interview outline. Following 2 rounds of Delphi expert consultation (n=5) and 3 discussions among the research team, along with practical feedback from 2 preliminary interviews (n=2), we developed a semistructured interview outline consisting of 6 theoretical dimensions: (1) EE: Which mHealth apps have you used? Can you give us a brief overview of your learning and usage experience? (2) PE: What is your attitude toward mHealth apps? Have they met your expectations? (3) SI: What is the attitude of your family, friends, or health care providers toward your usage of mHealth apps, and how does their attitude influence your usage? (4) FC and PR: What factors do you think influence your usage of mHealth apps? (5) In your opinion, what kind of mHealth app would be most suitable for you and your peers to use? (6) What kind of help would you like to receive when using mHealth apps? The research process strictly followed qualitative research standards. All interview questions were optimized for readability (Flesch-Kincaid index≤6.0) and cognitive adaptability testing to ensure that they were understandable to older interviewees.

Probing questions within the interview allowed participants to raise unexpected issues and provided flexibility to follow-up on these issues. By asking follow-up questions based on the answers to previous questions, interviewees were encouraged to freely share their experiences. The interview guide was used to ensure that all topics were covered.

Data Analysis

This study employed a systematic qualitative data analysis approach: initially, interview data was anonymized and assigned identification codes, while data collection and analysis occurred concurrently. Transcription was finalized within 24 hours post-interviews, with accuracy confirmed by a research team member and observational notes incorporated as supplementary data. The Colaizzi’s method [31] was employed alongside NVivo 14.0 software to facilitate the analysis, resulting in a thematic map established through a 3-tier quality control process involving analysis by principal investigators, expert supervision, and review by the research team. Disputed content was addressed through consensus meetings to ensure the research’s rigor and the interpretation’s reliability.

Rigor

The content of the interview outline was determined through literature review and theoretical framework, and semistructured interviews were conducted with the target group in advance (not included in the study) to ensure that the interview outline was rigorous and easy to understand. The interviews were carried out by master’s degree nursing students trained in qualitative research methodologies. To ensure the reliability of the collected data, we employed a combination of prolonged exposure, comprehensive analysis, diverse information sources, and various data collection methods, including interviews, field notes, and member checking by colleagues and participants. In the data analysis, efforts were made to incorporate the interviewees’ emotions while minimizing the influence of the researchers’ preconceived notions.

Results

Overview of Data and Analysis

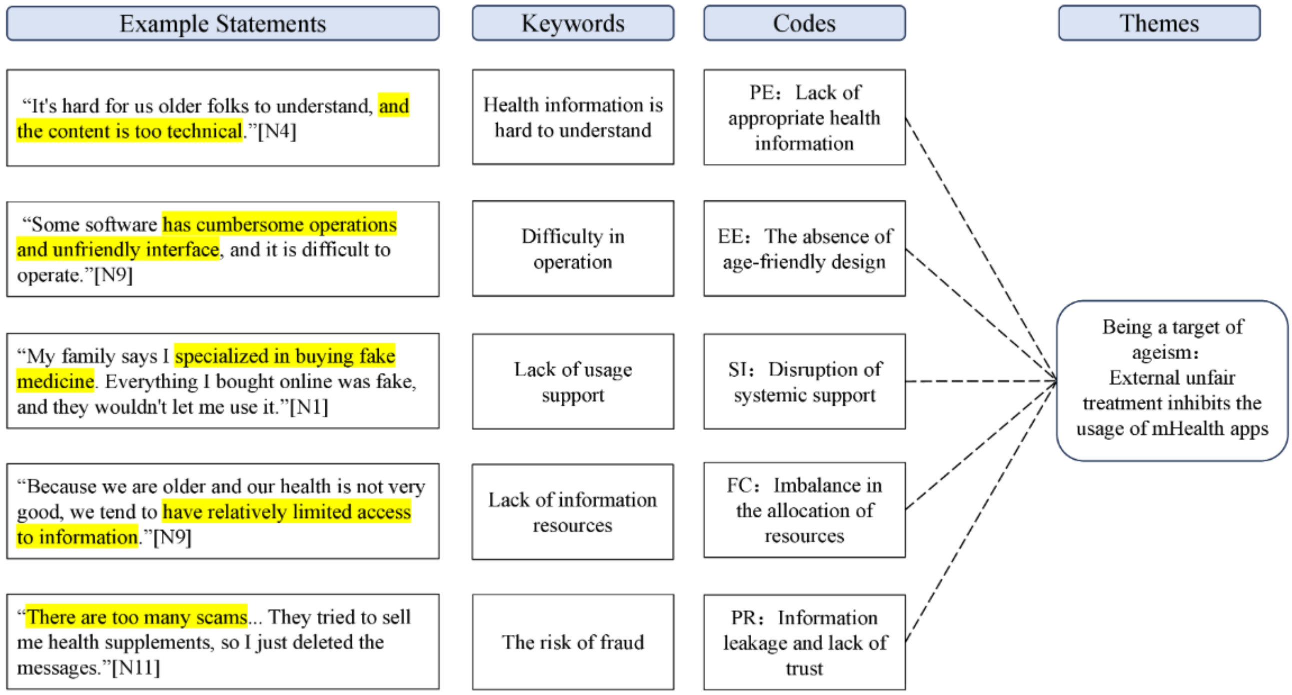
The study achieved data saturation after conducting interviews with 12 participants, at which point no new codes emerged, leading to the termination of the interviews. A total of 12 older adults were included in the study. The total duration of the interviews was approximately 370 minutes, with the transcripts of the relevant themes totaling around 50,000 words. Table 1 presents the general information of the 12 interviewees. The analysis process is shown in Figure 2, which only showed the analysis process of Theme 3.

Table . General information of the interviewees (n=12).

Code	Gender	Age, y	Monthly household income[a]	Education	Living arrangement	Marital status	Health insurance	Self-reported health status
N1	Female	66	3000 - 5000 (US $420-700)	Middle school	Three-generation co-residence	Married	URBMI[b]	Excellent
N2	Female	67	3000 - 5000 (US $420-700)	Middle school	Living with spouse	Married	UEBMI	Excellent
N3	Female	77	5000 - 10,000 (US $700-1400)	High school	Living alone	Widowed	UEBMI[c]	Good
N4	Male	73	5000 - 10,000 (US $700-1400)	Middle school	Living with spouse	Married	URBMI	Fair
N5	Male	75	1000 - 3000 (US $140-420)	High school	Living with spouse	Married	URBMI	Good
N6	Female	69	5000 - 10,000 (US $700-1400)	High school	Living alone	Divorced	URBMI	Fair
N7	Female	65	5000 - 10,000 (US $700-1400)	High school	Living with spouse	Married	URBMI	Good
N8	Male	72	3000 - 5000 (US $420-700)	High school	Living with children	Married	NCMS[d]	Fair
N9	Female	74	3000 - 5000 (US $420-700)	High school	Living with spouse	Married	UEBMI	Good
N10	Male	63	1000 - 3000 (US $140-420)	Middle school	Three-generation co-residence	Married	NCMS	Fair
N11	Male	75	5000 - 10,000 (US $700-1400)	High school	Living with spouse	Married	URBMI	Fair
N12	Female	66	5000 - 10,000 (US $700-1400)	Middle school	Living with spouse	Married	UEBMI	Good

[a]A currency exchange rate of CNY 1 = US$ 0.14 (as of December 2025) is applicable for converting monthly household income.
[b]URBMI: Urban Resident Basic Medical Insurance.
[c]UEBMI: Urban Employees Basic Medical Insurance.
[d]NCMS: New Cooperative Medical Scheme.

Figure 2. Process of thematic analysis (theme 3). EE: effort expectation; FC: facilitation condition; PE: performance expectation; PR: perceived risk; SI: social influence.



Theme 1: Stereotype Embodiment

PE: Age Labeling Weakens the Perceived Value of mHealth Apps

The internalization of age-related stereotypes caused participants to associate advanced age with a decrease in technological skills, resulting in a self-perception of "technological uselessness." The reduction in self-efficacy significantly impaired their acknowledgment of the potential health management advantages provided by mHealth apps. As N11 articulated:

> *I'm 75 years old. Why should I learn anything new? I'm getting old. These mHealth apps won't help much.*

It is worth noting that some interviewees actively use technology to compensate for age-related cognitive decline, reflecting compensatory strategies in individual behavior under the threat of stereotypes.

> *If there is something wrong with my body, I will look for answers.* [N9]

> *The doctor also talked about it (health-related knowledge), but we don't have very excellent recollections, so we can't remember anything. That's why we look for information online on our own.* [N10]

EE: Magnification Technology Usage Barriers Attributed to Age

Participants frequently ascribe operational challenges to the unavoidable cognitive and physical decline linked to aging, rather than to design deficiencies or insufficient guidance, thus exacerbating perceived obstacles to usage. N12 stated:

> *It's best if it's simple (in terms of operation and functions). We (elderly people) sometimes don't really understand (complex) programs.*

N11's attribution to educational background: "We have a relatively low level of education, so we don't really know how to use many of the functions, and it's quite difficult to learn," forming self-imposed barriers to technical learning. This cognitive pattern of attributing operational failures to age is essentially a concrete manifestation of implicit age stereotypes in the field of technology.

SI: Family Environment Reinforces Negative Expectations About Technology

External stereotypes are internalized as self-perceptions through significant individuals, particularly family members. Respondents reported that specific attitudes held by family members—such as the belief that "older adults are easily misled" (N10) and the idea that basic mobile phones are adequate for seniors (N8)—contributed to their negative self-perceptions about technological abilities. This indicates that the perpetuation of negative stereotypes within the social milieu substantially impedes the inclination to engage with technology.

> *My family told me not to believe what I read online because it's all scams aimed at old people, and they're worried that I'll get scammed.* [N10]

> *My kids said that basic phones for seniors are fine, so why do I need a smartphone?* [N8]

FC: Self-Imposed Constraints Hinder the Utilization of External Resources

Internalized stereotypes are evident in 2 specific behavioral patterns: intentional avoidance of acquiring new technologies and an overdependence on intergenerational support. The self-identification as "unlearnable" served as a significant obstacle, hindering the effective conversion of available external support resources into realized technical skills. As a result, a disconnect arose between the availability of facilitating conditions and their actual behavioral application, thereby constraining the practical adoption of mHealth apps.

> *We're not like young people; we don't want to learn new functions; these functions are enough for us.* [N12]
>
> *I'm lazy. I won't learn how to use (mHealth apps) unless I have to. My kids will help us if I need it.* [N5]

Theme 2: Stereotype Threat

PE: Avoidance of Perceived Value of mHealth Apps

Internalized age stereotypes diminish older individuals' confidence in the health advantages of technology, thereby creating a cycle of self-fulfilling prophecy. N12 said,

> *I can't find the answers I want (in the software) … I like going to the hospital to visit a doctor more because I trust what they say more.*

This preference indicates a strategic avoidance of digital health information, motivated by a concern that unsuccessful attempts may reinforce the stereotype that older individuals are unable to effectively use technology. Meanwhile, N10's prudent disposition ("My mind works slowly, so I'm not very willing to try new things unless they're truly useful") illustrated the inhibition of perceived technological efficacy induced by stereotype threat, stemming from the allocation of working memory resources to anxiety.

EE: Psychological Attributions for Avoiding Technical Learning

Older adults attributed operational challenges to the inevitability of aging, creating a cognitive pattern of technical inertia. The assertions "I'm getting older and don't want to learn new things" in N7 and "I'm not familiar with it (mHealth apps) and am too lazy to learn; my son will assist me" in N4 suggest that age-related labels generate negative psychological expectations and significantly reduce the motivation to develop technological skills. This attribution corresponds with the behavioral pattern of stereotype threat characterized by low expectations and low investment, leading to additional impairment of executive processes, such as cognitive flexibility.

SI: Usage Avoidance Under Interpersonal Pressure

External stereotypes were internalized as self-restrictive ideas through intergenerational exchanges. N8 expressed that his children's belief that "the elderly are easily deceived" made him feel "afraid to look, lest my children get into trouble," while N10 stopped asking questions after encountering his children's "impatient" attitude, illustrating how adverse feedback within the familial context reinforced avoidance behavior. This social pressure may be viewed as a manifestation of "group identity threat," leading older adults to preserve interpersonal harmony by refraining from technological involvement.

PR: Safety Concerns Intensify the Crisis of Confidence

Technical risk issues combined with traditional age biases resulted in a dual barrier to trust. The primary concerns for older adults using mHealth apps were the security of personal information and the protection of property. N3's awareness of concealed charges ("they charge you after asking a few questions"), N2's apprehension and concern regarding potential fraud ("you will be scammed if you click on it"), and N12's doubt regarding the accuracy of information ("it's a mix of truth and falsehood, so it's better to go to the hospital") illustrated the spillover effect of stereotype threat; anxiety consistently undermined trust in technology, even when detached from specific contexts.

Theme 3: Being a Target of Ageism

PE: Lack of Appropriate Health Information

Systematically disregarding the cognitive traits of older adults resulted in ineffective health content. N3 indicated that "the software content (in the software) is too difficult to understand… too technical and hard to remember," while N4 complained that "the content is too technical." These perceptions indicate a failure in information design to accommodate age-related changes in cognitive processing, such as declines in working memory capacity, thereby reducing the perceived usefulness and accessibility of the information provided. N5 revealed deceptive advertising ("the recommended products claim to lower blood sugar, but there is no scientific basis for this"), which decreased trust in technology and highlighted the exploitation of older adults for profit.

EE: The Absence of Age-Friendly Design

There were considerable age-related deficiencies in interface interaction. N4 emphasized the lack of physiological adaptation, stating that "There is no version specifically designed for the elderly... the font should be enlarged." N8 emphasized the necessity for multimodal requirements, asserting that "the voice version should be directly audible." N9 criticized the "cumbersome operations and unfriendly interface," specifically highlighting the design issue of prioritizing "technology-centric" over "user-centric" in product development. The lack of age-appropriate design increased the learning burden for older adults and reduced the user experience.

SI: Disruption of Systemic Support

The intergenerational influence within families significantly impacted technology usage. N1 was designated as "specialized in buying fake medicine" and prohibited from using mHealth apps, while N4 encountered limitations due to his children's concerns regarding possible financial exploitation, demonstrating that protective measures resulted in technological deprivation.

The health care system offered minimal professional support. N10 noted, "The medical staff didn't mention it and told me to go to the hospital," whereas N1 stated, "Sometimes when I ask them, they get a little impatient."

FC: Imbalance in the Allocation of Resources

Public health education resources were insufficiently integrated into digital platforms. The lack of resources intensified the technological marginalization of older adults. Some older adults exhibited confidence in official information sources and a desire to gain knowledge.

There are no such platforms available on a regular basis, so I don't know what software is available or which software is suitable for us. Apart from what the nurses at the hospital recommend, I don't know anything else. [N8]

Because we are older and our health is not very good, we tend to have relatively limited access to information. [N9]

PR: Information Leakage and Lack of Trust

Older adults often mentioned challenges associated with breaches of personal information, excessive promotional communications, and misleading marketing practices from unregistered medical institutions.

I don't know how they got my personal information. Many "doctors" send me text messages. [N10]

They recommend products that claim to cure diabetes, but that's a scam… We also try to verify the authenticity of such information. [N8]

There are too many scams… They tried to sell me health supplements, so I just deleted the messages. [N11]

Discussion

Principal Findings

This study examined the ageism encountered by older adults in the community while using mHealth apps and the mechanisms through which it impacts them. The primary findings indicate that ageism obstructs the adoption and utilization of mHealth apps via 3 principal pathways of the RAM model: Stereotype internalization in older adults leads to the attribution of operational difficulties to aging, resulting in a self-perception of technological incompetence. This perception diminishes their acknowledgment of the benefits of mHealth and establishes a challenging learning threshold to surpass. Stereotype threat induces anxiety avoidance, characterized by fears of privacy breaches, worries regarding operational failures, and excessive protection or negative feedback from family members. These factors collectively result in the active avoidance of technology exploration and usage. External unfair treatment, evident in age-inappropriate design flaws in apps, including complex interfaces, limited information availability, and insufficient support systems, creates barriers to usage and diminishes the perceived value of these apps. Research indicates that ageism operates at various levels, including individual cognition, social interaction, and the technological environment, creating substantial barriers that impede older adults' integration into digital health.

Reconstructing Digital Health Cognition in Older Adults: A Dual-Pathway Intervention to Enhance mHealth App Usage Effectiveness

Older adults typically internalize the stereotype that "age dictates technological ability." This belief leads to attributing operational failures to age, avoiding the learning of new technologies, and concerns about inconveniencing others. Consequently, this significantly diminishes their willingness to use mHealth apps and undermines their sense of self-efficacy [32]. This cognitive pattern establishes obstacles in 3 dimensions—PE (diminished perceived health value), EE (reduced learning motivation), and FC (increased perceived learning costs)—illustrating the self-fulfilling prophecy of stereotype threat within the RAM model [33-35], thereby compromising well-being and impacting physical health [36]. Interventions should implement a dual approach to address this phenomenon. Positive narratives should be employed to highlight success stories among peers, such as in chronic disease management, underscoring that proficiency with technology is contingent upon practice rather than age, while minimizing the notion of "technological disadvantage." Meanwhile, an educational approach characterized by "low threshold+high feedback" should be adopted. This involves deconstructing essential functions, such as blood pressure monitoring and medication reminders, while offering voice navigation and immediate feedback to build successful experiences, improve self-efficacy, and encourage proactive health behaviors [37].

From Substitution to Empowerment: Developing a Novel Paradigm for Digital Health Support for Older Adults

Research indicates that benevolent ageism, characterized by overprotective family members and insufficient support from medical personnel, exacerbates older individuals' perceptions of incompetence [16]. The original intention may be to protect older adults; however, it implicitly stereotypes them as deficient in information, judgment, and technical skills, thereby diminishing their willingness to utilize mHealth apps [17,18]. This discrimination establishes a vicious cycle of "external rejection-self-rejection-behavioral withdrawal" through the "being a target of ageism" path in the RAM model and the "social influence" dimension in the UTAUT model. Studies indicate that the advice and trust that kids give notably affect older adults' willingness to adopt technology, while positive interactions between generations may improve perceptions of aging [38], Studies indicate that the advice and trust that kids give notably affect older adults' willingness to adopt technology, while positive interactions between generations may improve perceptions of aging [39]. Digital reverse mentoring should be promoted to enhance older adults' information literacy, while improving medical staff's negative perceptions of older adults [40]. The objective is to shift the support framework from "substitution" to "empowerment and accompaniment," motivating older adults to assume control over device operation, acknowledging their advancements promptly, eliminating their self-perception as "technologically disadvantaged," and enhancing their sense of worth as participants to digital health.

Eliminating Invisible Technological Barriers: A Dual-Path Approach to Aging-Friendly Design and Risk Prevention in mHealth Apps

Research indicates that only 40% of mHealth apps incorporate older adults in the design process [41], and challenges such as intricate interfaces, superfluous operations, and insufficient information availability highlight invisible ageism at the technology level. Analysis based on the UTAUT model: in the EE dimension, design deficiencies such as small font sizes and poor-quality push notifications elevate learning costs [42]; in the PE dimension, an overabundance of technical terms and insufficient personalized guidance diminish perceived utility [43]; in the PR dimension, apprehensions regarding privacy breaches and misleading advertisements undermine usage intentions [44,45], all of which exacerbate older adults' self-perceptions of technological discrimination [41,46]. We advocate adopting a dual-faceted strategy: in terms of technology, implement age-appropriate features such as adjustable font sizes and voice navigation [47], create specialized and accessible health information [48], and engage older adults in the initial design stages to mitigate stereotypes [41]. In terms of risk prevention and control, establish data protection methods to eliminate hazardous information [49], while strengthening governmental regulation to create a trustworthy environment. By systematically optimizing design and management processes, obstacles to utilization by older adults can be efficiently reduced, consequently enhancing digital health inclusivity.

Multistakeholder Collaborative Empowerment: Constructing a Social Support System for Older Adults' Digital Health

Findings indicate that older adults frequently avoid engaging with mHealth apps due to a lack of informational resources and prevailing stereotypes (eg, "incapable of learning," "poor judgment") [50], which results in a diminished perception of resource support in the FC dimension. A multitiered support network is essential: medical institutions should incorporate mHealth guidance into health management services, and medical staff must acquire aging-friendly instructional skills to establish trust through recommendations from authoritative institutions. Research indicates that 55.5% of older internet users have encountered online risks [51], while also demonstrating significant vigilance [52]. Governments need to encourage platforms to enhance the digital environment, optimize recommendation algorithms, and implement antifraud training to bolster information discrimination capabilities. Through the combination of professional support from the medical system and social risk prevention and control, it is feasible to address older adults' demand for authoritative information while mitigating digital risks. This approach enhances the perceived benefits of mHealth apps and promotes healthy aging.

This study systematically elucidates the psychosocial mechanisms by which age discrimination affects older adults' adoption of mHealth apps, integrating and extending the UTAUT and RAM models. This study extends the application of the UTAUT model, confirming its efficacy and potential as a qualitative research framework for examining the subjective experiences of older adults. The integration of the RAM model with UTAUT elucidates the impact of ageism's psychosocial mechanisms on critical aspects of technology acceptance, providing fresh insights into the digital health barriers faced by older adults. These findings have practical implications for improving the delivery of digital health services in the communities and hospitals of Shanghai. Community health promotion activities must challenge stereotypes regarding older adults, enhancing their confidence through peer modeling and prompt feedback. Second, families and health care institutions ought to transition from "substitute operation" to "empowering accompaniment," thereby reducing psychological barriers through enhanced communication and initial guidance. Ultimately, the optimization of technology and services should be closely aligned with distinct user requirements, including interface complexity, information credibility, and the availability of continuous support.

Limitations

First, the sample size of this qualitative study is relatively small. Although small, the sample has reached theoretical saturation, meaning that adding more participants is unlikely to yield new insights. The current sample is drawn exclusively from a single community in Shanghai, which may introduce geographical and cultural biases that limit the generalizability of the findings. Second, despite using purposive sampling, selection bias may still be present, as participants who are more positive or more persistent in their usage of mHealth apps are more likely to participate. To mitigate this effect, we deliberately recruited participants who reported lower levels of mHealth apps usage. Additionally, the study was limited to older adults living in the community and did not include older adults in institutions such as nursing homes, which may have overlooked the experiences of ageism among more vulnerable groups. Therefore, future studies should include a broader population for comparative analysis to gain a more comprehensive understanding of the experiences of ageism in the use of mHealth apps among different groups.

In addition, this study used qualitative research methods, which are inherently subjective. To improve the scientific rigor of our research, we referenced qualitative research quality standards and focused on the following areas: researchers engaged in all interviews, recordings, and transcription processes to establish trust with participants, thereby enhancing credibility. Purposeful sampling was utilized to enhance the transferability of findings, accompanied by comprehensive descriptions of the sample characteristics. To ensure dependability, all interviews were audio-recorded, and comprehensive interview notes, transcripts, and research reflection journals were preserved. To ensure confirmability, 2 researchers performed independent data analyses, sought third-party consensus in cases of discrepancies, and provided organized data to research participants for verification. Although we used rigorous qualitative analysis to minimize bias, incorporating quantitative methods could provide more objective data, thereby improving the credibility and scientific rigor of the results. Future studies should adopt a mixed-method approach, combining qualitative and quantitative data to gain a more comprehensive understanding, provide objective indicators to enhance the reliability of the research results, and quantify the impact of ageism on the usage of mHealth apps. Finally, the duration of this study may limit our ability to track long-term changes in older adults' experiences of ageism. Future studies should include long-term follow-up to assess the temporal changes in ageism experiences and technology acceptance behavior, thereby providing more comprehensive insights and practical guidance for intervention design.

Conclusions

In this study, based on the extended UTAUT and RAM models, we identified 3 significant issues related to ageism in the usage of mHealth apps among older adults in the community: internalized stereotypes, benevolent ageism, and aging-friendly design flaws. We propose intervention strategies throughout 4 dimensions: psychological cognition, support systems, technological optimization, and social networks. This research is confined to a singular community sample located in Shanghai. Future studies should enhance the geographical and cultural diversity of the sample, implement longitudinal tracking and quantitative methods, and perform comprehensive research on the causal relationship between ageism and mHealth apps usage behavior to establish a foundation for creating an inclusive digital health environment.

Funding

This work was supported by the Open fund of key laboratory of ministry of education for long-term care of the older adults (LNYB-2023- ‑ 12), the Shanghai Educational Science Research Project (C2023136), and the Shanghai Shenkang Hospital Development Center Project, the fifth batch of standardized management and promotion projects—Acute and Critical Illnesses Integrated Chinese and Medicine Professional Nursing Technique Training Base (SHDC22024209).

Data Availability

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JS (lead), WL (equal), YX (supporting)
Data curation: JS (lead), YL (equal)
Formal analysis: JS (lead), YL (equal), YX (supporting), CZ (supporting)
Funding acquisition: WL
Investigation: JS
Methodology: JS (lead), WL (equal), YX (supporting)
Project administration: WL (lead), WZ (supporting)
Resources: WL
Supervision: WL (lead), WZ (supporting)
Validation: JS (lead), WL (supporting)
Visualization: JS (lead), YL (equal)
Writing – original draft: JS (lead), YL (equal), YX (supporting), CZ (supporting)
Writing – review & editing: WL (lead), WZ (supporting)

Conflicts of Interest

None declared.

References

1. Population size decline narrows, population quality continuously improves. National Bureau of Statistics. 2025. URL: https://www.stats.gov.cn/sj/sjjd/202501/t20250117_1958337.html [accessed 2025-11-16]
2. The 55th statistical report on internet development in China [Article in Chinese]. China Internet Network Information Center (CNNIC). 2024. URL: https://www.cnnic.net.cn/NMediaFile/2025/0428/MAIN17458061595875K4FP1NEUO.pdf [accessed 2025-12-17]
3. Su B, Li D, Xie J, et al. Chronic disease in China: geographic and socioeconomic determinants among persons aged 60 and older. J Am Med Dir Assoc 2023 Feb;24(2):206-212. [doi: 10.1016/j.jamda.2022.10.002] [Medline: 36370750]
4. Spinean A, Mladin A, Carniciu S, Stănescu AMA, Serafinceanu C. Emerging methods for integrative management of chronic diseases: utilizing mHealth apps for lifestyle interventions. Nutrients 2025 Apr 29;17(9):1506. [doi: 10.3390/nu17091506] [Medline: 40362815]
5. Notice of the general office of the state council on issuing the medium- and long-term plan for the prevention and control of chronic diseases in China [Article in Chinese]. General Office of the State Council. 2017. URL: https://www.gov.cn/zhengce/content/2017-02/14/content_5167886.htm [accessed 2025-12-17]
6. Alkhaldi O, McMillan B, Maddah N, Ainsworth J. Interventions aimed at enhancing health care providers' behavior toward the prescription of mobile Health apps: systematic review. JMIR Mhealth Uhealth 2023 Feb 27;11:e43561. [doi: 10.2196/43561] [Medline: 36848202]
7. Sobrinho AS, Gomes GO, Bueno Júnior CR. Developing a multiprofessional mobile app to enhance health habits in older adults: user-centered approach. JMIR Form Res 2024 Apr 15;8:e54214. [doi: 10.2196/54214] [Medline: 38619865]
8. Askari M, Klaver NS, van Gestel TJ, van de Klundert J. Intention to use medical apps among older adults in the Netherlands: cross-sectional study. J Med Internet Res 2020 Sep 4;22(9):e18080. [doi: 10.2196/18080] [Medline: 32624465]

9. Ahmad NA, Mat Ludin AF, Shahar S, Mohd Noah SA, Mohd Tohit N. Willingness, perceived barriers and motivators in adopting mobile applications for health-related interventions among older adults: a scoping review. BMJ Open 2022 Mar 9;12(3):e054561. [doi: 10.1136/bmjopen-2021-054561] [Medline: 35264349]
10. Verloo H, Kampel T, Vidal N, Pereira F. Perceptions about technologies that help community-dwelling older adults remain at home: qualitative study. J Med Internet Res 2020 Jun 4;22(6):e17930. [doi: 10.2196/17930] [Medline: 32496197]
11. Mariano J, Marques S, Ramos MR, Gerardo F, de Vries H. Too old for computers? The longitudinal relationship between stereotype threat and computer use by older adults. Front Psychol 2020;11:568972. [doi: 10.3389/fpsyg.2020.568972] [Medline: 33123050]
12. McDonough CC, University of Massachusetts Lowell. The effect of ageism on the digital divide among older adults. GGM 2016 Jun 16;2(1):1-7 [FREE Full text] [doi: 10.24966/GGM-8662/100008]
13. Global report on ageism. World Health Organization. 2021. URL: https://www.who.int/teams/social-determinants-of-health/demographic-change-and-healthy-ageing/combatting-ageism/global-report-on-ageism [accessed 2025-12-17]
14. Officer A, Schneiders ML, Wu D, Nash P, Thiyagarajan JA, Beard JR. Valuing older people: time for a global campaign to combat ageism. Bull World Health Organ 2016 Oct 1;94(10):710-710A. [doi: 10.2471/BLT.16.184960] [Medline: 27843156]
15. Henry JD, Coundouris SP, Craik FIM, von Hippel C, Grainger SA. The cognitive tenacity of self-directed ageism. Trends Cogn Sci 2023 Aug;27(8):713-725. [doi: 10.1016/j.tics.2023.03.010] [Medline: 37147237]
16. Apriceno M, Lytle A, Monahan C, Macdonald J, Levy SR. Prioritizing health care and employment resources during COVID-19: roles of benevolent and hostile ageism. Gerontologist 2021 Jan 21;61(1):98-102. [doi: 10.1093/geront/gnaa165] [Medline: 33119089]
17. Cary LA, Chasteen AL, Remedios J. The ambivalent ageism scale: developing and validating a scale to measure benevolent and hostile ageism. Gerontologist 2016 Aug 12:gnw118. [doi: 10.1093/geront/gnw118]
18. Swift HJ, Abrams D, Lamont RA, Drury L. The risks of ageism model: how ageism and negative attitudes toward age can be a barrier to active aging. Soc Issues Policy Rev 2017 Jan;11(1):195-231 [FREE Full text] [doi: 10.1111/sipr.12031]
19. Jackson SE, Hackett RA, Steptoe A. Associations between age discrimination and health and wellbeing: cross-sectional and prospective analysis of the English longitudinal study of ageing. Lancet Public Health 2019 Apr;4(4):e200-e208. [doi: 10.1016/S2468-2667(19)30035-0] [Medline: 30954145]
20. Dix E, Van Dyck L, Adeyemo S, et al. Ageism in the mental health setting. Curr Psychiatry Rep 2024 Nov;26(11):583-590. [doi: 10.1007/s11920-024-01531-2] [Medline: 39278983]
21. Mannheim I, Köttl H. Ageism and (successful) digital engagement: a proposed theoretical model. Gerontologist 2024 Sep 1;64(9):gnae078. [doi: 10.1093/geront/gnae078] [Medline: 38874215]
22. Knuutila M, Lehti TE, Karppinen H, Kautiainen H, Strandberg TE, Pitkala KH. Associations of perceived poor societal treatment among the oldest-old. Arch Gerontol Geriatr 2021;93:104318. [doi: 10.1016/j.archger.2020.104318] [Medline: 33310658]
23. Tice-Brown D, Kelly P, Heyman JC, Phipps C, White-Ryan L, Davis HJ. Older adults' perceptions of ageism, discrimination, and racism. Soc Work Health Care 2024;63(6-7):415-432. [doi: 10.1080/00981389.2024.2365136] [Medline: 38899560]
24. Wang X, Wang Y, Zhang F, Ge D, Guo Z. Associations between social isolation, perceived ageism and subjective well-being among rural Chinese older adults: a cross-sectional study. Geriatr Nurs 2024;59:598-603. [doi: 10.1016/j.gerinurse.2024.08.014] [Medline: 39178626]
25. Kaihlanen AM, Virtanen L, Buchert U, et al. Towards digital health equity - a qualitative study of the challenges experienced by vulnerable groups in using digital health services in the COVID-19 era. BMC Health Serv Res 2022 Feb 12;22(1):188. [doi: 10.1186/s12913-022-07584-4] [Medline: 35151302]
26. Schroeder T, Seaman K, Nguyen AD, Gewald H, Georgiou A. Social determinants of mobile Health app adoption - a qualitative study of older adults' perceptions in Australia. Stud Health Technol Inform 2023 Jun 22;304:81-85. [doi: 10.3233/SHTI230376] [Medline: 37347575]
27. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
28. Maswadi K, Ghani NA, Hamid S. Factors influencing the elderly's behavioural intention to use smart home technologies in Saudi Arabia. PLoS ONE 2022;17(8):e0272525. [doi: 10.1371/journal.pone.0272525] [Medline: 36040877]
29. Wu W, Zhang B, Li S, Liu H. Exploring factors of the willingness to accept AI-assisted learning environments: an empirical investigation based on the UTAUT model and perceived risk theory. Front Psychol 2022;13:870777. [doi: 10.3389/fpsyg.2022.870777] [Medline: 35814061]
30. Cimperman M, Makovec Brenčič M, Trkman P. Analyzing older users' home telehealth services acceptance behavior-applying an extended UTAUT model. Int J Med Inform 2016 Jun;90:22-31. [doi: 10.1016/j.ijmedinf.2016.03.002] [Medline: 27103194]
31. Englander M. The phenomenological method in qualitative psychology and psychiatry. Int J Qual Stud Health Well-being 2016;11(1):30682. [doi: 10.3402/qhw.v11.30682] [Medline: 26968361]
32. Kim J, Jeon SW, Byun H, Yi E. Exploring E-Health literacy and technology-use anxiety among older adults in Korea. Healthcare (Basel) 2023 May 25;11(11):1556. [doi: 10.3390/healthcare11111556] [Medline: 37297696]

33. Okun S, Ayalon L. The paradox of subjective age: age(ing) in the self-presentation of older adults. Int Psychogeriatr 2023 Oct;35(10):566-575. [doi: 10.1017/S1041610222000667] [Medline: 35968843]
34. Ostermeier R, Rothermund K. Age differences in age stereotypes: the role of life domain and cultural context. GeroPsych (Bern) 2022 Dec;35(4):177-188. [doi: 10.1024/1662-9647/a000272]
35. Marquet M, Missotten P, Dardenne B, Adam S. Interactions between stereotype threat, subjective aging, and memory in older adults. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2019 Jan;26(1):121-143. [doi: 10.1080/13825585.2017.1413166] [Medline: 29221428]
36. Bergman YS, Palgi Y. Ageism, personal and others' perceptions of age awareness, and their interactive effect on subjective accelerated aging. J Appl Gerontol 2021 Dec;40(12):1876-1880. [doi: 10.1177/0733464820967209] [Medline: 33095076]
37. Wang Y, Song Y, Zhu Y, Ji H, Wang A. Association of eHealth literacy with health promotion behaviors of community-dwelling older people: the chain mediating role of self-efficacy and self-care ability. Int J Environ Res Public Health 2022 May 17;19(10):6092. [doi: 10.3390/ijerph19106092] [Medline: 35627627]
38. Almulhem JA. Factors, barriers, and recommendations related to mobile health acceptance among the elderly in Saudi Arabia: a qualitative study. Healthcare (Basel) 2023 Nov 23;11(23):3024. [doi: 10.3390/healthcare11233024] [Medline: 38063592]
39. Mikton C, de la Fuente-Núñez V, Officer A, Krug E. Ageism: a social determinant of health that has come of age. Lancet 2021 Apr 10;397(10282):1333-1334. [doi: 10.1016/S0140-6736(21)00524-9] [Medline: 33743845]
40. Crutzen C, Missotten P, Adam S, Schroyen S. Does caring lead to stigmatisation? The perception of older people among healthcare professionals and the general population: a cross-sectional study. Int J Older People Nurs 2022 Sep;17(5):e12457. [doi: 10.1111/opn.12457] [Medline: 35267232]
41. Mannheim I, Wouters EJM, Köttl H, van Boekel LC, Brankaert R, van Zaalen Y. Ageism in the discourse and practice of designing digital technology for older persons: a scoping review. Gerontologist 2023 Aug 24;63(7):1188-1200. [doi: 10.1093/geront/gnac144] [Medline: 36130318]
42. Herkert C, Graat-Verboom L, Gilsing-Fernhout J, Schols M, Kemps HMC. Home-based exercise program for patients with combined advanced chronic cardiac and pulmonary diseases: exploratory study. JMIR Form Res 2021 Nov 9;5(11):e28634. [doi: 10.2196/28634] [Medline: 34751655]
43. Kim S, Chow BC, Park S, Liu H. The usage of digital health technology among older adults in Hong Kong and the role of technology readiness and eHealth literacy: path analysis. J Med Internet Res 2023 Apr 12;25:e41915. [doi: 10.2196/41915] [Medline: 37043274]
44. Klaver NS, van de Klundert J, van den Broek R, Askari M. Relationship between perceived risks of using mHealth applications and the intention to use them among older adults in the Netherlands: cross-sectional study. JMIR Mhealth Uhealth 2021 Aug 30;9(8):e26845. [doi: 10.2196/26845] [Medline: 34459745]
45. Quan - Haase A, Ho D. Online privacy concerns and privacy protection strategies among older adults in East York, Canada. J Assoc Inf Sci Technol 2020 Sep;71(9):1089-1102 [FREE Full text] [doi: 10.1002/asi.24364]
46. Shew A. Ableism, technoableism, and future AI. IEEE Technol Soc Mag 2020 Mar;39(1):40-85. [doi: 10.1109/MTS.2020.2967492]
47. Kebede AS, Ozolins LL, Holst H, Galvin K. Digital engagement of older adults: scoping review. J Med Internet Res 2022 Dec 7;24(12):e40192. [doi: 10.2196/40192] [Medline: 36477006]
48. Kraaijkamp JJM, van Dam van Isselt EF, Persoon A, Versluis A, Chavannes NH, Achterberg WP. eHealth in geriatric rehabilitation: systematic review of effectiveness, feasibility, and usability. J Med Internet Res 2021 Aug 19;23(8):e24015. [doi: 10.2196/24015] [Medline: 34420918]
49. Bonomi L, Huang Y, Ohno-Machado L. Privacy challenges and research opportunities for genomic data sharing. Nat Genet 2020 Jul;52(7):646-654. [doi: 10.1038/s41588-020-0651-0] [Medline: 32601475]
50. Choi YK, Thompson HJ, Demiris G. Internet-of-Things smart home technology to support aging-in-place: older adults' perceptions and attitudes. J Gerontol Nurs 2021 Apr;47(4):15-21. [doi: 10.3928/00989134-20210310-03] [Medline: 34038251]
51. Kemp S, Erades Pérez N. Consumer fraud against older adults in digital society: examining victimization and its impact. Int J Environ Res Public Health 2023 Apr 5;20(7):5404. [doi: 10.3390/ijerph20075404] [Medline: 37048017]
52. Nolte J, Hanoch Y, Wood S, Hengerer D. Susceptibility to COVID-19 scams: the roles of age, individual difference measures, and scam-related perceptions. Front Psychol 2021;12:789883. [doi: 10.3389/fpsyg.2021.789883] [Medline: 34975685]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
EE: effort expectation
FC: facilitation conditions
mHealth: mobile health
PE: performance expectation
PR: perceived risk

RAM: risks of ageism model
SI: social influence
UTAUT: unified theory of acceptance and use of technology

Edited by Y Luo; submitted 21.Jun.2025; peer-reviewed by A Hu, Y Sun; accepted 08.Dec.2025; published 07.Jan.2026.

Please cite as:
Sun J, Liu Y, Zhang C, Xing Y, Zhou W, Luan W
Experiences of Ageism in mHealth App Usage Among Older Adults: Interview Study Among Older Adults Based on Extended Unified Theory of Acceptance and Use of Technology and Risks of Ageism Models
JMIR Aging 2026;9:e79457
URL: https://aging.jmir.org/2026/1/e79457
doi:10.2196/79457

Original Paper

Long-Term Effects of Mobile-Based Metamemory Cognitive Training in Older Adults With Mild Cognitive Impairment: 15-Month Prospective Single-Arm Longitudinal Study

Jung-In Lim[1], MA; Yeeun Byeon[2], MS; Sunyoung Kang[3], MD; Hyeonjin Kim[4], PhD; Keun You Kim[3], MD, PhD; Lukas Stenzel[5], PhD; So Yeon Jeon[4,6], MD, PhD; Jun-Young Lee[4,6], MD, PhD

[1]Interdisciplinary Program in Cognitive Science, Seoul National University, Seoul, Republic of Korea
[2]Department of Medical Device Development, Seoul National University College of Medicine, Seoul, Republic of Korea
[3]Department of Psychiatry, Institute of Behavioral Science in Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea
[4]Department of Psychiatry, Seoul Metropolitan Government-Seoul National University Boramae Medical Center, Seoul, Republic of Korea
[5]Cogthera GmbH, Munich, Germany
[6]Department of Psychiatry, Seoul National University College of Medicine, Seoul, Republic of Korea

Corresponding Author:
Jun-Young Lee, MD, PhD
Department of Psychiatry
Seoul Metropolitan Government-Seoul National University Boramae Medical Center
20, Boramae-ro 5-gil, Dongjak-gu
Seoul,
Republic of Korea
Phone: 82 2 870 2462
Email: benji@snu.ac.kr

Abstract

Background: Mild cognitive impairment (MCI) is an intermediate state between normal aging and dementia, characterized by subjective cognitive decline and objective memory impairment. Cognitive training has consistently shown short-term benefits for individuals with MCI, but evidence on the long-term effectiveness is extremely limited. Given the progressive nature of MCI and the need for sustainable strategies to delay cognitive decline, research on the long-term impact of cognitive training is necessary and timely. Mobile-based platforms offer a promising solution by enhancing accessibility and adherence, but their durability of effect over extended periods remains underexplored.

Objective: This study aimed to evaluate the long-term effects of a mobile-based cognitive training app on the cognitive function of older adults with MCI.

Methods: In total, 28 older adults with MCI used Cogthera, a mobile cognitive training app based on metamemory training. Participants completed 2 training sessions daily for 3 months, and 9 (32%) continued for an additional 12 months. Cognitive function and quality of life were assessed using the Alzheimer's Disease Assessment Scale-Cognitive Subscale 14 and EQ-5D-5L.

Results: Cognitive function improved over 15 months, as measured by Alzheimer's Disease Assessment Scale-Cognitive Subscale ($F_{2,35.56}$=7.08; P=.003). EQ-5D-5L scores increased at 3 months but did not show sustained change at 15 months ($F_{2,42.14}$=3.40; P=.04). Greater cognitive improvements were associated with younger age, higher functional status, and lower baseline cognitive function.

Conclusions: This study showed that long-term use of a mobile-based metamemory cognitive training app was associated with cognitive improvements over 15 months. Although limited by the small sample size and the absence of a control group, these findings suggest potential for mobile cognitive training as a sustainable intervention that warrants validation in larger trials.

(JMIR Aging 2026;9:e81648) doi:10.2196/81648

KEYWORDS

Alzheimer disease; cognitive training; digital technology; mild cognitive impairment; metamemory

Introduction

Background

Mild cognitive impairment (MCI) is an intermediate stage between normal aging and Alzheimer disease (AD). It has emerged as a key target for early intervention to delay AD. MCI is characterized by subjective cognitive decline along with objective memory impairment without functional impairment [1]. The prevalence of MCI is 6.7% among individuals aged between 60 and 64 years and 25.5% among those aged between 80 and 84 years [2]. In addition, approximately 10% of individuals with MCI progress to AD annually [3].

Nonpharmacological interventions, such as physical activity, social engagement, and cognitive training, are widely recommended to delay AD onset in MCI [4-7]. Cognitive training has demonstrated superior efficacy in improving cognitive function in short-term treatment [8-11]. Despite these short-term benefits, few studies have examined whether cognitive training produces sustained effects beyond 1 year because of methodological difficulties [12]. Traditional cognitive interventions also rely on in-person, paper-based sessions, which can limit accessibility and adherence [13]. Digital cognitive training based on a mobile app enables home-based, self-paced engagement and has demonstrated greater participation [14].

We have previously reported the short-term efficacy of metamemory training (MMT) [15,16]. Metamemory refers to an individual's awareness and understanding of their memory functions, including their contents and processes [17,18]. MMT is based on this concept and has been shown to be effective in teaching mnemonic strategies and improving cognitive function [19-22]. However, its long-term efficacy may need to be investigated, especially with a mobile cognitive training app.

Objectives

This study aimed to examine the long-term effects of a cognitive training app based on MMT in older adults with MCI. We hypothesized that prolonged use of the app would lead to sustained improvements in cognitive function.

Methods

Participants

Participants were recruited from 1 memory clinic. Inclusion criteria were as follows: (1) aged between 55 and 85 years; (2) diagnosed with MCI by trained psychiatrists or neurologists according to the criteria proposed by Petersen [23]; (3) ownership of a personal smartphone and no difficulties using mobile apps; (4) if taking cognitive enhancers (acetylcholinesterase inhibitors or memantine), a stable dose maintained for at least 12 weeks before randomization; (5) availability of a caregiver who spends more than 8 hours per week with the participant and agrees to assist with follow-up and clinical evaluations; (6) ability to independently make phone calls to a caregiver using a smartphone; (7) no difficulties with reading or writing in Korean; and (8) adequate vision and hearing to participate in the clinical trial. MCI diagnosis was based on the following criteria by Petersen [23]: (1) reported concerns regarding cognitive changes, (2) impairment in 1 or more cognitive domains, (3) preservation of independence in functional abilities, and (4) absence of dementia [24]. Participants underwent cognitive assessments conducted by clinical psychologists using the Clinical Dementia Rating (CDR) and the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Assessment Battery (CERAD-NP). Eligibility criteria required a global CDR score of 0.5 and performance at least 1.0 SD below the mean in 1 or more memory domains of the CERAD-NP.

Exclusion criteria included the presence of serious physical illness or psychiatric disorders that could interfere with study procedures; the use of medications known to affect cognitive function (except those taken consistently for at least 3 months); and any neurological or medical conditions associated with cognitive decline other than MCI, such as stroke, central nervous system infection, head trauma, alcohol dependence, or depression. Of the 40 participants screened, 10 (25%) were excluded, 1 (3%) dropped out, and 1 (3%) experienced a stroke. Data from 28 (70%) participants were analyzed.

Ethical Considerations

This study was reviewed and approved by the institutional review board of Seoul Metropolitan Government-Seoul National University Boramae Medical Center (20-2022-48). All participants provided informed consent before participation. All collected data were deidentified and stored securely, and no personally identifiable information was accessible to the research team. Participants received KRW 200,000 (US $136) as compensation for their participation.

Intervention

Participants underwent cognitive training using Cogthera, a smartphone-based app based on MMT developed by Youn et al [16]. The structure and content of the Cogthera intervention, including its implementation of metamemory-based strategies, have been previously described in detail. The program was designed to enhance key cognitive functions, including attention, imagery, and association, which support effective memory encoding and retrieval. Attention training helped participants focus and concentrate on target information to facilitate deeper encoding. Imagery training encouraged vivid visualization to reinforce memory consolidation and retrieval. Association training promoted the integration of new information into existing semantic memory networks. Throughout the training process, Cogthera provides personalized feedback and dynamically adjusts task difficulty, enabling users to observe and evaluate their cognitive processes independently.

A 15-month single-arm longitudinal study was conducted to investigate the long-term efficacy of mobile-based cognitive training. Participants completed a 3-month cognitive training program using Cogthera, with 9 (32%) of the 28 participants continuing an additional 12 months. The training program consisted of 2 daily sessions, 7 days per week, with each session lasting approximately 15 minutes. The first session comprised 3 core cognitive exercises targeting attention, imagery, and association. The second session included 4 additional exercises, selected from 9 available options, excluding those already used in the first session. A personalized algorithm determined daily exercise composition to optimize engagement. All participants were able to use the Cogthera program independently without external assistance. The detailed content of each exercise can be found in Multimedia Appendix 1.

Measures

Alzheimer's Disease Assessment Scale-Cognitive Subscale 14

The Alzheimer's Disease Assessment Scale-Cognitive Subscale 14 (ADAS-Cog 14) was used to assess the severity of cognitive dysfunction [25]. This measure consists of fourteen tasks: (1) word recall, (2) commands, (3) constructional praxis, (4) delayed word recall, (5) naming, (6) ideational praxis, (7) orientation, (8) word recognition, (9) maze, (10) number cancelation, (11) remembering instructions, (12) comprehension, (13) word finding difficulty, and (14) spoken language ability. The ADAS-Cog 14 subdomains were categorized into 3 cognitive domains: memory, language, and praxis [26]. The total ADAS-Cog 14 score ranges from 0 to 90 points, with higher scores indicating greater cognitive impairment, as the score reflects the number of errors made across tasks.

CERAD-NP Assessment

The CERAD-NP was administered to assess cognitive function [27,28]. This battery included the following neuropsychological tests: (1) verbal fluency, (2) Boston Naming Test, (3) Mini-Mental State Examination, (4) word list learning, (5) constructional praxis, (6) word list recall, (7) word list recognition, (8) constructional praxis recall, (9) Trail Making Test, and (10) Stroop test. The CERAD-NP total score was calculated as the sum of raw scores, with higher scores indicating better cognitive function.

CDR Scale

The CDR scale was used to stage dementia severity [29]. This scale is informed by semistructured interviews conducted with both participants and informants, covering 6 domains: memory, orientation, judgment and problem-solving, community affairs, home and hobbies, and personal care. Each domain is scored from 0 to 3, with the scores used in calculating the global CDR score and the CDR-Sum of Boxes (CDR-SB). The global CDR score is an ordinal scale ranked from 0 to 3 as follows: 0=no cognitive impairment, 0.5=questionable or MCI, 1=mild dementia, 2=moderate dementia, 3=severe dementia. The CDR-SB is a continuous measure ranging from 0 to 18, calculated by summing the individual domain scores. Higher scores on both the global CDR and CDR-SB indicate greater cognitive and functional impairment.

EQ-5D-5L Scale

The EQ-5D-5L was administered to evaluate health-related quality of life [30]. This measure consists of five dimensions: (1) mobility, (2) self-care, (3) usual activities, (4) pain and discomfort, and (5) anxiety and depression. Each dimension is rated on a 5-level scale, with response options ranging from *no problem* to *extreme problem*. Responses were converted into a single index score using a national-specific value set [31]. This value set, derived from stated preference data collected from the general population, assigns weights to each health dimension level. The index score ranges from 0 to 1, with higher scores indicating better overall health status. In contrast, a lower score on each subdimension reflects a higher level of health in that specific domain.

Data Collection

Participants who provided informed consent were screened using global CDR and CERAD-NP. During the initial visit, participants underwent cognitive function assessments using ADAS-Cog 14 and completed self-reported questionnaires, such as EQ-5D-5L. After completing these assessments, they were provided with a smartphone preinstalled with Cogthera. Follow-up data on ADAS-Cog 14 and EQ-5D-5L were collected after 3 months of using Cogthera. In addition, further assessments were conducted on 9 (32%) of the 28 participants who continued using Cogthera for 15 months. Training adherence was assessed using compliance, which was defined as the proportion of completed sessions relative to the total number of assigned sessions. The assigned frequency was 2 sessions per day, and compliance was calculated separately for the initial 3-month period and the 12-month extension.

Statistical Analysis

Linear mixed model analyses were performed to assess changes in ADAS-Cog 14 and EQ-5D-5L scores over time while accounting for both intraindividual and interindividual variations in longitudinal data. The models included time (baseline, 3 mo, and 15 mo) as a fixed factor and participant as a random effect, allowing individual variability in intercepts. To evaluate the overall effect of time, type Ⅲ ANOVA was conducted.

To identify explanatory variables predicting intervention effects, the least absolute shrinkage and selection operator (LASSO) regression was used. This approach was used to identify the most relevant predictors while minimizing overfitting [32]. The outcome variable was the change in ADAS-Cog 14 at follow-up, with LASSO models including age, sex, education, CDR-SB, baseline ADAS-Cog 14 scores, and baseline EQ-5D-5L scores as predictors. All variables were standardized by centering each variable around the mean and scaling by the SD. The regularization parameter, λ, was optimized using 9-fold cross-validation to balance model complexity and predictive performance. The optimal λ value was then applied to estimate the coefficients of the selected predictors.

To contextualize cognitive changes observed in this study, descriptive data from the placebo cohort of EXPEDITION studies were used as a historical reference [33,34]. EXPEDITION and EXPEDITION-2 were multicenter, double-blind, phase 3 trials of solanezumab, including 663 patients with MCI treated with a placebo. Cognitive function was assessed using ADAS-Cog 14 at baseline and at 6 follow-up points every 3 months over 18 months. Welch *t* test (2-tailed) was conducted to compare data from this study with the EXPEDITION placebo cohort at baseline, 3 months, and 15 months, accounting for unequal variances and sample sizes.

All statistical analyses were conducted using R software (version 4.4.1; R Foundation for Statistical Computing). Statistical significance was set at $P<.05$, with all tests being 2-tailed. For the post hoc analysis of the linear mixed model, Bonferroni-adjusted *P* values were reported. Effect sizes for the fixed time effects were also reported, with partial eta-squared (η^2_p) computed from the corresponding *F* statistics. Missing data were addressed through complete case analysis without imputation. To ensure analytic independence, all data access, monitoring, and statistical analyses were conducted exclusively by authors not affiliated with the company. Company-affiliated authors had no access to raw data and did not participate in data cleaning, analytic decisions, or interpretation of results.

Results

Demographic Characteristics

A total of 28 participants were included in the analysis (Table 1). The mean age of participants was 72.8 (SD 6.7) years, and the mean education level was 11.4 (SD 5.0) years. The sample comprised a higher proportion of female individuals (20/28, 71%) and was predominantly composed of nondrinkers (27/28, 96%) and nonsmokers (26/28, 93%). The mean CDR-SB score was 2.0 (SD 1.1), ranging from 0.5 to 4.0.

Table 1. Baseline demographic data for the participants (N=28).

Demographic characteristic	Values
Age (y), mean (SD)	72.8 (6.7)
Sex, n (%)	
Female	20 (71)
Male	8 (29)
Education (y), mean (SD)	11.4 (5.0)
Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Assessment Battery (*z* score[a]), mean (SD)	
Word list learning	–0.5 (0.9)
Word list recall	–1.6 (1.0)
Word list recognition	–1.7 (1.4)
Constructional praxis recall	–1.3 (0.9)
Clinical Dementia Rating-Sum of Boxes (score), mean (SD)	2.0 (1.1)

[a]A standardized score indicating how many SDs a value is from the mean.

Training Adherence

During the initial 3-month training period, participants completed an average of 85.3% (SD 23.6%) of assigned sessions. Among 9 (32%) of the 28 participants who continued for an additional 12 months, mean compliance during the extension period was 51.7% (SD 25.1%).

Intervention Effects of Cogthera

Linear mixed model analysis revealed a significant decrease in total ADAS-Cog 14 scores over time, indicating improved cognitive function ($F_{2,35.56}$=7.08; *P*=.003; η^2_p=0.28; Table 2). Specifically, a significant decrease in total ADAS-Cog 14 scores was observed over 15 months (Bonferroni-adjusted *P*=.001), while the change over 3 months was not significant (Bonferroni-adjusted *P*=.27; Table 3).

Table 2. Effects of time on cognitive function and quality of life.

	Baseline (n=28), mean (SD)	3 mo (n=28), mean (SD)	15 mo (n=9), mean (SD)	*F* test (*df*)	*P* value
Alzheimer's Disease Assessment Scale-Cognitive Subscale[a]					
Total score	28.18 (8.21)	27.07 (7.73)	25.56 (8.28)	7.08 (2, 35.56)	.003
Memory[b]	22.25 (6.24)	21.86 (6.29)	20.44 (8.56)	2.53 (2, 35.83)	.09
Language[c]	0.86 (0.93)	0.64 (0.83)	0.56 (0.73)	1.64 (2, 35.83)	.21
Praxis[d]	2.32 (1.19)	1.96 (1.10)	1.33 (1.00)	7.11 (2, 39.33)	.002
EQ-5D-5L					
Index score	0.81 (0.08)	0.85 (0.04)	0.84 (0.07)	3.40 (2, 42.14)	.04

[a]Lower scores represent better performance.

[b]Sum of word recall, delayed word recall, orientation, and word recognition task scores.

[c]Sum of naming, remembering instructions, comprehension, word finding difficulty, and spoken language ability task scores.

[d]Sum of commands, constructional praxis, and ideational praxis task scores.

Table 3. Longitudinal changes in cognitive function and quality of life.

	Estimate (SE)	*t* test (*df*)	*P* value[a]
Alzheimer's Disease Assessment Scale-Cognitive Subscale			
Total score			
3 months vs baseline	–1.11 (0.72)	–1.53 (35.06)	.27
15 months vs baseline	–4.33 (1.16)	–3.74 (35.96)	.001
Memory			
3 months vs baseline	–0.39 (0.68)	–0.58 (35.17)	.99
15 months vs baseline	–2.42 (1.08)	–2.25 (36.38)	.06
Language			
3 months vs baseline	–0.21 (0.12)	–1.81 (34.65)	.16
15 months vs baseline	–0.09 (0.19)	–0.47 (36.83)	.99
Praxis			
3 months vs baseline	–0.36 (0.23)	–1.53 (35.78)	.27
15 months vs baseline	–1.35 (0.36)	–3.76 (42.63)	.001
EQ-5D-5L			
Index score			
3 months vs baseline	0.04 (0.02)	2.58 (35.29)	.03
15 months vs baseline	0.03 (0.02)	1.27 (49.19)	.42

[a]Bonferroni-adjusted *P* values.

In the ADAS-Cog 14 subdomain analysis, praxis scores, representing executive function, showed a significant decline over time ($F_{2,39.33}$=7.11; *P*=.002; η^2_p=0.27), particularly over 15 months (Bonferroni-adjusted *P*=.001). The reduction in memory scores was marginal ($F_{2,35.83}$=2.53; *P*=.09; η^2_p=0.12), with a trend-level decrease observed at 15 months (Bonferroni-adjusted *P*=.06). The overall effect of time on language scores was not significant.

A significant change in the EQ-5D-5L index score was observed over time ($F_{2,42.14}$=3.40; *P*=.04; η^2_p=0.14), with a statistically significant increase from baseline to 3 months (Bonferroni-adjusted *P*=.03). Among the EQ-5D-5L subdimensions, mobility showed a marginal overall effect of time ($F_{2,\ 39.05}$=2.59; *P*=.09; η^2_p=0.12), with a trend-level decrease observed at 3 months (Bonferroni-adjusted *P*=.06). Detailed results for the EQ-5D-5L subdimensions are provided in Tables S1 and S2 in Multimedia Appendix 2).

Factors Associated With Intervention Effects on Cognitive Function

According to the LASSO results, CDR-SB and baseline ADAS-Cog 14 scores predicted score changes over 3 months, while age, CDR-SB, and baseline ADAS-Cog 14 scores predicted score changes over 15 months. The multiple linear regression model, which included sex, CDR-SB, and baseline ADAS-Cog 14 scores, explained a significant proportion of the variance in score changes over 3 months (R^2=0.40; adjusted R^2=0.33; *P*=.01; Table 4). Specifically, lower CDR-SB scores (β=–0.51; *P*=.02), and higher baseline ADAS-Cog 14 scores significantly predicted greater improvements (β=0.61; *P*=.01).

Table 4. Factors predicting cognitive function improvements.

	Values, estimate (SE)	β	*t* test *(df)*	R^2 (adjusted R^2)	*P* value
Change over 3 months[a]				0.40 (0.33)	.01
Sex	2.60 (1.32)	0.69	1.97		.06
CDR-SB[b]	–1.78 (0.70)	–0.51	–2.53		.02
ADAS-Cog 14[c]	0.28 (0.09)	0.61	3.01		.01
Change over 15 months[d]				0.81 (0.70)	.03
Age	–0.44 (0.14)	–0.62	–3.23		.02
CDR-SB	–4.80 (1.39)	–1.10	–3.47		.02
ADAS-Cog 14	0.61 (0.14)	1.05	4.25		.01

[a]Score obtained by subtracting the Alzheimer's Disease Assessment Scale-Cognitive Subscale score at 3 months from the baseline Alzheimer's Disease Assessment Scale-Cognitive Subscale score.

[b]CDR-SB: Clinical Dementia Rating-Sum of Boxes.

[c]ADAS-Cog 14: Alzheimer's Disease Assessment Scale-Cognitive Subscale.

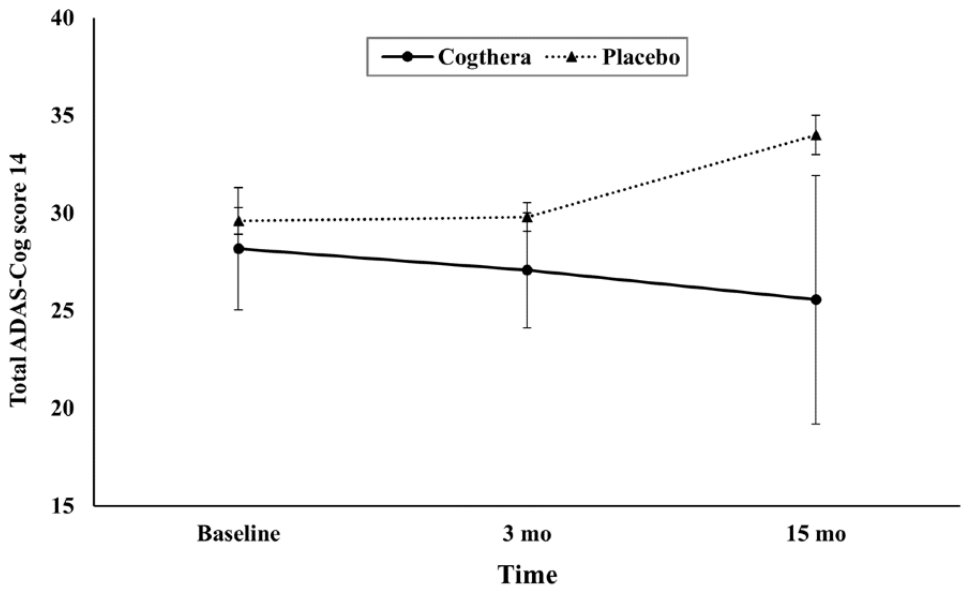
[d]Score obtained by subtracting the Alzheimer's Disease Assessment Scale-Cognitive Subscale score at 15 months from the baseline Alzheimer's Disease Assessment Scale-Cognitive Subscale score.

The regression model, which included age, CDR-SB, and baseline ADAS-Cog 14 scores, accounted for a significant proportion of the variance in score changes observed over 15 months (R^2=0.81; adjusted R^2=0.70; *P*=.03). Specifically, younger age (β=–0.62; *P*=.02), lower CDR-SB scores (β=–1.10; *P*=.02), and higher baseline ADAS-Cog 14 scores significantly predicted greater improvements (β=1.05; *P*=.01).

Comparison With Historical Placebo Data

The mean and SD data from a previously reported placebo cohort were used for comparison at each time point. No significant differences were observed between the Cogthera group and the placebo group at baseline ($t_{29.68}$=0.89; *P*=.38; Figure 1). At 3 months, ADAS-Cog 14 scores in the Cogthera group were marginally lower than those in the placebo group ($t_{30.63}$=1.81; *P*=.08). By 15 months, the difference became statistically significant, suggesting greater cognitive improvement in the Cogthera group ($t_{8.55}$=3.01; *P*=.02).

Figure 1. Comparison of cognitive function between the Cogthera group and the historical placebo group. Error bars represent 95% CIs of the mean Alzheimer's Disease Assessment Scale-Cognitive Subscale 14 (ADAS-Cog 14) scores at each time point.



Discussion

Principal Findings

This study examined the long-term effects of a mobile-based cognitive training app in older adults with MCI. ADAS-Cog 14 scores decreased over time, and EQ-5D-5L scores increased in the early phase of training. Age, CDR-SB, and baseline ADAS-Cog 14 scores were associated with the degree of cognitive change. Although comparisons with a historical placebo cohort are limited by differences in design and sample characteristics, the observed reduction in ADAS-Cog 14 scores over 15 months suggests a pattern of cognitive change associated with long-term training use.

The primary finding of this study is the reduction in ADAS-Cog 14 total, memory, and praxis scores associated with long-term use of Cogthera. While no meaningful change was observed at 3 months, a clearer reduction emerged at 15 months. This delayed pattern may reflect the cumulative effects of training, but it could also be influenced by practice effects, familiarity with test procedures, or natural fluctuations in cognitive performance. These possibilities underscore the need for cautious interpretation, particularly in the absence of a control group. Even so, the overall pattern is consistent with previous studies, suggesting that longer training durations are more likely to yield detectable cognitive improvements [35]. Unlike other ADAS-Cog 14 subdomains, language scores did not show meaningful change. The small numerical variation observed is likely to reflect measurement variability or stabilization rather than the true effect of the intervention. This aligns with previous research indicating that language tends to remain relatively preserved in the early stage of cognitive decline compared with other ADAS-Cog 14 subdomains [36,37].

Another finding is a rise in the EQ-5D-5L index score during the first 3 months of training. This early change is consistent with previous studies reporting short-term benefits in quality of life among individuals with MCI, but the effect did not persist at 15 months in this study [38-41]. This pattern suggests that changes in perceived well-being may be limited to the initial phase of training. Such early gains may also reflect nonspecific factors, including placebo effects or heightened engagement at the beginning of training.

Predictor analyses revealed that both short- and long-term improvements in cognitive function were associated with lower baseline cognitive function and higher functional status. Participants with higher baseline ADAS-Cog 14 scores, which indicate poorer initial cognitive performance, and those with lower CDR-SB scores, which indicate less functional impairment, tended to show larger reductions over time. Previous research has reported conflicting findings regarding the relationship between baseline cognitive performance and cognitive training effectiveness [42]. Two competing hypotheses have been proposed to explain this relationship: the compensation effect and the magnification effect. The compensation effect suggests that individuals with higher baseline cognitive performance benefit less from training due to limited room for improvement. In contrast, the magnification effect proposes that they benefit more by leveraging greater cognitive efficiency and plasticity. The findings of this study align with the compensation effect, as participants with lower baseline cognitive performance showed greater training-related gains. This pattern suggests that Cogthera may have provided compensatory cognitive stimulation. Such stimulation could have helped individuals with lower baseline cognitive performance make better use of their remaining cognitive resources. CDR-SB reflects the pathological progression of cognitive and functional decline and reliably differentiates MCI from very early AD while predicting future progression [43-45]. The association between lower CDR-SB scores and greater cognitive gains may suggest that individuals with preserved daily function but greater cognitive vulnerability were more responsive to training. These observations are broadly consistent with compensatory mechanisms, while other explanations remain possible.

Younger age was associated with modestly better cognitive outcomes over the 15-month period. One possible explanation is that younger participants may engage more consistently in long-term digital interventions. Previous research has shown that they tend to complete more training sessions and demonstrate higher adherence. Such patterns of adherence may help account for the cognitive changes observed in this group [46-48].

These findings suggest that MMT delivered through a mobile app may offer a scalable and accessible approach for individuals with MCI. Mobile apps provide a practical alternative to traditional cognitive training by allowing users to engage in cognitive exercises at their convenience. As memory decline has already begun in individuals with MCI, this approach may serve as a supportive digital complement to existing strategies. A potential strength of this study is the observation that cognitive changes became more evident with prolonged use, which may help inform future work on long-term digital training. Given the lack of known adverse effects, mobile-based cognitive training could represent a feasible and well-tolerated adjunctive approach for older adults with MCI. It may be particularly useful in cases where pharmacological treatments carry a higher risk of side effects.

Limitations

Despite the observational insights provided by this study, several limitations warrant careful consideration. First, the absence of a control group limits internal validity, as it is not possible to distinguish training-related changes from those that may occur naturally over time. Without contemporaneous controls, causal inferences cannot be drawn.

Second, although 28 participants initiated the study, only 9 (32%) continued to the 15-month assessment. Because long-term participation was voluntary, those who remained may differ in motivation or engagement. However, baseline comparisons between participants who continued to 15 months and those who participated only in the initial 3-month period showed no significant differences in demographic characteristics, cognitive measures, quality of life scores, or training adherence, as presented in Table S3 in Multimedia Appendix 2. In addition, baseline-to–3-month change scores did not differ between the 2 groups, indicating that early improvement was not greater among those who remained in the study. These findings suggest that systematic baseline differences were limited, although unmeasured factors related to motivation or persistence may still have influenced long-term retention. The small number of long-term participants also restricts the generalizability of the 15-month findings.

Finally, the overall sample size was small, constraining statistical power and limiting the ability to detect subtle effects. Future studies with larger, randomized, and more diverse samples will be essential to determine whether long-term mobile-based cognitive training produces reliable and sustained effects in individuals with MCI.

Conclusions

This study suggests that a mobile-based MMT app may offer supportive benefits for cognitive function in older adults with MCI. Although methodological limitations constrain definitive interpretation, the observed patterns indicate that mobile-based cognitive training could be a feasible and accessible approach for individuals seeking strategies to maintain cognitive health. Continued evaluation in larger, controlled studies will help determine the extent and durability of these potential effects.

Funding

This work was supported by the Technology Innovation Program (or Industrial Strategic Technology Development Program—the bioindustry technology development; grant 20018143, development and clinical application of digital therapeutic devices user experience, untact clinical cloud platforms, and real-world evidence technologies based on the brain and body of the elderly) funded by the Ministry of Trade, Industry, and Energy (South Korea).

Data Availability

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

Conflicts of Interest

YB is employed by Emocog Inc. SK was employed in the R&D department of Emocog Inc. during the conduct of the study and remained employed until August 31, 2025. JYL is the chief executive officer of Emocog Inc. LS is the chief operating officer of Cogthera GmbH, a subsidiary of Emocog Inc. Emocog Inc owns the app evaluated in this study and has a commercial interest in its future distribution. Cogthera GmbH is responsible for the commercialization and planned distribution of a modified version of this app in Germany. However, YB, SK, JYL, and LS were not involved in the data collection or analysis, which was conducted independently by other authors. All other authors declare no conflicts of interest.

Multimedia Appendix 1

Description of the cognitive exercises included in Cogthera.

[XLSX File (Microsoft Excel File), 11 KB - aging_v9i1e81648_app1.xlsx]

Multimedia Appendix 2

Supplementary EQ-5D-5L subdimension results and baseline user comparisons.

[DOCX File , 29 KB - aging_v9i1e81648_app2.docx]

References

1. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Arch Neurol 1999 Mar 01;56(3):303-308. [doi: 10.1001/archneur.56.3.303] [Medline: 10190820]
2. Petersen RC, Lopez O, Armstrong MJ, Getchius TS, Ganguli M, Gloss D, et al. Practice guideline update summary: mild cognitive impairment [RETIRED]: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 2018 Jan 16;90(3):126-135 [FREE Full text] [doi: 10.1212/WNL.0000000000004826] [Medline: 29282327]
3. Mitchell AJ, Shiri-Feshki M. Rate of progression of mild cognitive impairment to dementia--meta-analysis of 41 robust inception cohort studies. Acta Psychiatr Scand 2009 Apr 03;119(4):252-265. [doi: 10.1111/j.1600-0447.2008.01326.x] [Medline: 19236314]
4. Baker LD, Frank LL, Foster-Schubert K, Green PS, Wilkinson CW, McTiernan A, et al. Effects of aerobic exercise on mild cognitive impairment: a controlled trial. Arch Neurol 2010 Jan 01;67(1):71-79 [FREE Full text] [doi: 10.1001/archneurol.2009.307] [Medline: 20065132]
5. Suzuki T, Shimada H, Makizako H, Doi T, Yoshida D, Tsutsumimoto K, et al. Effects of multicomponent exercise on cognitive function in older adults with amnestic mild cognitive impairment: a randomized controlled trial. BMC Neurol 2012 Oct 31;12:128 [FREE Full text] [doi: 10.1186/1471-2377-12-128] [Medline: 23113898]
6. Li H, Li J, Li N, Li B, Wang P, Zhou T. Cognitive intervention for persons with mild cognitive impairment: a meta-analysis. Ageing Res Rev 2011 Apr;10(2):285-296. [doi: 10.1016/j.arr.2010.11.003] [Medline: 21130185]

7. Hughes TF, Flatt JD, Fu B, Chang CC, Ganguli M. Engagement in social activities and progression from mild to severe cognitive impairment: the MYHAT study. Int Psychogeriatr 2013 Apr;25(4):587-595 [FREE Full text] [doi: 10.1017/S1041610212002086] [Medline: 23257280]
8. Gavelin HM, Dong C, Minkov R, Bahar-Fuchs A, Ellis KA, Lautenschlager NT, et al. Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. Ageing Res Rev 2021 Mar;66:101232. [doi: 10.1016/j.arr.2020.101232] [Medline: 33249177]
9. Intzandt B, Vrinceanu T, Huck J, Vincent T, Montero-Odasso M, Gauthier CJ, et al. Comparing the effect of cognitive vs. exercise training on brain MRI outcomes in healthy older adults: a systematic review. Neurosci Biobehav Rev 2021 Sep;128:511-533. [doi: 10.1016/j.neubiorev.2021.07.003] [Medline: 34245760]
10. Raimo S, Cropano M, Gaita M, Maggi G, Cavallo ND, Roldan-Tapia MD, et al. The efficacy of cognitive training on neuropsychological outcomes in mild cognitive impairment: a meta-analysis. Brain Sci 2023 Oct 25;13(11):1510 [FREE Full text] [doi: 10.3390/brainsci13111510] [Medline: 38002471]
11. Hampstead BM, Stringer AY, Iordan AD, Ploutz-Snyder R, Sathian K. Toward rational use of cognitive training in those with mild cognitive impairment. Alzheimers Dement 2023 Mar 06;19(3):933-945 [FREE Full text] [doi: 10.1002/alz.12718] [Medline: 35791724]
12. Belleville S, Cuesta M, Bier N, Brodeur C, Gauthier S, Gilbert B, et al. Five-year effects of cognitive training in individuals with mild cognitive impairment. Alzheimers Dement (Amst) 2024 Sep 06;16(3):e12626. [doi: 10.1002/dad2.12626] [Medline: 39246830]
13. Chan AT, Ip RT, Tran JY, Chan JY, Tsoi KK. Computerized cognitive training for memory functions in mild cognitive impairment or dementia: a systematic review and meta-analysis. NPJ Digit Med 2024 Jan 03;7(1):1 [FREE Full text] [doi: 10.1038/s41746-023-00987-5] [Medline: 38172429]
14. Rossetto F, Isernia S, Realdon O, Borgnis F, Blasi V, Pagliari C, et al. A digital health home intervention for people within the Alzheimer's disease continuum: results from the Ability-TelerehABILITation pilot randomized controlled trial. Ann Med 2023 Dec 17;55(1):1080-1091 [FREE Full text] [doi: 10.1080/07853890.2023.2185672] [Medline: 36929703]
15. Youn JH, Ryu SH, Lee JY, Park S, Cho SJ, Kwon H, et al. Brain structural changes after multi-strategic metamemory training in older adults with subjective memory complaints: a randomized controlled trial. Brain Behav 2019 May 27;9(5):e01278 [FREE Full text] [doi: 10.1002/brb3.1278] [Medline: 30916450]
16. Youn JH, Park S, Lee JY, Cho SJ, Kim J, Ryu SH. Cognitive improvement in older adults with mild cognitive impairment: evidence from a multi-strategic metamemory training. J Clin Med 2020 Jan 28;9(2):362 [FREE Full text] [doi: 10.3390/jcm9020362] [Medline: 32013035]
17. McDougall GJ, Kang J. Memory self-efficacy and memory performance in older males. Int J Mens Health 2003 May 1;2(2):131-147 [FREE Full text] [doi: 10.3149/jmh.0202.131] [Medline: 19043600]
18. Flavell JH. Metacognition and cognitive monitoring: a new area of cognitive–developmental inquiry. Am Psychol 1979 Oct;34(10):906-911. [doi: 10.1037/0003-066x.34.10.906]
19. Tarricone P. The Taxonomy of Metacognition. London, UK: Psychology Press; 2011.
20. Kim J, Shin E, Han K, Park S, Youn JH, Jin G, et al. Efficacy of smart speaker-based metamemory training in older adults: case-control cohort study. J Med Internet Res 2021 Feb 16;23(2):e20177 [FREE Full text] [doi: 10.2196/20177] [Medline: 33591276]
21. Sella E, Carbone E, Vincenzi M, Toffalini E, Borella E. Efficacy of memory training interventions targeting metacognition for older adults: a systematic review and meta-analysis. Aging Ment Health 2023 Apr 11;27(4):674-694. [doi: 10.1080/13607863.2022.2122931] [Medline: 36218025]
22. Youn JH, Lee JY, Kim S, Ryu SH. Multistrategic memory training with the metamemory concept in healthy older adults. Psychiatry Investig 2011 Dec;8(4):354-361 [FREE Full text] [doi: 10.4306/pi.2011.8.4.354] [Medline: 22216046]
23. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med 2004 Sep 20;256(3):183-194 [FREE Full text] [doi: 10.1111/j.1365-2796.2004.01388.x] [Medline: 15324362]
24. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer’s disease: recommendations from the National Institute on Aging-Alzheimer’s Association Workgroups on Diagnostic Guidelines for Alzheimer’s Disease. Focus 2013 Jan 1;11(1):96-106. [doi: 10.1176/appi.focus.11.1.96]
25. Mohs RC, Knopman D, Petersen RC, Ferris SH, Ernesto C, Grundman M, et al. Development of cognitive instruments for use in clinical trials of antidementia drugs: additions to the Alzheimer's Disease Assessment Scale that broaden its scope. The Alzheimer's Disease Cooperative Study. Alzheimer Dis Assoc Disord 1997;11 Suppl 2:S13-S21. [Medline: 9236948]
26. Verma N, Beretvas SN, Pascual B, Masdeu JC, Markey MK, Alzheimer’s Disease Neuroimaging Initiative. New scoring methodology improves the sensitivity of the Alzheimer's Disease Assessment Scale-Cognitive subscale (ADAS-Cog) in clinical trials. Alzheimers Res Ther 2015 Nov 12;7(1):64 [FREE Full text] [doi: 10.1186/s13195-015-0151-0] [Medline: 26560146]
27. Morris JC, Heyman A, Mohs RC, Hughes JP, van Belle G, Fillenbaum G, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 1989 Sep;39(9):1159-1165. [doi: 10.1212/wnl.39.9.1159] [Medline: 2771064]

28. Lee JH, Lee KU, Lee DY, Kim KW, Jhoo JH, Kim JH, et al. Development of the Korean version of the Consortium to Establish a Registry for Alzheimer's Disease Assessment Packet (CERAD-K): clinical and neuropsychological assessment batteries. J Gerontol B Psychol Sci Soc Sci 2002 Jan 01;57(1):P47-P53. [doi: 10.1093/geronb/57.1.p47] [Medline: 11773223]
29. Morris JC. The Clinical Dementia Rating (CDR): current version and scoring rules. Neurology 1993 Nov;43(11):2412-2414. [doi: 10.1212/wnl.43.11.2412-a] [Medline: 8232972]
30. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec 9;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
31. Kim SH, Ahn J, Ock M, Shin S, Park J, Luo N, et al. The EQ-5D-5L valuation study in Korea. Qual Life Res 2016 Jul 10;25(7):1845-1852. [doi: 10.1007/s11136-015-1205-2] [Medline: 26961008]
32. Hassan MM, Hassan MM, Yasmin F, Khan MA, Zaman S, Galibuzzaman, et al. A comparative assessment of machine learning algorithms with the Least Absolute Shrinkage and Selection Operator for breast cancer detection and prediction. Decis Anal J 2023 Jun;7:100245. [doi: 10.1016/j.dajour.2023.100245]
33. Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B, Joffe S, et al. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. N Engl J Med 2014 Jan 23;370(4):311-321. [doi: 10.1056/nejmoa1312889]
34. Liu-Seifert H, Siemers E, Price K, Han B, Selzler KJ, Henley D, et al. Cognitive impairment precedes and predicts functional impairment in mild Alzheimer's disease. J Alzheimers Dis 2015;47(1):205-214 [FREE Full text] [doi: 10.3233/JAD-142508] [Medline: 26402769]
35. Eggenberger P, Schumacher V, Angst M, Theill N, de Bruin ED. Does multicomponent physical exercise with simultaneous cognitive training boost cognitive performance in older adults? A 6-month randomized controlled trial with a 1-year follow-up. Clin Interv Aging 2015;10:1335-1349 [FREE Full text] [doi: 10.2147/CIA.S87732] [Medline: 26316729]
36. Raghavan N, Samtani MN, Farnum M, Yang E, Novak G, Grundman M, et al. The ADAS-Cog revisited: novel composite scales based on ADAS-Cog to improve efficiency in MCI and early AD trials. Alzheimers Dement 2013 Feb 05;9(1 Suppl):S21-S31 [FREE Full text] [doi: 10.1016/j.jalz.2012.05.2187] [Medline: 23127469]
37. Cogo-Moreira H, Krance SH, Wu CY, Lanctôt KL, Herrmann N, Black SE, et al. State, trait, and accumulated features of the Alzheimer's Disease Assessment Scale Cognitive Subscale (ADAS-Cog) in mild Alzheimer's disease. Alzheimers Dement (N Y) 2023 Mar 26;9(1):e12376 [FREE Full text] [doi: 10.1002/trc2.12376] [Medline: 36994227]
38. Maki Y, Yamaguchi T, Yamagami T, Murai T, Hachisuka K, Miyamae F, et al. The impact of subjective memory complaints on quality of life in community-dwelling older adults. Psychogeriatrics 2014 Sep 21;14(3):175-181. [doi: 10.1111/psyg.12056] [Medline: 25142381]
39. Giuli C, Papa R, Lattanzio F, Postacchini D. The effects of cognitive training for elderly: results from my mind project. Rejuvenation Res 2016 Dec;19(6):485-494 [FREE Full text] [doi: 10.1089/rej.2015.1791] [Medline: 26952713]
40. Han JW, Lee H, Hong JW, Kim K, Kim T, Byun HJ, et al. Multimodal cognitive enhancement therapy for patients with mild cognitive impairment and mild dementia: a multi- center, randomized, controlled, double-blind, crossover trial. J Alzheimers Dis 2017;55(2):787-796. [doi: 10.3233/JAD-160619] [Medline: 27802233]
41. Fu E, Compton M. Non-pharmacological interventions that improve quality of life in people with mild cognitive impairment. Biomed J Sci Tech Res 2021 Oct 28;39(4):31543-31547 [FREE Full text] [doi: 10.26717/BJSTR.2021.39.006338]
42. Shaw J, Hosseini SH. The effect of baseline performance and age on cognitive training improvements in older adults: a qualitative review. J Prev Alzheimers Dis 2021 Jan;8(1):100-109 [FREE Full text] [doi: 10.14283/jpad.2020.55] [Medline: 33336231]
43. Perneczky R, Hartmann J, Grimmer T, Drzezga A, Kurz A. Cerebral metabolic correlates of the clinical dementia rating scale in mild cognitive impairment. J Geriatr Psychiatry Neurol 2007 Jun 01;20(2):84-88. [doi: 10.1177/0891988706297093] [Medline: 17548777]
44. O'Bryant SE, Lacritz LH, Hall J, Waring SC, Chan W, Khodr ZG, et al. Validation of the new interpretive guidelines for the clinical dementia rating scale sum of boxes score in the national Alzheimer's coordinating center database. Arch Neurol 2010 Jun 01;67(6):746-749 [FREE Full text] [doi: 10.1001/archneurol.2010.115] [Medline: 20558394]
45. Tzeng RC, Yang YW, Hsu KC, Chang HT, Chiu PY. Sum of boxes of the clinical dementia rating scale highly predicts conversion or reversion in predementia stages. Front Aging Neurosci 2022 Sep 23;14:1021792 [FREE Full text] [doi: 10.3389/fnagi.2022.1021792] [Medline: 36212036]
46. Vaghefi I, Tulu B. The continued use of mobile health apps: insights from a longitudinal study. JMIR Mhealth Uhealth 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]
47. Belleville S, Gilbert B, Fontaine F, Gagnon L, Ménard E, Gauthier S. Improvement of episodic memory in persons with mild cognitive impairment and healthy older adults: evidence from a cognitive intervention program. Dement Geriatr Cogn Disord 2006 Nov 3;22(5-6):486-499. [doi: 10.1159/000096316] [Medline: 17050952]
48. Turunen M, Hokkanen L, Bäckman L, Stigsdotter-Neely A, Hänninen T, Paajanen T, et al. Computer-based cognitive training for older adults: determinants of adherence. PLoS One 2019 Jul 10;14(7):e0219541 [FREE Full text] [doi: 10.1371/journal.pone.0219541] [Medline: 31291337]

Abbreviations

AD: Alzheimer disease
ADAS-Cog 14: Alzheimer's Disease Assessment Scale-Cognitive Subscale 14
CDR: Clinical Dementia Rating
CDR-SB: Clinical Dementia Rating-Sum of Boxes
CERAD-NP: Consortium to Establish a Registry for Alzheimer's Disease Neuropsychological Assessment Battery
LASSO: least absolute shrinkage and selection operator
MCI: mild cognitive impairment
MMT: metamemory training

Edited by M O'Connell; submitted 31.Jul.2025; peer-reviewed by AS Rigaud, P Kooncumchoo; comments to author 19.Oct.2025; accepted 05.Dec.2025; published 02.Jan.2026.

Please cite as:
Lim JI, Byeon Y, Kang S, Kim H, Kim KY, Stenzel L, Jeon SY, Lee JY
Long-Term Effects of Mobile-Based Metamemory Cognitive Training in Older Adults With Mild Cognitive Impairment: 15-Month Prospective Single-Arm Longitudinal Study
JMIR Aging 2026;9:e81648
URL: https://aging.jmir.org/2026/1/e81648
doi:10.2196/81648
PMID:

Impact of 4 Weeks or More Immersive Virtual Reality on Quality of Life and Physical Activity in Older Adults: Systematic Review and Meta-Analysis

Iria Trillo-Charlín[1,2], MSc; Javier Bravo-Aparicio[2,3], MSc; Juan Avendaño-Coy[1,2], PhD; Héctor Beltrán-Alacreu[2,3], PhD

[1]Toledo Physiotherapy Research Group (GIFTO), Faculty of Physical Therapy and Nursing, University of Castilla-La Mancha, Toledo, Spain

[2]Toledo Physiotherapy Research Group (GIFTO), Castilla-La Mancha Health Research Institute (IDISCAM), Toledo, Spain

[3]Pain, Mental Health, Exercise and Technology Research Group (PAIN+MET), Faculty of Physical Therapy and Nursing, University of Castilla-La Mancha, Avenue de Carlos III, 45004 Toledo, Toledo, Spain

Corresponding Author:
Javier Bravo-Aparicio, MSc
Toledo Physiotherapy Research Group (GIFTO), Castilla-La Mancha Health Research Institute (IDISCAM), Toledo, Spain

Abstract

Background: Population aging poses significant public health challenges. Older adults often face multimorbidity, functional decline, and diminished quality of life. While physical activity can mitigate these effects, adherence remains low. Immersive virtual reality (IVR) has emerged as a promising, engaging tool to promote physical and cognitive health in this population.

Objective: The review aims to evaluate the effectiveness of IVR interventions lasting 4 weeks or more on quality of life, physical activity, pain, perceived effort, and adverse events in older adults.

Methods: A systematic review and meta-analysis were conducted following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and Cochrane guidelines. Literature was searched across PubMed, Web of Science, PEDro, and Scopus, as well as sources of gray literature. Eligible studies included randomized controlled trials involving participants aged >60 years, using IVR via head-mounted display. Outcomes assessed included quality of life, physical activity, pain, perceived effort, and adverse events. Risk of bias and evidence certainty were assessed using Risk of Bias 2.0 and GRADE (Grades of Recommendation Assessment Development and Evaluation), respectively.

Results: A total of 14 studies with 839 participants were included in the qualitative synthesis, of which 8 were eligible for quantitative meta-analysis. The pooled analysis showed a statistically significant moderate effect of IVR on quality of life (standardized mean difference [SMD]=0.48, 95% CI 0.1-0.8; *P*=.007), particularly in interventions lasting 10 to 12 weeks or involving more than 600 minutes of exposure. For physical activity, no significant differences were found between IVR and control groups (SMD=–0.2, 95% CI –0.7 to 0.4; *P*=.50). Evidence for secondary outcomes (pain, perceived exertion, and adverse events) was limited and largely qualitative, with inconsistent findings. Pain outcomes, assessed in 2 studies, indicated reductions in the IVR group, especially when multimodal approaches were used. Perceived effort was not systematically measured. Adverse events were generally mild, with cybersickness being the most reported issue.

Conclusions: IVR interventions of 4 weeks or more appear to moderately improve quality of life in older adults, especially those with clinical vulnerabilities or in institutional settings. Although effects on physical activity were not significant, trends suggest potential with appropriate program design. Preliminary findings support IVR's use in pain reduction, particularly when incorporating emotional and multisensory elements. The low incidence of adverse events suggests good tolerability. Overall, IVR is a promising and safe tool to support healthy aging, though further high-quality studies are needed to confirm these findings and assess long-term outcomes.

(JMIR Aging 2026;9:e80820) doi:10.2196/80820

KEYWORDS

adverse events; immersive virtual reality; older adults; pain; perceived effort; physical activity; PRISMA; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; quality of life

Introduction

By 2030, the population of adults aged ≥65 years is expected to reach 994 million (12% worldwide), intensifying pressures on public-health systems [1]. From a biological perspective, aging is defined as a complex and progressive process that affects multiple systems, leading to a decline in functional capacity and increased vulnerability to various pathologies [2]. Aging is often accompanied by impairments such as loss of muscle strength, cognitive decline, and a heightened risk of falls and disability, all of which negatively impact the autonomy and quality of life of older adults [2-4].

In this context, it is crucial to identify effective strategies that support health and well-being in later life, with physical activity being one of the fundamental pillars for promoting healthy aging [5]. In older adults, physical activity is associated with better physical function, preservation of cognition, and enhanced quality of life. It is also linked to lower disability in activities of daily living, fewer falls, and a reduced risk of neurodegenerative diseases such as dementia [6-9]. Moreover, it has been linked to decreased mortality rates, and it is a leading modifiable determinant of healthy aging with a dose–response association to all-cause and cardiovascular mortality in older adults [10,11]. However, despite the available evidence, the proportion of older adults who meet physical activity recommendations remains suboptimal [11]. Several factors contribute to this, including lack of awareness of the benefits, fear of pain or falling, low motivation, and environmental barriers [12].

In this regard, immersive virtual reality (IVR) emerges as a potential nonpharmacological therapeutic alternative. Through 3D-simulated environments experienced via devices known as head-mounted displays, IVR offers an immersive experience in which users can interact with the virtual environment [13]. IVR's sensorimotor immersion and playful elements may support engagement with physical activity, but current evidence is exploratory and based on a small feasibility trial [14]. Because IVR may enhance motivation and engagement, it is clinically and pragmatically important to test whether multiweek IVR programs can change physical activity behavior. Accordingly, our protocol designated physical activity as a coprimary outcome [15]. Additionally, its design allows for the adjustment of difficulty levels, making it a customizable alternative [16,17]—a feature considered essential in exercise prescription [18].

On the other hand, IVR has also proven useful in interventions aimed at cognitive stimulation and reminiscence, thereby expanding its range of applications in the context of active aging [19,20].

Currently, evidence suggests that IVR may benefit balance, mobility, cognition, and psychological well-being. However, some trials have reported mixed or modest effects [20-23]. These findings highlight the need for standardized outcome measures and adequately dosed programs in future research. These interventions have been implemented in both healthy individuals [21,24] and those with chronic medical conditions [22] or mild cognitive impairment [20]. However, existing reviews often involve short intervention periods [25]. This limits the ability to observe sustained long-term effects and may inflate immediate benefits [15,21]. Therefore, the focus is placed on programs delivered over periods that map onto the recall horizons of core quality-of-life instruments, a window that also coincides with the early emergence of neural and motor adaptations in response to training [26-30]. Trials in older adults have reported measurable gains over such intervals—cognition, balance, and functional outcomes [31-34]. Importantly, centering analyses on these exposure windows reduces the susceptibility of very brief interventions to novelty and Hawthorne effects [35]. In line with the World Health Organization's (WHO) healthy aging agenda, which prioritizes functional capacity and well-being as central goals in aging societies, we selected quality of life and physical activity as primary outcomes. These measures are patient-centered indicators of healthy aging and modifiable behavior, making them particularly relevant assessment criteria [36].

Prior reviews in older adults have primarily established that virtual reality (VR) is acceptable and feasible, while noting that evidence for effectiveness remains limited [15]. Other syntheses have focused on exergames and outcomes such as balance in long-term care facilities, aggregating heterogeneous exposure durations and without a prespecified minimum dose [37]. More recent randomized controlled trial (RCT)-only reviews of head-mounted display VR suggest benefits for physical activity and broader well-being, but do not isolate ≥4-week programs as an a priori inclusion criterion [38]. To our knowledge, no systematic review has focused exclusively on IVR interventions in older adults with a minimum duration of ≥4 weeks as an a priori inclusion criterion.

The objective of this systematic review and meta-analysis is to evaluate the effectiveness of IVR interventions lasting 4 weeks or more on the quality of life, physical activity, pain, perceived effort, and adverse events in older adults.

Methods

The protocol for this systematic review and meta-analysis was registered in the International Prospective Register of Systematic Reviews (CRD420251019170). It was conducted following the recommendations of the Cochrane Collaboration and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Checklist 1) [39]

Search Strategy

A bibliographic search was completed between March 15, 2025, and June 15, 2025, in all the following databases: PubMed, Web of Science, PEDro, and Scopus. Additionally, we systematically searched for gray literature. This included searches in ClinicalTrials.gov, the WHO International Clinical Trials Registry Platform, ProQuest Dissertations & Theses Global, and medRxiv. Search strategies used are available in Multimedia Appendix 1 [14,18,24,29,40-53].

We also carried out a "snowball" search to identify additional studies by searching the reference lists of publications eligible for full-text review.

Inclusion and Exclusion Criteria

Inclusion criteria were defined according to the PICOS (population, intervention, comparison, outcome, study design) framework: (1) we included studies only if all participants were aged ≥60 years at baseline; (2) IVR intervention, alone or combined with other therapies, lasting 4 weeks or more; (3) the comparison group can be control, placebo, or another type of intervention; (4) outcome measures related to physical activity, quality of life, pain intensity, perceived effort, and adverse events; and (5) randomized controlled trials (RCTs), crossover clinical trials, and randomized mixed methods studies.

Exclusion criteria were: (1) studies with participants aged under 60 years, although the average age is 60 years or older; and (2) studies whose intervention is defined as immersive but does not use the head-mounted display.

No language restrictions were applied while searching.

Study Selection

Once the research question was defined, studies were identified and screened accordingly. Following the search strategy, all references were imported into Rayyan (Rayyan Systems Inc) to exclude duplicate studies [54].

Two independent researchers (IT-C and JB-A) conducted the study selection based on the predefined inclusion and exclusion criteria. In cases of disagreement, a third researcher (HB-A) was consulted to resolve discrepancies and reach a consensus.

Data Extraction

First, two researchers (IT-C and JB-A) independently extracted key information from the included studies. In case of discrepancies, a third researcher was consulted to resolve disagreements. The extracted data included: first author and year of publication, country, study design, sample size, age, diagnosis, setting, type of intervention of the groups, hardware and software of VR used, time of intervention (total amount of time in h, session time, frequency, and number of wk), outcomes, and follow-up period.

In the second phase, quantitative data for both primary and secondary outcomes were extracted. For the primary outcomes—physical activity and quality of life—as well as for the secondary outcomes of pain, perceived effort, and adverse events, mean and SD values were collected. When studies reported both change scores and final values, the final values were prioritized for analysis. When information regarding any of the above was unclear, we contacted authors of the reports to provide further details.

If data were only available in graphs, the graph digitization software GraphGrabber 2.0.2 (Quintessa Ltd) was used for extraction [55].

Risk of Bias

The methodological quality of the included studies was independently assessed by 2 researchers using the Cochrane Risk of Bias 2.0 tool for both parallel and crossover designs [56]. In instances where discrepancies arose between the two primary researchers, a third independent researcher was consulted to resolve disagreements. This tool evaluates the risk of bias across 5 domains: randomization process, missing outcome data, measurement of outcomes, selection of reported results, and deviations from intended interventions.

In addition, selective reporting will be judged by comparing published outcomes with registered protocols when available.

Main Outcomes

The primary outcomes of this study were quality of life and physical activity. Any objective or self-reported measure of these outcomes was considered eligible for inclusion, provided the instruments used demonstrated evidence of validity and reliability. No restrictions were placed on the number or timing of assessment time points (eg, baseline, postintervention, and follow-up).

The secondary outcomes included pain intensity, perceived effort, and adverse events. The same eligibility and selection criteria were applied, with preference given to the numeric rating scale (NRS) and visual analog scale (VAS) for pain intensity [57].

Statistical Analysis

A meta-analysis was conducted to estimate the overall effect of IVR interventions on quality of life and physical activity levels in older adults. As the included studies used different scales to assess these outcomes, the standardized mean difference (SMD) was used as the effect size measure. In all cases, higher scores indicated better outcomes, except for the study by Rodríguez-Fuentes et al [43], which used the Parkinson Disease Questionnaire (PDQ-39) scale, where lower scores reflect a better quality of life.

All hypothesis tests were 2-tailed, with statistical significance set at α=.05. Effect estimates are reported with 95% CIs. The inverse variance method was applied using a random-effects model, considering the expected clinical and methodological heterogeneity among studies. For each outcome, 95% CI were calculated, and a significance level of P<.05 was established. The magnitude of the effect was interpreted according to the Cohen criteria: small effect (SMD approximately 0.2), moderate (approximately 0.5), and large (≥0.8).

Statistical heterogeneity was assessed using the chi-square test, the ^{2} estimate, and the I^2 statistic. Heterogeneity was established as low for I^2=25%, moderate for I^2=50%, and high for I^2=75%. Subgroup analyses were conducted based on type of control group (active vs passive), intervention duration in weeks (6 - 8wk vs 10 - 12 wk), and total exposure time (180 - 480 min vs 600 - 1800 min). In addition, a sensitivity analysis was conducted by sequentially excluding each individual study to assess its impact on the overall effect size and heterogeneity.

All statistical analyses were performed using Review Manager (RevMan) version 5.4.1. The certainty of the evidence will be assessed using the GRADE (Grades of Recommendation Assessment Development and Evaluation) approach [58].

Deviations From Protocol

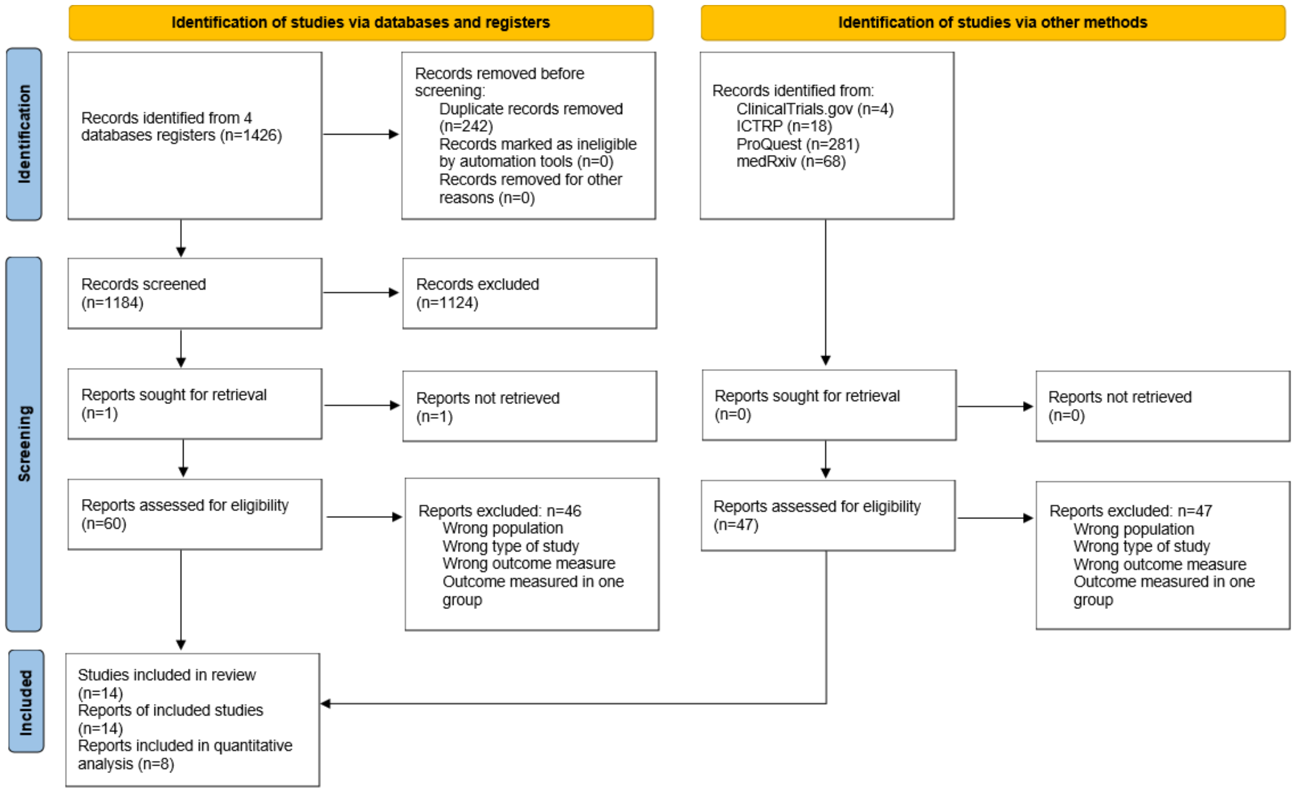
Several deviations from the International Prospective Register of Systematic Reviews protocol occurred. The mental component of the 12-item Short Form Survey (SF-12) was not analyzed to avoid conceptual overlap (Multimedia Appendix 1 [14,18,24,29,40-53]). Publication bias assessment was planned but not performed due to <10 studies per outcome (Multimedia Appendix 1 [14,18,24,29,40-53]). When required, means and SDs were estimated from median (IQR) following validated procedures (Multimedia Appendix 1 [14,18,24,29,40-53]). Subgroup analyses were restricted to comparator type, intervention duration, and total exposure; definitions and limitations are detailed in Multimedia Appendix 1 [14,18,24,29,40-53].

Results

Study Selection

The selection process is illustrated in the PRISMA 2020 flow diagram (Figure 1) in this review.

Figure 1. Flow diagram in this review. ICTRP: International Clinical Trials Registry Platform.



We found 1426 records in database searching. After duplicate removal, we screened 1184 records, from which we reviewed 60 full-text documents. Later, we searched documents that cited any of the initially included studies as well as the references of the initially included studies. However, no extra studies that fulfilled the inclusion criteria were found in these searches.

Ultimately, 14 trials met the inclusion criteria and were included in the systematic review. Of these, 8 studies were included in the meta-analysis. The remaining 6 studies were excluded from the quantitative synthesis because they did not report outcome measures corresponding to the primary endpoints analyzed in the meta-analysis (ie, quality of life or physical activity).

Study Characteristics

This systematic review included 14 RCTs published between 2021 and 2025 [14,19,25,43-53], with a total of 839 older adults and sample sizes ranging from 9 to 293 participants. While most studies included older adults without specific diagnoses [14,19,25,44,45], others focused on populations with cognitive impairment or frailty [46-48,50], balance and mobility issues [51], Parkinson disease [43], knee osteoarthritis [53], chronic low back pain [52], or individuals with joint arthroplasty [49].

Regarding settings, 10 studies were conducted in community centers or nursing homes [19,43-48,50-52], 3 in laboratory conditions [44,52], and 2 at home [14,53].

Most studies focused on physical activity delivered via VR, using environments that required movement [14,44,51]. In some cases, VR-based physical activity was combined with usual care or conventional rehabilitation, such as occupational therapy or kinesiotherapy [45,49]. Other interventions emphasized cognitive training delivered via VR [46], sometimes including underdesk ergometers for simultaneous stimulation [47,48]. Some protocols included education sessions, followed by structured VR-based exercise programs [25,53].

A few studies explored more therapeutic applications of VR, such as immersive reminiscence therapy [19], or multimodal pain management programs that integrated psychoeducation and movement therapy [52]. One study used VR to simulate daily living environments, aiming to promote autonomy in participants with cognitive frailty [50].

Control conditions varied: some used usual care [19,45,47,50], others applied active comparators such as conventional rehabilitation or group-based exercise [14,43,48,49,51-53] and a few used no-intervention or educational controls [25,44,46].

Session durations typically ranged from 15 to 60 minutes, with frequencies varying from 1 to 5 times per week over periods from 4 to 12 weeks. The longest interventions, by Lo et al [53] and Rodríguez-Fuentes et al [43], lasted 12 weeks.

Detailed tables summarizing the characteristics of each included study are provided in Table 1 (study characteristics) and Table 2 (intervention characteristics). In addition, detailed specifications of software, hardware, and outcome instruments are provided in Multimedia Appendix 1 [14,18,24,29,40-53].

Table . Study characteristics.

Study ID	Study design	Sample size	Age (y)	Diagnosis	Setting
Barsasella et al (2021) [44]	RCT[a]	n=60; VRG[b] n=29; CG[c] n=31	NI[d]	NSD[e]	Laboratory-based
Campo-Prieto et al (2022) [45]	RCT	n=24; VRG n=13; CG n=11	VRG: mean 85.08 (SD 8.48) CG: mean 84.82 (SD 8.1)	NSD	Clinical-based
Chiu et al (2023) [46]	RCT	n=60; VRG n=30; CG n=30	VRG: mean 80.7 (SD 8.8) CG: mean 80 (SD 7.9)	Cognitive impairment	Clinical-based
Drazich et al (2023) [24]	Pilot RCT	n=20; VRG n=10; CG n=10	Mean 74.1 (SD 6.5)	NSD	Clinical-based
Kershner et al (2024) [14]	Pilot RCT	n=9; VRG n=5; CG n=4	Mean 66.8 (SD 4.8)	NSD	Home-based
Khirallah Abd el Fatah et al (2024) [18]	RCT	n=60. VRG n=20; RT group n=20; CG n=20	Mean 66.68 (SD 4.22)	NSD	Clinical-based
Kwan et al (2021) [48]	Pilot RCT	n=15; VRG n=9; CG n=8	Median: 74 (IQR 9,5)	Cognitive frailty	Clinical-based
Kwan et al (2024) [47]	RCT	n=293; VRG n=146; CG n=147	Mean 74.5 (SD 6.8)	Cognitive frailty	Clinical-based
Lo et al (2024) [53]	Mixed methods pilot RCT	n=30; VRG n=15; CG n=15	Median: 63.5 (IQR 61.8 - 66.3)	Knee osteoarthritis	Home-based
Mazurek et al (2023) [49]	RCT	n=68; VRG n=34; CG n=34	Mean 69.59 (SD 6.16)	Hip or knee joint arthroplasty surgeries	Laboratory-based
Rodríguez-Fuentes et al (2024) [43]	RCT	n=52; VRG n=30; CG n=22	Mean 70.79 (SD 6.59)	Parkinson	Parkinson Association
Sekar et al (2024) [51]	RCT	n=60; VRG=30; CG=30	NI	Balance and mobility issues	Clinical-based
Stamm et al (2022) [52]	Pilot RCT	n=22; VRG n=11; CG n=11	VRG: mean 75.0 (SD 5.8) CG: mean 75.5 (SD 4.39)	Chronic low back pain	Laboratory-based
Zheng et al (2025) [50]	RCT	n=66; VRG n=33; CG n=33	Mean 80.20 (SD 9.14)	Cognitive frailty	Clinical-based

[a]RCT: randomized controlled trial.
[b]VRG: virtual reality group.
[c]CG: control group.

[d]NI: no information.

[e]NSD: no specific diagnosis.

Table . Intervention characteristics.

Study ID	Experimental group (s)	Control group	Frequency	Follow-up
Barsasella et al (2021) [44]	PA[a] via VR[b]	No intervention	15’ 2 t/w[c] 6 weeks TT:[d] 180’	—[e]
Campo-Prieto et al (2022) [45]	Usual care+PA via VR	Usual care (occupational therapy and memory workshops)	6’ 3 t/w 10 weeks TT: 180’	—
Chiu et al (2023) [46]	Cognitive training intervention via VR	No intervention	60’ 1 t/w 8 weeks TT: 480’	—
Drazich et al (2023) [24]	PA education+VR	1 session of PA education	40’ 2 t/w 8 weeks TT: 640’	—
Kershner et al (2024) [14]	PA via VR	PA via group videoconference	45 - 60’ minimum/w 4 weeks TT: 180 - 240’ minimum	—
Khirallah Abd el Fatah et al (2024) [18]	EXP 1: IVR[f] reminiscence therapy EXP 2: Traditional reminiscence therapy	Usual care (daily personal care, primary nursing care, medical care)	30 - 45’ 2 t/w 6 weeks TT: 360 - 540’	3 months
Kwan et al (2021) [48]	Motor and cognitive training on VR+underdesk ergometer	Cognitive training on tablet computers and motor training cycling on ergometer	30’ 2 t/w 8 weeks TT: 480’	—
Kwan et al (2024) [47]	Motor and cognitive training on VR+underdesk ergometer	Usual care (activities provided by the community centers)	20 - 30’ 2 t/w 8 weeks TT: 320 - 480’	—
Lo et al (2024) [53]	Health talk+lower limb exercises via VR	Lower limb exercises	30’ 5 t/w 12 weeks TT: 1800’	—
Mazurek et al (2023) [49]	Relaxing VR+conventional rehabilitation	Conventional rehabilitation (kinesiotherapy, ergotherapy, laser therapy/magnetic therapy/electrotherapy)	VR sessions: 20’ 2 t/w 4 weeks TT: 160’	—
Rodríguez-Fuentes et al (2024) [43]	Cycloergometer+VR	Static Cycling using Smart Cycloergometers	25’ 2 t/w 12 weeks TT: 600’	—
Sekar et al (2024) [51]	Balance and mobility exercises with VR	Balance and mobility exercises	2 t/w 8 weeks	—

Study ID	Experimental group (s)	Control group	Frequency	Follow-up
Stamm et al (2022) [52]	Movement therapy and psychoeducation via VR	Conventional multimodal pain therapy (chair-based group exercises and psychoeducation units)	30’ 3 t/w 4 weeks TT: 360’	—
Zheng et al (2025) [50]	Scenarios with daily environments via VR	Usual care (nursing care, and routine activities like finger exercises and holiday paper cutting)	45’ 2 t/w 12 weeks TT: 1080’	—

[a]PA: physical activity.
[b]VR: virtual reality.
[c]t/w: times/week.
[d]TT: total time.
[e]—: not available.
[f]IVR: immersive virtual reality.

Main Outcomes

Quality of Life

Among all 6 studies assessed quality of life using various validated tools such as EuroQol VAS [44,53], SF-12 [45,50], World Health Organization Quality of Life Scale Brief Version (WHOQOL-BREF) [43,46] and PDQ-39 [46,50].

Overall, 5 studies reported improvements in quality of life following IVR interventions. Chiu et al [46] and Zheng et al [50] found the most notable effects, with WHOQOL-BREF scores nearly doubling and significant gains in SF-12 mental health, respectively. Only Barsasella et al [44] found no overall differences in EuroQol 5D-3L scores between groups, though improvements were noted in specific domains such as pain and anxiety.

Physical Activity

Among all 3 studies assessed physical activity using both self-report and objective measures (eg, Yale Physical Activity Survey, Garmin Vivosmart 4, and ActivPAL accelerometry) [14,25,53].

Drazich et al [24] observed modest gains in vigorous activity within the VR group, despite stable weekly activity levels. Kershner et al [14] reported greater gains in steps and vigorous activity for the video conference group, though the VR group improved more in functional capacity. Lo et al [53] found slightly higher metabolic equivalent of tasks in the VR group, but without significant differences.

Intensity of Pain

Out of all 2 studies addressed pain. Lo et al [53] used the NRS and Western Ontario and McMaster Universities Osteoarthritis Index subscale in patients with osteoarthritis, showing reductions in both scores within the VR group (NRS: 5.93-4.78; Western Ontario and McMaster Universities Osteoarthritis Index: 189.5-160.2), with minimal changes in controls. However, these differences did not reach statistical significance when comparing IVR to control, suggesting no conclusive effect attributable to the intervention. Mazurek et al [49] reported a significant drop in VAS scores in the VR group (5.27-0.88), exceeding improvements in the control group.

Perceived Effort

None of the included studies systematically assessed perceived effort using validated tools such as the Borg rating of perceived exertion scale or comparable measures.

Adverse Events

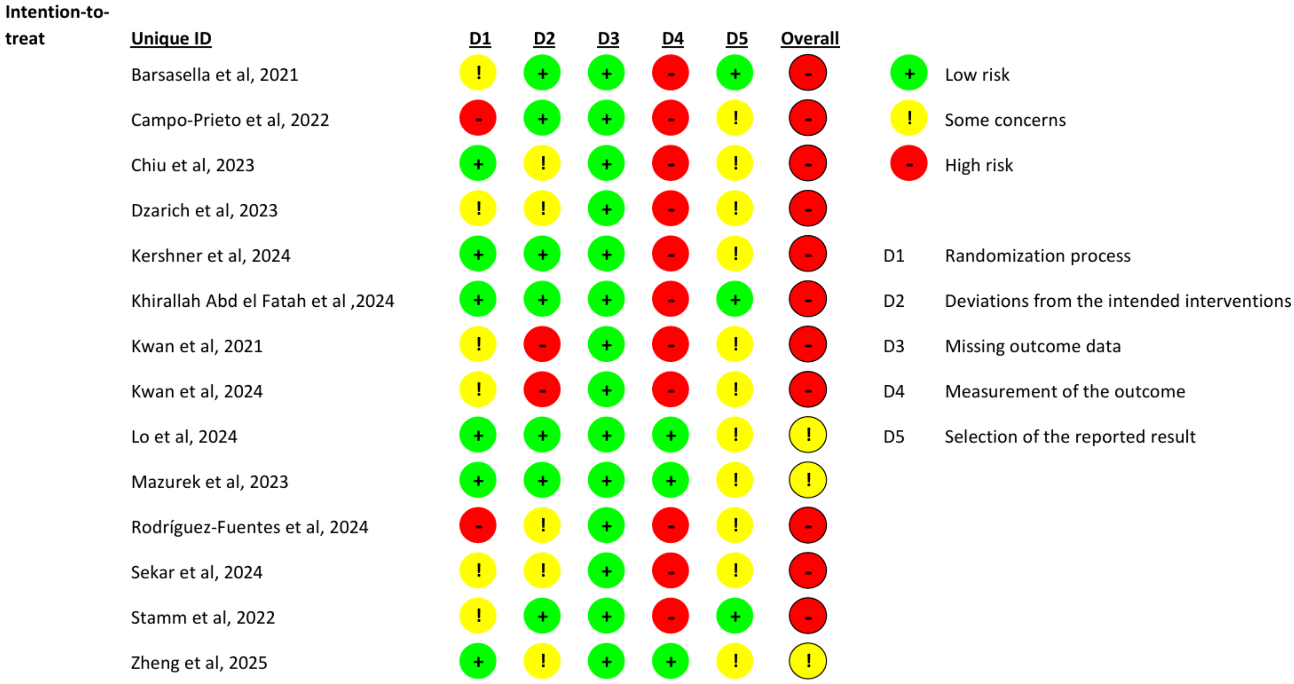
Adverse events were reported in 7 of the included studies, focusing on cybersickness [14,19,25,45,47,48,53]. Overall, most studies reported minimal or mild adverse events. In the pilot study of Kwan et al [47,48], 1 participant in the VR group withdrew early due to persistent symptoms of cybersickness, while the RCT reported low incidence rates (0.7%-3%) across 293 participants. Kershner et al [14] observed mild symptoms present in some participants. Lo et al [53] found that 5 out of 15 participants in the VR group reported mild adverse events such as dizziness or visual fatigue, though these did not result in discontinuation.

Risk of Bias in Studies

Most studies were judged to be at high risk of bias, except for 3 studies that showed a moderate risk [49,50,53]. All studies demonstrated low risk in domain 3 (missing outcome data). In domain 4 (measurement of the outcome), all but 3 studies were rated as high risk [49,50,53]. In domain 5 (selection of the reported result), 3 studies were rated as low risk, while the remaining studies were judged to have unclear risk.

As shown in Figure 2, the agreement rate achieved between the 2 researchers who completed risk of bias assessment was 81.43%. In case of disagreement, a third researcher resolved it.

Figure 2. Risk of bias [14,18,24,43-53].

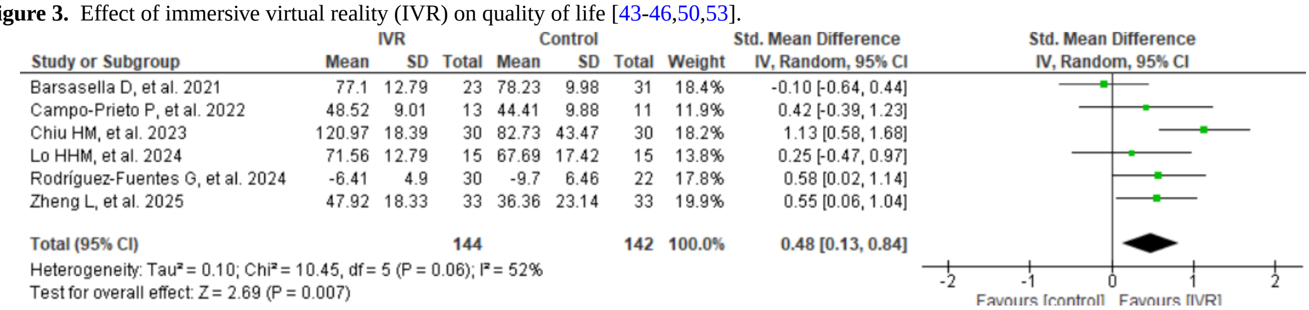


Results of Syntheses

Quality of Life

As shown in Figure 3 (effect of IVR on quality of life), a total of 6 RCTs including 286 participants (IVR group: n=144; control group: n=142) were synthesized to examine the effect of IVR interventions on quality of life in older adults. The meta-analysis yielded a statistically significant moderate effect in favor of IVR compared to control conditions (SMD=0.48, 95% CI 0.1-0.8; *P*=.007). Heterogeneity was moderate (I^2=52%; *P*=.06; ^{2}=0.10), indicating some variability across studies.

Figure 3. Effect of immersive virtual reality (IVR) on quality of life [43-46,50,53].



As shown in Table 3 subgroup analyses were performed to explore the potential influence of comparator type, intervention duration in weeks, and total intervention time in minutes.

Table . Subgroup analysis.

Subgroup	Studies, n	Participants, n	Random-effects model		Heterogeneity I^2[b] (%)	Between subgroups difference	
			SMD[a] (95% CI)	*P* value		Chi-square (*df*)	*P* value
Comparator type						0.02 (1)	.89
Passive control	4	110	0.50 (−0.04 to 1.04)	.07	71		
Active control	2	82	0.48 (0.12 to 0.84)	.05	0		
Weeks of intervention						0.00 (1)	.97
6 - 8	2	120	0.51 (−0.69 to 1.72)	.40	90		
10 - 12	4	172	0.48 (0.18 to 0.79)	.002	0		
Minutes of intervention						0.00 (1)	.99
180 - 480	3	138	0.49 (−0.31 to 1.28)	.23	80		
600 - 1800	3	286	0.49 (0.13 to 0.84)	.003	0		

[a]SMD: standardized mean difference.

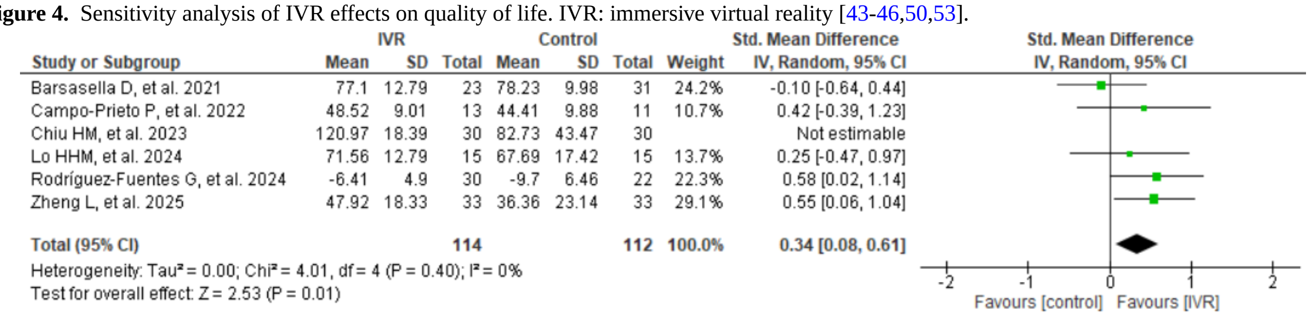
[b]I^2: inconsistency index.

Within-subgroup analyses showed statistically significant effects in some conditions. For instance, a significant effect was observed in the subgroup receiving 600 - 1800 minutes of intervention (SMD=0.49, 95% CI 0.13-0.84; *P*=.003), in contrast to the nonsignificant effect in the 180 - 480 minutes group. Similarly, the 10 - 12 week intervention subgroup showed a statistically significant moderate effect (SMD=0.48, 95% CI 0.18-0.79; *P*=.002), while the 6 - 8 week group did not (*P*=.40).

Regarding comparator type, significant effects were observed in the active control group (*P*=.05), but not in the passive control group (*P*=.07). However, despite these within-subgroup differences, the overall between-subgroup comparisons were not statistically significant for any of the 3 variables assessed: comparator type (*P*=.89), duration in weeks (*P*=.97), or total minutes of intervention (*P*=.99). These findings suggest that while effect sizes may vary descriptively across subgroups, such differences are not supported statistically.

As shown in Figure 4 (sensitivity analysis of IVR effects on quality of life), a sensitivity analysis was conducted excluding the study by Chiu et al [46], which presented the largest effect size (SMD=1.1). When this study was removed, the overall effect in favor of IVR remained statistically significant (SMD=0.3, 95% CI 0.08-0.6; *P*=.01), although the magnitude of the effect was reduced. Notably, heterogeneity was eliminated (I^2=0%; ^{2}=0.00), suggesting that this study contributed substantially to the observed heterogeneity in the main analysis.

Figure 4. Sensitivity analysis of IVR effects on quality of life. IVR: immersive virtual reality [43-46,50,53].

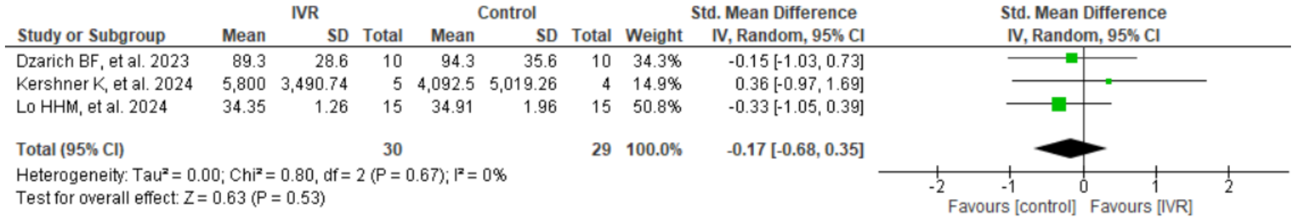


Physical Activity

Among all 3 RCTs involving a total of 59 participants (IVR group: n=30; control group: n=29), IVR was assessed on physical activity in older adults. The meta-analysis revealed no statistically significant difference between groups (SMD=−0.2, 95% CI −0.7 to 0.4; *P*=.50).

As shown in Figure 5 (effect of IVR on physical activity), heterogeneity among studies was minimal (I^2=0%; *P*=.70; ^{2}=0.00), indicating high consistency in the direction and magnitude of effects across studies. Due to the small number of studies, no subgroup or sensitivity analyses were performed for this outcome.

Figure 5. Effect of immersive virtual reality (IVR) on physical activity [14,24,53].



Certainty of Evidence

As shown in Figure 6 (assessment of evidence according to GRADE), the certainty of the evidence for the outcome quality of life was rated as low. Downgrading was applied due to very serious concerns related to the risk of bias. For the outcome of physical activity, the certainty of the evidence was rated as very low. Downgrading was applied due to serious risk of bias, very serious imprecision, and strong suspicion of publication bias.

Figure 6. Assessment of evidence according to GRADE (Grades of Recommendation Assessment Development and Evaluation). IVR: immersive virtual reality; SMD: standardized mean difference. *Out of 6 studies, 4 contributing to this outcome were judged to have high risk of bias, particularly due to issues in the randomization process, deviations from intended interventions, and selective reporting. The remaining 2 had some concerns. **Out of the 3 studies, 2 were judged to have high risk of bias, particularly in randomization and outcome assessment. The overall certainty was downgraded due to methodological limitations. ***The CI was wide and included both meaningful benefit and harm. The small sample size and lack of statistical significance led to downgrading by 2 levels. ****Only 3 studies were included in the meta-analysis for physical activity, which prevents formal assessment of publication bias. Due to the small number of trials and likelihood of selective reporting, strong suspicion of publication bias was considered.

Certainty assessment						N of patients		Effect		Certainty
N of studies	Study design	Risk of bias	Inconsistency	Imprecision	Other considerations	IVR	Control group	Relative (95% CI)	Absolute (95% CI)	
Quality of life										
6	Randomized trials	Very serious*	Not serious	Not serious	None	144	142	-	SMD 0.48 higher (0.13 higher to 0.84 higher)	Low*
Physical activity										
3	Randomized trials	Serious**	Not serious	Very serious***	Publication vias strongly suspected****	30	29	-	SMD 0.17 lower (0.68 lower to 0.35 higher)	Very low**,***,****

Discussion

Principal Findings

To our knowledge, this review is the first to focus exclusively on IVR interventions in older adults with a minimum duration of ≥4 weeks as a predefined inclusion criterion. The findings reveal a statistically significant moderate effect of IVR on quality of life, especially in individuals with clinical vulnerability or living in institutional settings. In contrast, no significant effects were observed for physical activity, and the direction of the results across included studies was inconsistent. Although some isolated findings indicated minor improvements in specific physical activity parameters, these were not replicated across trials and were derived from interventions of limited intensity or duration. As such, current evidence does not support definitive conclusions regarding the impact of IVR on physical activity, though it highlights important considerations for future program design. On the other hand, given the limited number of studies and heterogeneous measurement, secondary outcomes (pain, perceived exertion, and adverse events) could not be robustly quantified. We retained these outcomes because they were prespecified in our protocol and are clinically salient, but current evidence is insufficient to support firm conclusions. Additionally, preliminary results indicate potential benefits for pain reduction, and IVR was well tolerated, with low incidence of mild adverse events.

Quality of Life

Our findings suggest a moderate, statistically significant improvement in quality of life with IVR versus control; however, the certainty of evidence is low due to risk of bias. The observed effect size—approaching the 0.5 threshold commonly regarded as clinically meaningful in geriatric interventions—suggests a relevant improvement in this population. The psychological sense of presence elicited by IVR may diminish the perception of aversive stimuli such as pain or anxiety and promote emotional regulation, intrinsic motivation, and well-being [59,60].

Notably, the most pronounced benefits were evident among older adults with clinical diagnoses, indicating efficacy in individuals with functional vulnerabilities. This observation is consistent with previous research reporting cognitive gains in populations with mild cognitive impairment [61]. Moreover, interventions implemented in residential or institutional environments tended to yield more consistent and favorable outcomes. In support of this, Li et al [62] reported that IVR experiences in nursing homes enhanced not only well-being and social engagement but were also perceived as meaningful and motivating by participants.

Regarding intervention characteristics, programs incorporating cognitive or functional components generated more substantial improvements than those centered exclusively on physical exercise. This may reflect the inherently multidimensional nature of quality of life, which integrates cognitive, emotional, and social dimensions in addition to physical health [63].

In terms of duration, longer interventions (10 - 12 wk or exceeding 600 min) could be associated with more reliable improvements in quality of life. This finding is aligned with results from Vasodi et al [64], who reported that extended IVR programs led to better outcomes in older adults' mood and well-being, potentially due to increased engagement, gradual adaptation, and the consolidation of behavioral changes over time.

Physical Activity

The effectiveness of IVR in promoting physical activity among older adults appears to be highly contingent on intervention design; however, in our review, the evidence is very uncertain. The meta-analysis included only 3 randomized trials and showed no statistically significant differences between IVR and control groups, with inconsistent directions of effect across studies. Coupled with the very low certainty of evidence, small total sample size, and wide confidence intervals, these findings should be interpreted as hypothesis-generating rather than decision-informing. These findings align with previous evidence suggesting that brief, lab-based programs are generally insufficient to generate sustained behavioral change. For instance, studies by Lo et al [53] and Drazich et al [24] reported only modest or transient increases in activity, likely due to seated, low-intensity exercises and short durations. In contrast, research in younger adults shows IVR can elicit greater physiological responses—such as increased oxygen consumption and enjoyment—when compared to traditional 2D formats, likely due to enhanced emotional engagement [65].

The success of longer, home-based interventions like those reported by Dinet and Nouchi [66] may reflect the critical importance of habit formation and environmental integration, which require extended exposure periods and real-world application contexts. However, such outcomes seem contingent on the intervention's ability to integrate into daily routines, adapt to user capacity, and sustain motivation over time.

A further limitation is the reliance on self-reported physical activity measures, which are subject to bias. Future research should use wearable devices with validated protocols for objective monitoring. Overall, IVR can support physical activity, but only if programs are engaging, adaptable, and promote long-term autonomous use.

Intensity of Pain

The effects of IVR on pain intensity in older adults show heterogeneous results, probably conditioned by the design of the interventions and the characteristics of the participants. In the pilot study by Lo et al [53] with older people with knee osteoarthritis, the differences in pain intensity did not reach statistical significance compared to the control group, which may be attributed to the limited format of the intervention, focusing exclusively on strength without aerobic elements or relevant visual distracters.

In contrast, the study by Mazurek et al [49] shows more robust results: after 8 sessions of IVR with psychotherapeutic approach and immersive relaxation, a significant reduction in pain was observed, with a significant difference compared to the control group. This effect could be explained by the inclusion of psychological components, such as attentional distraction, emotional reinforcement, and the use of therapeutic metaphors, which have been associated with a downward modulation of pain in previous neurobiological studies [67,68].

Additional studies reinforce this approach. Li et al [69] observed a significant reduction in chronic low back pain in older adults after 8 weeks of IVR combined with functional exercise, highlighting increased adherence and reduced analgesic use in the IVR group. Taken together, the evidence suggests that IVR may be effective in reducing pain in older adults, especially when applied with a multisensory, emotional, and adaptive approach.

Adverse Events

The results of this review suggest that IVR is generally well tolerated by older adults, with a low incidence of adverse events, mainly related to mild symptoms of cybersickness (dizziness, nausea, and visual fatigue). Studies such as those by Campo-Prieto et al [45], Drazich et al [24], or Khirallah et al [18] reported no adverse events, while others, such as Lo et al [53], did observe mild symptoms in one third of participants, with no related dropouts. In Kwan et al [47,48], one withdrawal due to persistent cybersickness was reported, but in their later study, with a larger sample size, the incidence was low (0.7% - 3%). These findings are consistent with previous reviews. Weech et al [70], in a systematic review, identified that symptoms of cybersickness are common in immersive environments, but their severity tends to be mild and dependent on factors such as content type, duration of exposure, and individual characteristics. Stanney et al [71] emphasized the importance of individualized visor fit, especially interpupillary distance, noting that poor fit significantly increases discomfort, especially in women.

Overall, the evidence suggests good overall tolerability, although not without some episodes.

Clinical Implications

The findings suggest that IVR may be a valuable clinical tool to enhance the quality of life in older adults, particularly those with functional limitations or in institutional settings. Given the predominance of high risk of bias across several trials and low to very low certainty by GRADE, the pooled effects should be interpreted as signals of possible benefit, not as precise estimates for clinical decision-making.

Interventions combining cognitive, physical, and motivational elements show greater effectiveness. Although physical activity outcomes were not significant, appropriate program design appears crucial for adherence. Preliminary evidence also supports IVR's use in pain management, especially with emotionally engaging approaches. Its low rate of adverse events indicates good tolerability. With proper device adjustment and supervision, IVR can be progressively integrated into geriatric rehabilitation across various care environments.

From a policy perspective, our findings align with the UN Decade of Healthy Aging, particularly its priority on maintaining functional ability and person-centered care. IVR could act as an engagement-enhancing tool when embedded within routine health and social care services [72]. Operationally, its integration should be guided by WHO's integrated care for older people pathways in primary and community care and adhere to the principles of the Global Strategy on Digital Health—namely interoperability, equity, and evidence-based implementation [72].

Limitations

This systematic review and meta-analysis present several limitations related to its design and execution. First, although the search strategy was comprehensive and included 4 major databases, relevant studies indexed in other sources may have been missed. However, no language restrictions were applied, which mitigates selection bias. Second, the number of studies eligible for inclusion in the meta-analysis was limited, particularly for some outcomes such as physical activity and pain, reducing the statistical power and precision of the effect estimates. Additionally, due to the low number of included trials per outcome (n<10), we could not perform funnel plots or the Egger test to formally assess publication bias. Third, the heterogeneity in outcome measurement instruments and reporting formats across studies made it necessary to apply specific criteria for data inclusion, such as prioritizing final values over change scores and estimating means and SD when only medians and interquartile ranges were reported. These decisions, although methodologically justified, may introduce some degree of imprecision. Finally, some planned subgroup analyses could not be conducted due to insufficient data availability. While sensitivity analyses were performed where possible, the overall ability to explore sources of heterogeneity was limited.

A wide range of instruments was used to assess quality of life and physical activity (SF-12, EuroQol-5 Dimensions, WHOQOL-BREF, PDQ-39, accelerometers, and self-report surveys). While we synthesized conceptually similar constructs, instrument heterogeneity reduces scale-specific interpretability and likely contributed to between-study variance.

Future RCTs should address current evidence gaps by prioritizing: (1) the use of standardized and psychometrically validated outcome measures, particularly for physical activity and perceived effort; (2) objective monitoring of physical activity levels through wearable devices to reduce reliance on self-report; (3) long-term follow-up assessments to determine the durability of IVR effects on quality of life and functional outcomes; and (4) cost-effectiveness analyses to evaluate the feasibility of implementing IVR programs in real-world geriatric care settings; and (5) future studies should focus on the analysis of contextual factors—such as delivery model (standalone IVR vs adjunct), supervision intensity, and care setting—as potential effect modifiers. In addition, the main bias in the review was due to the unblinding of assessors and nonregistered protocols; therefore, future studies should (1) blind outcome assessors, and (2) prospectively register protocols with a prespecified analysis plan to prevent selective reporting.

Conclusions

IVR interventions lasting 4 weeks or more appear to moderately improve quality of life in older adults, especially those with clinical vulnerability or living in institutional settings. In contrast, no significant effects were observed for physical activity, and available evidence does not support a consistent trend in favor of IVR. Further research is needed to determine whether specific program designs could enhance its impact in this domain. Pain, perceived effort, and adverse events are included as secondary outcomes, but the evidence is sparse and largely qualitative. Policies and previous research also support the potential of IVR in reducing pain, particularly when using multisensory and emotionally engaging approaches, although the certainty of evidence is low to very low according to GRADE; therefore, these findings should be interpreted with caution. Overall, IVR is well tolerated and shows promise as a safe, adaptable, and motivating tool to support healthy aging, warranting further research in diverse settings.

Funding

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

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study information.
[DOCX File, 45 KB - aging_v9i1e80820_app1.docx]

Checklist 1
PRISMA checklist.
[DOCX File, 274 KB - aging_v9i1e80820_app2.docx]

References

1. World population prospects 2022: summary of results. : United Nations; 2022 URL: https://www.un.org/development/desa/pd/content/World-Population-Prospects-2022 [accessed 2025-12-12]
2. McPhee JS, French DP, Jackson D, Nazroo J, Pendleton N, Degens H. Physical activity in older age: perspectives for healthy ageing and frailty. Biogerontology 2016 Jun;17(3):567-580. [doi: 10.1007/s10522-016-9641-0] [Medline: 26936444]
3. Bellantuono I. Find drugs that delay many diseases of old age. Nature New Biol 2018 Feb 15;554(7692):293-295. [doi: 10.1038/d41586-018-01668-0] [Medline: 29446384]
4. Aging and health [Web page in Spanish]. World Health Organization. URL: https://www.who.int/es/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-12-24]
5. Daskalopoulou C, Stubbs B, Kralj C, Koukounari A, Prince M, Prina AM. Physical activity and healthy ageing: a systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev 2017 Sep;38(6–17):6-17. [doi: 10.1016/j.arr.2017.06.003] [Medline: 28648951]
6. Cunningham C, O' Sullivan R, Caserotti P, Tully MA. Consequences of physical inactivity in older adults: a systematic review of reviews and meta-analyses. Scand J Med Sci Sports 2020 May;30(5):816-827. [doi: 10.1111/sms.13616] [Medline: 32020713]
7. Roberts CE, Phillips LH, Cooper CL, Gray S, Allan JL. Effect of different types of physical activity on activities of daily living in older adults: systematic review and meta-analysis. J Aging Phys Act 2017 Oct 1;25(4):653-670. [doi: 10.1123/japa.2016-0201] [Medline: 28181837]
8. Alanazi MA. The role of physical activity in adjunctive nursing management of neuro-degenerative diseases among older adults: a systematic review of interventional studies. Life (Basel) 2024 May 7;14(5):597. [doi: 10.3390/life14050597] [Medline: 38792618]
9. Olanrewaju O, Kelly S, Cowan A, Brayne C, Lafortune L. Physical activity in community dwelling older people: a systematic review of reviews of interventions and context. PLoS ONE 2016;11(12):e0168614. [doi: 10.1371/journal.pone.0168614] [Medline: 27997604]
10. Fukushima N, Kikuchi H, Sato H, et al. Dose-response relationship of physical activity with all-cause mortality among older adults: an umbrella review. J Am Med Dir Assoc 2024 Mar;25(3):417-430. [doi: 10.1016/j.jamda.2023.09.028] [Medline: 37925162]
11. Picorelli AMA, Pereira LSM, Pereira DS, Felício D, Sherrington C. Adherence to exercise programs for older people is influenced by program characteristics and personal factors: a systematic review. J Physiother 2014 Sep;60(3):151-156. [doi: 10.1016/j.jphys.2014.06.012] [Medline: 25092418]
12. Schutzer KA, Graves BS. Barriers and motivations to exercise in older adults. Prev Med 2004 Nov;39(5):1056-1061. [doi: 10.1016/j.ypmed.2004.04.003] [Medline: 15475041]
13. Grassini S, Laumann K. Immersive visual technologies and human health. Presented at: ECCE '21: Proceedings of the 32nd European Conference on Cognitive Ergonomics; Apr 26-29, 2021. [doi: 10.1145/3452853.3452856]
14. Kershner K, Morton D, Robison J, N'dah KW, Fanning J. Assessing the feasibility and acceptability of virtual reality for remote group-mediated physical activity in older adults: pilot randomized controlled trial. JMIR Form Res 2024 Nov 8;8:e53156. [doi: 10.2196/53156] [Medline: 39514256]
15. Van Veelen N, Boonekamp R, Schoonderwoerd T, et al. Tailored immersion: implementing personalized components into virtual reality for veterans with post-traumatic stress disorder. Eur Psychiatr 2022 Jun;65(S1):S675-S675. [doi: 10.1192/j.eurpsy.2022.1737]
16. Cucinella SL, de Winter JCF, Grauwmeijer E, Evers M, Marchal-Crespo L. Towards personalized immersive virtual reality neurorehabilitation: a human-centered design. J Neuroeng Rehabil 2025 Jan 20;22(1):7. [doi: 10.1186/s12984-024-01489-5] [Medline: 39833912]
17. Wada T, Matsumoto H, Hagino H. Customized exercise programs implemented by physical therapists improve exercise-related self-efficacy and promote behavioral changes in elderly individuals without regular exercise: a randomized controlled trial. BMC Public Health 2019 Jul 9;19(1):917. [doi: 10.1186/s12889-019-7270-7] [Medline: 31288781]
18. Khirallah Abd El Fatah N, Abdelwahab Khedr M, Alshammari M, Mabrouk Abdelaziz Elgarhy S. Effect of immersive virtual reality reminiscence versus traditional reminiscence therapy on cognitive function and psychological well-being among older adults in assisted living facilities: a randomized controlled trial. Geriatr Nurs 2024;55:191-203. [doi: 10.1016/j.gerinurse.2023.11.010] [Medline: 38007908]
19. Porras-Garcia B, Rojas-Rincón J, Adams A, Garolera M, Chang R. Immersive virtual reality cognitive training for improving cognition and depressive symptoms among older adults. current evidence and future recommendations. a systematic review. Cyberpsychol Behav Soc Netw 2024 Oct;27(10):692-703. [doi: 10.1089/cyber.2024.0090] [Medline: 39180434]
20. Lee J, Phu S, Lord SR, Okubo Y. Effects of immersive virtual reality training on balance, gait and mobility in older adults: a systematic review and meta-analysis. Gait Posture 2024 May;110:129-137. [doi: 10.1016/j.gaitpost.2024.03.009] [Medline: 38581933]
21. Maheta B, Kraft A, Interrante N, et al. Using virtual reality to improve outcomes related to quality of life among older adults with serious illnesses: systematic review of randomized controlled trials. J Med Internet Res 2025 Feb 26;27:e54452. [doi: 10.2196/54452] [Medline: 40009834]

22. Dermody G, Whitehead L, Wilson G, Glass C. The role of virtual reality in improving health outcomes for community-dwelling older adults: systematic review. J Med Internet Res 2020 Jun 1;22(6):e17331. [doi: 10.2196/17331] [Medline: 32478662]
23. Ke Z, Wei M, Yang F, et al. The effectiveness of immersive virtual reality on the psychology of older adults: a systematic review and meta-analysis of randomized controlled trials. Gen Hosp Psychiatry 2025;94(86–96):86-96. [doi: 10.1016/j.genhosppsych.2025.02.024] [Medline: 40036985]
24. Drazich BF, Anokye D, Zhu S, et al. Motivating older adults through immersive virtual exercise (motive): a randomized pilot study. Geriatr Nurs 2023;54:229-236. [doi: 10.1016/j.gerinurse.2023.09.019] [Medline: 37844539]
25. Doré B, Gaudreault A, Everard G, et al. Acceptability, feasibility, and effectiveness of immersive virtual technologies to promote exercise in older adults: a systematic review and meta-analysis. Sensors (Basel) 2023 Feb 24;23(5):2506. [doi: 10.3390/s23052506] [Medline: 36904709]
26. Chen PJ, Hsu HF, Chen KM, Belcastro F. VR exergame interventions among older adults living in long-term care facilities: a systematic review with meta-analysis. Ann Phys Rehabil Med 2023 Apr;66(3):101702. [doi: 10.1016/j.rehab.2022.101702] [Medline: 36028201]
27. Kim KA, Ahn JA. Effectiveness of immersive virtual reality simulation programs using head-mounted displays in promoting physical activity in older adults: a systematic review. Clin Simul Nurs 2024 Sep;94:101593. [doi: 10.1016/j.ecns.2024.101593]
28. John W, Kosinski M, Gandek B. SF-36 Health Survey: Manual & Interpretation Guide 1993. URL: https://czresearch.com/info/SF36_healthsurvey_ch6.pdf [accessed 2025-12-12]
29. Ware JE, Kosinski M, Diane TB. How to Score SF-12 Physical and Mental Health Summary Scales, 2nd edition: The Health Institute, New England Medical Center; 1995.
30. Rong W, Geok SK, Samsudin S, Zhao Y, Ma H, Zhang X. Effects of strength training on neuromuscular adaptations in the development of maximal strength: a systematic review and meta-analysis. Sci Rep 2025 Jun 2;15(1):19315. [doi: 10.1038/s41598-025-03070-z] [Medline: 40456806]
31. Škarabot J, Brownstein CG, Casolo A, Del Vecchio A, Ansdell P. The knowns and unknowns of neural adaptations to resistance training. Eur J Appl Physiol 2021 Mar;121(3):675-685. [doi: 10.1007/s00421-020-04567-3] [Medline: 33355714]
32. Brown N, Bubeck D, Haeufle DFB, et al. Weekly time course of neuro-muscular adaptation to intensive strength training. Front Physiol 2017;8(JUN):329. [doi: 10.3389/fphys.2017.00329] [Medline: 28642711]
33. Nouchi R, Taki Y, Takeuchi H, et al. Four weeks of combination exercise training improved executive functions, episodic memory, and processing speed in healthy elderly people: evidence from a randomized controlled trial. Age (Dordr) 2014 Apr;36(2):787-799. [doi: 10.1007/s11357-013-9588-x] [Medline: 24065294]
34. Jagdhane S, Kanekar N, Aruin AS. The effect of a four-week balance training program on anticipatory postural adjustments in older adults: a pilot feasibility study. Curr Aging Sci 2016;9(4):295-300. [doi: 10.2174/1874609809666160413113443] [Medline: 27071477]
35. Yue T, Yan X, Lv Y, Ren X, Zhang S, Qi F. Effects of immersive virtual reality-based tennis training on balance in older adults. Gerontology 2025;71(10):861-871. [doi: 10.1159/000547548] [Medline: 40695260]
36. Sadowski I, Meilleur-Bédard M, Khoury B. A novel virtual reality-based nature meditation program for older adults' mental health: results from a pilot randomized controlled trial. Clin Gerontol 2025;48(4):663-683. [doi: 10.1080/07317115.2025.2482089] [Medline: 40152067]
37. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol 2014 Mar;67(3):267-277. [doi: 10.1016/j.jclinepi.2013.08.015] [Medline: 24275499]
38. WHO's work on the UN decade of healthy ageing (2021-2030). World Health Organization. URL: https://www.who.int/initiatives/decade-of-healthy-ageing [accessed 2025-12-12]
39. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
40. Shi J, Luo D, Weng H, et al. Optimally estimating the sample standard deviation from the five‐number summary. Res Synth Methods 2020 Sep;11(5):641-654. [doi: 10.1002/jrsm.1429]
41. Luo D, Wan X, Liu J, Tong T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. Stat Methods Med Res 2018 Jun;27(6):1785-1805. [doi: 10.1177/0962280216669183] [Medline: 27683581]
42. Cheng LJ, Tan RLY, Luo N. Measurement properties of the EQ VAS around the globe: a systematic review and meta-regression analysis. Value Health 2021 Aug;24(8):1223-1233. [doi: 10.1016/j.jval.2021.02.003] [Medline: 34372988]
43. Rodríguez-Fuentes G, Campo-Prieto P, Cancela-Carral JM. Immersive virtual reality high-intensity aerobic training to slow parkinson's disease: the ReViPark program. Appl Sci ;14(11):4708. [doi: 10.3390/app14114708]
44. Barsasella D, Liu MF, Malwade S, et al. Effects of virtual reality sessions on the quality of life, happiness, and functional fitness among the older people: a randomized controlled trial from taiwan. Comput Methods Programs Biomed 2021 Mar;200:105892. [doi: 10.1016/j.cmpb.2020.105892] [Medline: 33280934]

45. Campo-Prieto P, Cancela-Carral JM, Rodríguez-Fuentes G. Feasibility and effects of an immersive virtual reality exergame program on physical functions in institutionalized older adults: a randomized clinical trial. Sensors (Basel) 2022 Sep 6;22(18):6742. [doi: 10.3390/s22186742] [Medline: 36146092]
46. Chiu HM, Hsu MC, Ouyang WC. Effects of incorporating virtual reality training intervention into health care on cognitive function and wellbeing in older adults with cognitive impairment: a randomized controlled trial. Int J Hum Comput Stud 2023 Feb;170:102957. [doi: 10.1016/j.ijhcs.2022.102957]
47. Kwan RYC, Liu J, Sin OSK, et al. Effects of virtual reality motor-cognitive training for older people with cognitive frailty: multicentered randomized controlled trial. J Med Internet Res 2024 Sep 11;26:e57809. [doi: 10.2196/57809] [Medline: 39259959]
48. Kwan RYC, Liu JYW, Fong KNK, et al. Feasibility and effects of virtual reality motor-cognitive training in community-dwelling older people with cognitive frailty: pilot randomized controlled trial. JMIR Serious Games 2021 Aug 6;9(3):e28400. [doi: 10.2196/28400] [Medline: 34383662]
49. Mazurek J, Cieślik B, Wrzeciono A, Gajda R, Szczepańska-Gieracha J. Immersive virtual reality therapy is supportive for orthopedic rehabilitation among the elderly: a randomized controlled trial. J Clin Med 2023 Dec 14;12(24):7681. [doi: 10.3390/jcm12247681] [Medline: 38137750]
50. Zheng L, Li X, Xu Y, et al. Effects of virtual reality-based activities of daily living rehabilitation training in older adults with cognitive frailty and activities of daily living impairments: a randomized controlled trial. J Am Med Dir Assoc 2025 Feb;26(2):105397. [doi: 10.1016/j.jamda.2024.105397] [Medline: 39615543]
51. Sekar M, Suganthirababu P, Subramanian SS, et al. The effectiveness of virtual reality (VR) therapy on balance and mobility in elderly patients: a randomized controlled trial. Fiz Pol 2024 Dec 31;24(5):191-194. [doi: 10.56984/8ZG020C8UWP]
52. Stamm O, Dahms R, Reithinger N, Ruß A, Müller-Werdan U. Virtual reality exergame for supplementing multimodal pain therapy in older adults with chronic back pain: a randomized controlled pilot study. Virtual Real 2022;26(4):1291-1305. [doi: 10.1007/s10055-022-00629-3] [Medline: 35194374]
53. Lo HHM, Ng M, Fong PYH, et al. Examining the feasibility, acceptability, and preliminary efficacy of an immersive virtual reality-assisted lower limb strength training for knee osteoarthritis: mixed methods pilot randomized controlled trial. JMIR Serious Games 2024 Sep 27;12(2024):e52563. [doi: 10.2196/52563] [Medline: 39331525]
54. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
55. Graph Grabber 2.0.2. Quintessa. 2020. URL: https://www.quintessa.org/software/downloads-and-demos/graph-grabber-2.0.2 [accessed 2025-12-12]
56. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
57. Vicente Herrero MT, Delgado Bueno S, Bandrés Moyá F, Ramírez Iñiguez de la Torre MV, Capdevila García L. Pain assessment. Comparative review of scales and questionnaires [Title in Spanish]. Rev Soc Esp Dolor 2018;25(4):228-236. [doi: 10.20986/resed.2018.3632/2017]
58. McMaster University. GRADEpro. URL: https://www.gradepro.org/ [accessed 2025-12-12]
59. Riva G, Wiederhold BK, Mantovani F. Neuroscience of virtual reality: from virtual exposure to embodied medicine. Cyberpsychol Behav Soc Netw 2019 Jan;22(1):82-96. [doi: 10.1089/cyber.2017.29099.gri] [Medline: 30183347]
60. Chirico A, Lucidi F, De Laurentiis M, Milanese C, Napoli A, Giordano A. Virtual reality in health system: beyond entertainment. a mini-review on the efficacy of vr during cancer treatment. J Cell Physiol 2016 Feb;231(2):275-287. [doi: 10.1002/jcp.25117] [Medline: 26238976]
61. Yang Q, Zhang L, Chang F, Yang H, Chen B, Liu Z. Virtual reality interventions for older adults with mild cognitive impairment: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2025 Jan 10;27:e59195. [doi: 10.2196/59195] [Medline: 39793970]
62. Li Y, Wilke C, Shiyanov I, Muschalla B. Impact of virtual reality-based group activities on activity level and well-being among older adults in nursing homes: longitudinal exploratory study. JMIR Serious Games 2024 Mar 29;12:e50796. [doi: 10.2196/50796] [Medline: 38551635]
63. Liao Y, Hsu MHK, Xie FL, Dai HX, Liu M. The role of virtual reality on physical and psychological health among older adults: a systematic review. Edelweiss Appl Sci Technol 2024 Oct 25;8(6):2762-2777. [doi: 10.55214/25768484.v8i6.2556]
64. Vasodi E, Saatchian V, Dehghan Ghahfarokhi A. Virtual reality-based exercise interventions on quality of life, some balance factors and depression in older adults: a systematic review and meta-analysis of randomized controlled trials. Geriatr Nurs 2023;53:227-239. [doi: 10.1016/j.gerinurse.2023.07.019] [Medline: 37598426]
65. Barbour B, Sefton L, Bruce RM, Valmaggia L, Runswick OR. Acute psychological and physiological benefits of exercising with virtual reality. In: Sun F, editor. PLoS ONE 2024;19(12):e0314331. [doi: 10.1371/journal.pone.0314331] [Medline: 39693283]
66. Dinet J, Nouchi R. Promoting physical activity for elderly people with immersive virtual reality (IVR). Presented at: 13th International Conference on Applied Human Factors and Ergonomics (AHFE 2022); Jun 24-28, 2022. [doi: 10.54941/ahfe1001682]

67. Malloy KM, Milling LS. The effectiveness of virtual reality distraction for pain reduction: a systematic review. Clin Psychol Rev 2010 Dec;30(8):1011-1018. [doi: 10.1016/j.cpr.2010.07.001] [Medline: 20691523]
68. Hoffman HG, Chambers GT, Meyer WJ 3rd, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann Behav Med 2011 Apr;41(2):183-191. [doi: 10.1007/s12160-010-9248-7] [Medline: 21264690]
69. Li A, Montaño Z, Chen VJ, Gold JI. Virtual reality and pain management: current trends and future directions. Pain Manag 2011 Mar;1(2):147-157. [doi: 10.2217/pmt.10.15] [Medline: 21779307]
70. Weech S, Kenny S, Barnett-Cowan M. Presence and cybersickness in virtual reality are negatively related: a review. Front Psychol 2019;10:158. [doi: 10.3389/fpsyg.2019.00158] [Medline: 30778320]
71. Stanney K, Fidopiastis C, Foster L. Virtual reality is sexist: but it does not have to be. Front Robot AI 2020;7:4. [doi: 10.3389/frobt.2020.00004] [Medline: 33501173]
72. Integrated care for older people (ICOPE) : guidance for person-centred assessment and pathways in primary care, 2nd ed. : World Health Organization; 2024 URL: https://www.who.int/publications/i/item/9789240103726 [accessed 2024-12-12]

Abbreviations

GRADE: Grades of Recommendation Assessment Development and Evaluation
IVR: immersive virtual reality
NRS: numeric rating scale
PDQ-39: Parkinson Disease Questionnaire-39
PICOS: population, intervention, comparison, outcome, study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
RevMan: Review Manager
SF-12: 12-item Short Form Survey
SMD: standardized mean difference
VAS: visual analog scale
VR: virtual reality
WHO: World Health Organization
WHOQOL-BREF: World Health Organization Quality of Life Scale Brief Version

Edited by H Köttl; submitted 17.Jul.2025; peer-reviewed by JF Costa, ME Heidari, P Adedigba; revised version received 03.Nov.2025; accepted 03.Nov.2025; published 12.Jan.2026.

Please cite as:
Trillo-Charlín I, Bravo-Aparicio J, Avendaño-Coy J, Beltrán-Alacreu H
Impact of 4 Weeks or More Immersive Virtual Reality on Quality of Life and Physical Activity in Older Adults: Systematic Review and Meta-Analysis
JMIR Aging 2026;9:e80820
URL: https://aging.jmir.org/2026/1/e80820
doi:10.2196/80820

Review

Lessons Learned About Digital Health Tool Acceptability Among Rural Older Adults: Systematic Review Guided by the Technology Acceptance Model

Zachary M Siegel[1], BA; Ellie Quinkert[1], BS; Jiya Pai[2]; Corinne H Miller[3], MLiS; Marquita W Lewis[4], PhD, MPH, MS

[1]University of Kentucky, Lexington, KY, United States

[2]Case Western Reserve University, Cleveland, United States

[3]Northwestern University, Chicago, IL, United States

[4]Department of Medical Social Sciences, Northwestern University, Chicago, IL, United States

Corresponding Author:
Marquita W Lewis, PhD, MPH, MS
Department of Medical Social Sciences
Northwestern University
750 N. Lake Shore Drive
Chicago, IL, 60611
United States
Phone: 1 312 503 1986
Email: marquita.lewis-thames@northwestern.edu

Abstract

Background: Digital health tools are increasingly vital in rural health care due to widespread hospital closures and the rapid adoption of telehealth during the COVID-19 pandemic. Rural older adults, a uniquely vulnerable population, face barriers to accessing these tools due to rurality and usability challenges. Although a growing body of literature examines the acceptability and usability of digital tools among rural older adults, no study has synthesized this research to establish best practices.

Objective: This study aims to review existing literature on digital health tools for rural older adults, highlighting key lessons learned about their acceptability and identifying strategies to improve usability for this population.

Methods: Following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, this study reviewed literature that investigated the role of digital health tools on the health outcomes of rural older adults (ie, at least 60 years old). The literature was retrieved from 5 electronic databases through June 2023. This study and all reviewed literature were conducted in the United States. Guided by a systematic process, 2 reviewers assessed relevant articles for eligibility, analyzed data, and extracted relevant content. The extracted findings were organized according to the evidence-based technology acceptance model, which assesses the acceptability of a technology by its usefulness, ease of use, and intention to use.

Results: The preliminary title review produced 7728 results, and 38 eligible manuscripts were included in the final review. Studies included both rural older adults and providers of rural older adults as participants. Digital health tools included, but were not limited to, videoconferencing, phone calls, telehealth monitoring, telemedicine appointments, and computer-based interventions. Findings on the usefulness of digital health tools by rural older adults were mixed. While digital health tools were useful for overcoming barriers to accessing care, these tools were less useful for rural older adults with limited digital literacy. Additionally, some studies described that the technology was easy but difficult to use when faced with environmental barriers, equipment issues, and discomfort with the technology. Rural older adults often reported an intention to use the technology after the study. Yet, on a few occasions, participants who preferred in-person care visits or did not have buy-in on the technology reported no intention to use the technology again.

Conclusions: Our review highlights that rural older adults and their providers generally view digital health tools as acceptable for delivering care and, in some cases, as a viable alternative to in-person clinic visits. While certain barriers impacted the acceptance of these tools among rural older adults, many of these challenges were not directly linked to their age or rural location; thus, they are potentially applicable to urban older adults.

Trial Registration: PROSPERO CRD42021287924; https://www.crd.york.ac.uk/PROSPERO/view/CRD42021287924

(JMIR Aging 2026;9:e70012) doi:10.2196/70012

KEYWORDS

health services; older adult; rural health; digital health; social determinants of health

Introduction

Digital health tools facilitate communication between patients and health care providers and offer access to resources. These tools encompass a range of technologies, including mobile health apps, electronic health records, wearable devices, and telehealth services. Social distancing mandates related to COVID-19 facilitated increased funding to support improved access to broadband internet and the rapid uptake of digital health tools [1]. To increase digital tool access and use by rural residents, in the spring of 2022, the US Department of Health and Human Services announced a US $16.3 million expansion to telehealth care in the Title X Family Planning Program [2]. Thus, rural health care professionals and systems were able to integrate digital tool uptake in their care rapidly [3]. Telehealth uptake in health clinics and hospitals increased by 154% in March 2020 compared with March 2019 [4]. For many rural patients, digital health tools are an essential component of their health care management and will likely remain important for timely and continuous rural care coordination [1].

Rural older adults represent a vulnerable population at the intersection of aging and rural residency, facing well-documented yet preventable challenges in accessing health care [5]. Rural residents are rapidly aging in place. For instance, 25% of older adults live in a rural or small town, and this is expected to rise to 33% by 2030 [6]. Additionally, for many rural older adults, care management is complex, confusing, and further challenged by coordination between distant health care facilities [7]. Since 2010, over 160 rural hospitals have permanently closed their doors, reducing access to inpatient care, which is critical for improving rural community health [8]. Therefore, aging rural populations will increasingly experience limited access to specialty care and poorer health outcomes [9].

Digital health tools can potentially overcome care coordination challenges for rural older adults. Once rural older adults engage with digital health tools, they often find their experience satisfactory and, at times, comparable to in-person visits [10]. Rural older adults evaluated web-based consultations conducted by service providers with high efficiency and satisfaction scores [11]. Once older adults understand the technology, they often find it an acceptable mode of care when punctuated by in-person visits.

Despite high levels of satisfaction with digital tools by rural older adults, compared with urban older adults, this vulnerable population has reduced telehealth use [12,13]. Also, although rural residents are willing to adopt digital health tools [14,15], studies show that rural older adults report slower telehealth uptake than younger rural adults [16,17]. This is partially due to barriers that make using digital health tools difficult for rural older adults. Some of these barriers include technical literacy, lack of technical support, cost, ownership of technology, and visual acuity [13]. In a systematic review including rural adults aged 55 years and older who have used telehealth, older adults reported a willingness to learn how to use various digital tools, but 30% felt too inexperienced with technology to use them [18]. Similarly, in a sample of Medicare enrollees, rural cancer survivors had a significantly lower predicted probability of internet use for patient-provider communication when compared with urban cancer survivors with Medicare (28% vs 46%) [19]. Importantly, not all rural older adults will find digital health tools to be a favorable health care management tool. Yet, funding to increase broadband access and the threat of widening rural medical deserts will facilitate continued telehealth uptake of digital tools by health care systems, thereby reinforcing the increased uptake of digital health tools by rural older adults. Increasing the acceptability of digital health tools and reducing barriers to their uptake for rural older adults are essential for providing health care to rural older adults.

Given the rapid acceleration of digital health tools by rural health care providers, rural older adults find digital health tools helpful. Still, rural older adults have reduced uptake of these technologies compared with both urban older adults and younger rural adults. With the increasing use of digital health tools, understanding their acceptability among rural older adults is crucial for ensuring this vulnerable population stays engaged in their care management and coordination as reliance on these tools continues to grow. Details about the rural older adults' digital health tool acceptability and usage can inform tool intervention design, implementation, and evaluation. Existing research summarizes the effectiveness of services such as telehealth among older adults, but strategies to improve rural older adults' usage of digital health tools are limited [18]. Therefore, this study will systematically review the existing literature in the United States on rural adults' acceptability of digital health tools and assess lessons learned on digital health tool usage among rural older adults.

Methods

Study Design

The study was analyzed and reported in accordance with the Cochrane systematic review guidelines and the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines (Multimedia Appendix 1) [20,21].

Search Strategy and Data Sources/Protocol Registration

A trained librarian conducted searches in MEDLINE (Ovid), Cochrane Database of Systematic Reviews (Wiley), Embase (Elsevier), CINAHL (EBSCO), and PsycINFO (EBSCO) databases. This search included articles published in English through June 5, 2023. Keywords and subject headings related to the following topics were used to identify possible articles: rural residents, older adults, the use of technology-enhanced tools to navigate telehealth, and acceptability. See the supplementary materials for the complete search strategy. Following, we searched the references of eligible articles for additional relevant articles. The protocol was registered post hoc in the International Prospective Register of Systematic Reviews (PROSPERO CRD42021287924).

All articles eligible for data extraction underwent title and abstract review, and full-text review by a pair of reviewers. Reviewers independently assessed the articles based on the eligibility criteria (see below) using standardized procedures in the systematic review software Covidence (Covidence systematic review software, Melbourne, Victoria, Australia). A pair of reviewers discussed and resolved conflicts in weekly meetings. Each article was assessed for quality by 2 reviewers. Extracted content and Quality Assessments were reported using Microsoft Excel and Covidence.

Eligibility Criteria

We included research articles of investigations conducted in the United States that assessed a digital health tool's ability to connect patients with providers, where at least 25% of the sample population identified as rural, and at least 25% of the sample identified as at least 60 years old. Articles were excluded if they were a review (eg, systematic or scoping), withdrawn, a conference proceeding, an abstract, or a dissertation. We also excluded articles published before 1999, as we deemed that the technology or lessons learned from the technology over 25 years ago were antiquated.

Data Extraction

Paired reviewers conducted consensus meetings to agree upon the rationale for data extraction content and synthesize the results. Table 1 displays the extraction content. In short, paired reviewers reported each article's title, first author, and year of publication. The outcome variables collected from each study included the participants' age (average or mean), study design, and type of digital health tool technology (eg, videoconference or wearables). Following the technology acceptance model (TAM), we extracted data related to the core TAM domains: perceived usefulness, perceived ease of use, and intention to use [22]. The perceived usefulness domain describes how much technology improves a patient's performance. Perceived ease of use is the effort required to use the technology. Last, intention to use refers to a patient's willingness to use the technology. Given that we aim to synthesize acceptability and lessons learned, our analysis did not assess the effect of the outcomes. This systematic review did not need an exploration of heterogeneity or a sensitivity analysis.

Table 1. Summary of findings.

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Anderson et al [23]	Mean age 61 years	Community-based outpatient clinics in rural Southeast Texas	Qualitative	Videoconferencing	• It was accessible. • Videoconferencing was convenient and reduced transportation issues.	• Participants reported that there was too much information covered in the 2 self-management classes and that 2.5 hours was too long for a single videoconferencing session. • Clinicians reported that they had no time to assist when technology problems occurred. • Clinicians reported that there was only limited clinic space to hold videoconferencing sessions.	• Not applicable.
Barton et al [24]	28.5% of the sample is 60 years or older	67.3% of the sample is from rural Colorado	Cross-sectional	Phone call, videoconference	• Providers reported that ease of access to patient records, scheduling follow-up visits, and timely follow-ups were all better accomplished during digital visits than office visits (2-4.5× higher, all P<.001).	• Not applicable.	• Patients who did not engage in telehealth listed “preference to seeing the provider in person” and “telehealth not being an option” as reasons.
Bernacchi et al [25]	Mean age 60.2 years	Rural dwelling residents from Southeast US	Mixed methods	Videoconferencing	• Access to an oncology nurse during COVID-19 increased rural cancer patients’ access to care, information, and education. • Participants gave unfavorable scores to questions that asked about the use of the technology. For example, the lowest scoring items were “my health is better than it was before I used the technology” (—X = 3, SD = 0.89).	• Participants struggled with connecting to appointments due to a lack of equipment or discomfort with digital technology. • Three participants with insufficient broadband used nearby telehealth satellite sites at local clinics or hospitals. • Participants often relied on family members for connection due to limited internet experience or poor broadband signals.	• Participants were committed to overcoming barriers in order to speak with their nurse.

Study	Population characteristics: Age	Population characteristics: Rural	Study design	Technology	Usefulness	Ease of use	Intention to use
Bonsignore et al [26]	Mean age 72 years	Rural counties in Western North Carolina	Mixed methods	Telemonitoring	• TapCloud increased patient-provider relationship, accessibility to clinicians, increased response times, improved efficiency, and made medications more accessible. • Accessibility to a provider increased comfort, along with communication and preemptive management of problems.	• TapCloud's facilitation of direct, efficient contact with patients made it easy to use. • Patients and caregivers were particularly enthusiastic about how easy medication refills are with TapCloud. • Patients and caregivers find the TapCloud application intuitive, easy to use, and not time-intensive.	• Once coached on using the application, patients readily adopted the technology and often felt a sense of accomplishment in doing so.
Browning et al [27]	Mean age 81 years	Residents of rural Southwest Virginia	Retrospective quality improvement case series	Phone-based telehealth monitoring	• Participants felt nurses had better accessibility to vitals (88.9%). • 90% felt a better connection with their doctor.	• 77.8% found telehealth easy to use.	• Not applicable.
Collie et al [28]	Mean age 60.7 (SD 9.24) years	Resident of Intermountain region of Northeastern California	Pretest-posttest	Videoconferencing	• Not applicable.	• Participants adapted quickly to videoconferencing and reported no communication difficulties. • The facilitator faced challenges due to time lags and poor lighting.	• All said they would recommend it to other women with breast cancer.
Cummings et al [29]	Mean age 52.8 (SD 16.2) years	Rural counties in eastern North Carolina	Descriptive	Nonmydriatic retinal imaging telemedicine system	• 85% of images were "good or fair" when examined by a retinal specialist. • Retinal specialists reported "very certain or certain" in 84% of diagnoses.	• 96.3% of the participants were "very comfortable" or "comfortable" with the portable camera.	• Not applicable.
DeHart et al [30]	41.2% of the sample is 55 years or older	Rural resident from South-Eastern State	Mixed methods	Web and mobile-based telehealth with remote monitoring	• The application effects were seen through the increased coordination across agencies, technology fit with intraoffice demands, growing practice through engagement, flexible delivery, reduced provider travel, geographic access, reduced patient transport, saving time/visits for patients, reduced costs, and fewer challenges to sustainability.		• The most common patient-level challenge reported by providers was a lack of buy-in to use the service consistently.

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
						• Telehealth had organizational technology and space challenges. • Patient skills and comfort were a challenge. • Provider knowledge and skills were a challenge. • Rural patients struggled with technology access, often due to a lack of broadband technology in homes, but also due to technology skills for aging populations and those with less education.	
Demaerschalk et al [31]	Mean age 66.3 (SD 13.5) years	Patients at rural medical centers, more than 185 miles from a primary stroke center	Randomized control trial	Telestroke	• Results indicate effectiveness in the organization and structure of stroke telemedicine networks for extending stroke care in rural communities.	• Technical problems were noted in 20 of 27 (74%) of telemedicine consultations and in 0 of 27 telephone consultations. However, none of the issues kept the patient from being cared for.	• Not applicable.
Depatie and Bigbee [26]	Senior living facility serving adults aged 60 years or older	Rural northern California senior centers	Descriptive mixed methods	Mobile health technology		• Participants' comfort level with using email or the Internet to communicate with a health care provider was evenly split.	• Responses to interest in incorporating technology into daily life for health tracking and communication with a health care professional indicated that 33% had no interest, 23% were somewhat interested, and 13% indicated that they already used this technology.

Study	Population characteristics: Age	Population characteristics: Rural	Study design	Technology	Usefulness	Ease of use	Intention to use
					• 67% felt it was moderately important for the technology to monitor their health, but do not feel comfortable sharing with their provider. • 53% felt it was somewhat or extremely important to monitor their health and were comfortable sharing information with their provider. • 46% felt it was somewhat or extremely important to monitor their health as well as share information over the internet with their health care provider. • 44% indicated it was moderately, somewhat, or extremely important to use mobile health technology in combination with in-home nurse visits. • 40% felt it was somewhat or extremely important to connect patient education and support groups online. • 77% felt health technology had a clear benefit on their health.		
DeVido et al [32]	Mean age 54 (SD 19.4) years	Providers at a rural population-serving hospital	Case report	Telepsychiatry	• Telepsychiatry is capable of responding to many consultation questions in a hospital setting. • Cognitive test items were collected reliably using a camera function that interpreted physical images. • Telemedicine gave valuable clinical data regarding psychotic symptoms. • Telepsychiatry was applicable in accessing a diverse range of patients. • Communication between the resource nurse and referring physician was aided through telemedicine.	• Technological challenges included getting the cameras and software to operate reliably. • An ongoing challenge was the integrity of the image and audio. • Audio feed was compromised by Wi-Fi connectivity problems.	• Not applicable.

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Donahue et al [33]	Mean age 57.9 (SD 12.4) years	Rural North Carolina county	Cohort	Phone-based digital health care	• 84% of participants set a goal and reached at least one goal due to digital health technology.	• Not applicable.	• Providers expressed interest in continuing the phone coaching program with their participants, if resources were available. • 96% would recommend phone coaching to others. • 58% of participants remained engaged with phone coaching over the 12-month period (missed fewer than 3 consecutive monthly calls), 17% were less engaged (received at least one call), and 25% were not engaged.
Finley et al [34]	82% of the sample is 65 years or older	Three rural outpatient telehealth clinics, or one urban outpatient clinic in Arizona	Mixed methods	Telecardiology	• Patients perceived service as competent and began to see benefits in accessibility, reduced transportation, and general timesaving.	• Not applicable.	• When compared with the telehealth group, in-person participants more frequently expressed concerns about technology, about care not being face-to-face, and made more negative statements about quality. • Rural participants held the most positive attitudes toward telehealth, while suburban participants held the least positive attitudes toward telehealth.
Geller et al [35]	54.6% of the sample is 61 years or older	Patients and their providers from rural practices in Vermont	Quasi-experimental	Computer-based interactive interventions			• Not applicable.

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
					• Preliminary evidence that the technology could educate patients about screening. • Technology promoted patient-provider discussion, provider recommendations, and positive patient intentions to get a screening.	• Digitally everyone found the program easy to use. • Older participants and less educated participants found the sound (reading the information and questions on each screen) helpful compared with the younger and more educated participants.	
Gutierrez et al [36]	Mean age 65.2 years	Patients from a VA[a] hospital in rural Wisconsin	Mixed methods	Telehospitalist	• Communication benefits • Telehospitalists reported confidence that the diagnosis accuracy and quality were that of in-person. • Patient satisfaction showed improvement in care coordination (18%; P=.02).	• Connectivity problems were prevalent, although most providers were able to resort to a backup plan. • Internet connectivity was inconsistent, leading to disruption in video communications. • It was easy to contact bedside providers/telehospitalist: 100% telehospitalist, 42.9% physicians, 10% nurses, and 22.2% other staff.	• Not applicable.
Hatch et al [37]	47.4% of the sample is 65 years or older	33.1% of the sample resided in a rural area	Descriptive cross-sectional cohort	Telehealth	• Not applicable.	• Older and low-education patients are less familiar with the required technology and experience some technological limitations.	• Not applicable.
Hicks et al [38]	Mean age 68.9 years	Residents in rural midwestern state	Experimental	Telemonitoring	• 78.3% indicated that the use of telehealth technology improved their care. • 78.3% indicated it was very convenient. • 95.7% of the respondents indicated that the telehealth technology affected their relationship with their nurses positively.	• 95.7% indicated it was very easy to communicate with the agency personnel using the equipment. • 95.7% of the respondents indicated that the telehealth technology was very easy to use. • 87.0% indicated that the telehealth technology worked very well.	• 91.3% of the respondents indicated definitely and 8.7% indicated maybe when asked if they would use the digital health technology again.
Holloway et al [39]	Mean age 61.3 (SD 11.6) years	Residents in rural Montana	Pretest-posttest	PRISM[b] digital health care videoconferencing			• Not applicable.

Study	Population characteristics: Age	Population characteristics: Rural	Study design	Technology	Usefulness	Ease of use	Intention to use
					• Patients learned to work with the technology, enabling them to interact effectively with patients at other sites. • Participants felt that the ability to participate in a team approach to diabetes management without traveling was beneficial. • Compared with baseline, patients reported improvements in diabetes care of 30-200% 1 year after the intervention.	• Scheduling was a difficult and time-consuming task. • Staff members easily learned to use telehealth technology. • In most cases, patients were able to learn manual skills using telehealth technology. • 99% felt the technology picture and sound were clear. • 100% of patients felt comfortable learning health information using this technology and said they understood the information as if it were imparted in person.	
Khairat et al [40]	Mean age 77.8 years	66.7% of the sample resided in rural North Carolina	Cross-sectional	Videoconferencing follow-up care	• Geriatric patients did not have trouble using digital health tools. • Telemedicine platforms improve primary care by allowing providers to follow-up with their geriatric patients at a time and place that is most convenient for both groups.	• Not applicable.	• Not applicable.
Kulcsar et al [41]	Mean age 60.3 (SD 16.1) years	The majority of patients resided in rural New Hampshire or rural Vermont	Cross-sectional quality improvement	Telerheumatology	• Providers reported that 19% of the patient visits seen via Telerheumatology were inappropriate for the telemedicine visit type due to poor understanding of symptoms, symptom complexity, and the limited ability to perform a physical examination. • 81% of patients rated that they were comfortable with the provider's ability to examine them, thought the provider spent an adequate amount of time with them, and made an accurate diagnosis of their condition.	• 94% of patients felt that each individual member of the staff made check-in easy, was friendly, and competent with the videoconferencing equipment. • An area of greatest dissatisfaction for patients stemmed from problems with scheduling appointments (usually follow-up).	• About half (53%) of the patients either agreed or strongly agreed that they would like to be seen via telerheumatology again if given the option.
			Qualitative				

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Liu et al [42]	Mean age 67 years	Patients from a clinic in rural Wisconsin		Teleophthalmology	• Teleophthalmology was often preferred by patients because it was more comfortable than traditional eye examinations, since pharmacological pupil dilation was usually not needed. • Teleophthalmology is effective for increasing diabetic eye screening rates in rural populations.	• Tele-ophthalmology offered convenience through same-day scheduling, proximity to the patient's primary care provider, and short wait times. • Patients appreciated the quick, easy, and painless nature of tele-ophthalmology compared with traditional, in-person eye examinations.	• Participants reported being unaware of tele-ophthalmology prior to the study.
Locke et al [43]	Mean age 69.2 years	Residents of rural zip codes defined by the United States Census Bureau	Retrospective	Home computer video health technology	• 96% of participants preferred video telehealth rather than traditional training visits at the medical center. • 76% of participants would not have gotten additional inhaler training if not for the telehealth. • The telehealth inhaler training delivered via internet video telehealth demonstrated an improvement in technique overall. • The main benefits, as listed by participants of the program, were convenience, time-saving, and decreased travel expenses. • Inhaler training delivered via video telehealth by a pharmacist was well received. • The CHAT inhaler training program provided an alternative to in-person visits for rural patients with transportation burdens.	• Among 93 home telehealth program enrollees, 19 (20%) faced technical issues with the computer or video software that prevented participation. • A quarter of participants reported frequent technical problems, consistent with pharmacists noting issues in 149 (63%) of scheduled visits, and 19 visits (13%) were postponed or partially completed due to unresolved technical issues. • Common issues included patient errors or confusion (41% of visits) with the video telehealth program, particularly with logging into the Jabber program and basic computer skills, along with 11% experiencing computer/software issues and 25% having audio/video troubles. • Despite these challenges, over 90% of participants found the equipment easy to set up and appreciated the benefits of convenience, decreased travel time and expenses, and increased privacy.	• A majority (96%) preferred home video telehealth for inhaler training compared with going to the medical center for in-person training.
McIlhenny et al [44]		Patients of rural medical clinics	Quasi-experimental	Computer-based telemedicine	• Not applicable.		

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
	Mean age 63.8 (SD 12.46) years					• Nearly all participants found the information easy to access.	• Participants responded positively to the individualized education provided by the nurse educator and did not prefer receiving information from the internet.
Owolo et al [45]	Mean age 59.9 (SD 13.5) years	28.9% of the sample was considered rural	Retrospective cohort	Web-based telemedicine	• Patients who used the patient portal had 5.21 times higher odds of completing a video visit compared with patients who did not use the patient portal (95% CI: 1.28-21.23; P=.022). • The majority of telehealth visits were conducted over the phone; however, there was an increase in video visits in the post–initial surge period.	• 81.9% of physicians agreed that telemedicine was easy to use with a preference for imaging review, initial appointments, and postoperative care. • While telephone visits were still used at higher rates than video visits, the increased use of video visits potentially reflects a better organized infrastructure for performing this type of visit.	• Not applicable.
Robinson et al [46]	Mean age 71 (SD 6.8) years	Residents from rural New York	Cross-sectional	Telemedicine	• 72% found the in-home nurse visit “very helpful,” in contrast to 55% for telephone tutoring and 46% for the 39-page user’s manual. • 37.8% of the participants wanted additional telephone training to access the web.	• Twenty-two subjects (63%) rated the amount of training as “3” (about right) compared with their initial expectation. • Four responders (11%) rated the training lower than “3” (less than expected), while 9 participants (26%) rated it higher than “3” (more than expected).	• Not applicable.
Rodriguez et al [47]	Mean age 63 (SD 12) years	Patients and providers from rural medical centers and remote community-based outreach clinics in Pennsylvania	Mixed methods	Electronic consultations	• Improved communication, as it enabled effective information transfer and patient-centered care. • For PCPsc, time efficiency was the main reason for telehealth satisfaction. • E-consults improved access to specialty care, saving travel time, and enabling confident care.	• Not applicable.	• All patients indicated that they intended to use e-consults in the future. • The intent to use e-consults in the future focused primarily on quality of care and timeliness of care.
Schlittenhardt et al [48]	Mean age 54 years	78% of the sample from rural areas	Mixed methods	Tele-Continence care		• Tele-Continence Care implementation was uncomplicated.	

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
					• The option of telehealth as a follow-up appointment reduced the overall failure rate from 45% to 14.3%. • Patients appreciated the convenience and reduction in transportation costs.		• When asked if face-to-face visits would be preferred over telehealth visits, patients described a neutral opinion.
Schooley et al [49]	Mean age 63.5 years	Residents of rural Vermont	Mixed methods	Information technology mediums	• Telehealth programs, such as telephone triage, are accepted as an option to receive remote care and avoid travel. • The community care home telehealth enabled physiological data to be monitored remotely through landline phones amid staffing problems. • Email, computers, and assistive devices were imperative to aiding veterans with disabling conditions to communicate effectively while remote.	• Assessing mental health symptoms was challenging through telephone and email, as these methods do not allow for observing nonverbal cues, unlike advanced videoconferencing.	• 56.5% of respondents reported some level of likelihood of consulting with a doctor over the phone, compared with 13.2% who reported some level of likelihood of using the Internet to consult with a doctor. • 35% had an interest in communicating with their doctors and nurses via the Internet about their health care. • Younger veterans (age 41-55 years) were more likely than older veterans to report interest in the program.
Silvestrini et al [50]	Mean age 60 years	Patients referred from clinics in rural Washington, Oregon, or Alaska	Qualitative	TelePain	• TelePain reduced transportation barriers and travel costs. • Patients enjoy the convenience of its general accommodations.		

Study	Population characteristics: Age	Population characteristics: Rural	Study design	Technology	Usefulness	Ease of use	Intention to use
						• Patients generally did not have significant issues using telehealth technology, with many veterans rating TelePain video and audio quality as "good" or "fine." • TelePain staff effectively helped veterans navigate any technological problems, such as resolving audio issues by having the provider call the patient. • Some patients suggested improvements for telehealth, such as larger TVs, better monitors, or improved camera systems to enhance the video quality and overall experience.	• Many of these patients mentioned that they would like to continue to use TelePain because of its convenience, rather than having to travel far distances to receive pain care. • When asked if they would use TelePain again, one patient said, "Oh, of course. It's much easier than going clear to [the VA medical center]." Another patient replied, "Yeah, probably. It's easier for me than driving to [the VA] and back."
Strowd et al [51]	Mean age 44.5 (SD 24.1) years	26% of the sample resided in a rural zip code	Retrospective cohort	Video and phone-based digital technology	• 73% of respondents reported that clinical needs were met with the telehealth visit. • Patients completing the video visits were more likely to have needs met than those who did telephone-only visits (77% vs 71%, P=.34). This was further seen in both urban and rural communities (71% vs 80%, P=.27).	• The most common patient-reported barrier to scheduling a video visit was patient technology-related (44% of patients), which included lack or limited access to a smartphone or home computer (n=59), no camera for video (n=51), no internet availability (n=27), or other (n=12).	• 45% of the respondents would definitely consider a future telehealth visit, 28% might consider, 24% would only consider if required. • Patients who completed video visits were more likely to definitely consider a future telehealth visit compared with patients who completed a telephone-only visit (58% vs 38%, P=.02). • Patients from rural communities were more likely to definitely consider a future telehealth visit compared with those from urban communities (55% vs 42%, P=.05). • Not applicable.

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
Svistova et al [17]	A large portion of the sample is aged 65 years and older	Rural residents in Pennsylvania	Qualitative focus group	Zoom teleconferencing	• Technology has allowed service providers to adapt quickly to COVID-19. • The effects of telehealth eliminated the transportation barrier and reduced cancellations and missed appointments.	• Older adults faced significant barriers using telehealth due to unfamiliarity with technology, lack of recent devices, low comfort levels, and distrust leading providers to have fewer positive insights about serving them during COVID-19. • An insurance provider noted that older clients often lack the proper technology or skills for telehealth. • Using technology for health care communication is often uncomfortable for older adults.	
Switzer et al [52]	Mean age 70 years	Patients from one of 12 hospitals, 10 of which were in rural Georgia	Cohort	Telestroke	• Telestroke helped enroll patients into the acute stroke treatment program. • The initiation and treatment period for Telestroke patients was quicker compared with in-person patients.	• Not applicable.	• Not applicable.
Virmani et al [53]	Mean age 65.8 (SD 9.2) years	Residents in rural Arkansas	Cross-sectional	Web-based televideo	• 82% of participants reported televideo had a positive effect on their in-home visits. • 70% of participants reported that they liked that there were no travel arrangements. • 84% of participants reported that they liked the ability to stay in the comfort of their home. • In rural areas, the quality of the audio-video connectivity was enough to implement the routine clinical and research assessments.	• 28% preferred for in-person visits. • Audio-video quality was rated great for 60% of visits, with video slightly slow for 30%. • Only 14% of participants needed more than 5 minutes for setup; 42% needed additional time for understanding, locating, or completing research assessments.	• Most patients reported being more likely to participate in telemedicine research in the future. • 28% preferred for in-person visits.
Waymouth et al [54]	Providers of older adults	Providers in a rural area	Qualitative	Assistive and remote monitoring technology		• Not applicable.	• Not applicable.

Study	Population characteristics: Age	Population characteristics: Rural	Study design	Technology	Usefulness	Ease of use	Intention to use
					• Telehealth helped facilitate care for individuals with dementia through the use of assistive technology and remote monitoring technology such as fall sensors, webcams, alarm systems, GPS tracking, and smart home hubs.		
Weiner et al [55]	Mean age 69.7 (SD 12.8) years	Residents of Choctaw Nation of Oklahoma	Clinical trial	Videoconferencing	• Virtual care was used to overcome transportation burdens.	• Less than 3% no-show rate indicates the technology was easy to use.	• There were no refusals of videoconferencing for initial visits and 2 refusals of continued videoconferencing follow-up. • The no-show rate for all videoconferencing sessions in the past year was 3%.
West and Milio [56]	Mean age 68.3 years	Patients from a rural homecare organization	Mixed methods case study	Rural homecare organization telemedicine	• There were frequent issues with the telehealth device that caused disruptions. • Restrictions were seen in outdated telecommunications equipment, which impaired the transmission of audio and visual information. • An increase in productivity and lower costs needs to be implemented to meet the demand for these services.	• Informants reported that the complexity of the setup and the number of components confused and intimidated the nurses and patients. • Faulty equipment also caused frequent disruptions during telemedicine visits.	• Younger patients are more likely to receive telemedicine visits than those receiving traditional home care.
West et al [57]	55 years or older	Residents of rural upstate New York	Cohort clinical trial	Web-based telemedicine with videoconferencing	• The IDEATel project demonstrated the home televisits through videoconferencing were possible for rural underserved elderly adults with the right home education and behavior goal setting.	• Many participants initially had difficulty using the unit and required additional instruction.	• Not applicable.
Zulman et al [58]	Mean age 56 (SD 17) years	53% of the sample as rural	Mixed methods	Tablet-based health technology	• The VA's use of video telehealth tablets reached rural older adults effectively.	• Not applicable.	

Study	Population characteristics		Study design	Technology	Usefulness	Ease of use	Intention to use
	Age	Rural					
							• 81% of tablet recipients used their tablets during the evaluation period. • Tablet recipients were more likely to use their tablets if they were 45-64 years or 65 years old (compared with < 45 years). • Patients who did not use the telehealth tablet were largely younger with more chronic conditions and a lack of social support.

[a]VA: Veterans Health Administration.

[b]PRISM: Promoting Realistic Individual Self Management.

[c]PCP: primary care provider.

Quality Appraisal

The quality assessments were achieved through critical appraisal tools used in Joanna Briggs Institute Systematic Reviews [59] and the McGill Mixed Methods Appraisal Tool Version 2018 [60]. Two reviewers assessed the quality of each included article represented in the extracted data. Weekly meetings were held to discuss conflicts. Studies with a quality score of less than 50% were not included in the analysis, indicating evidence of reporting bias.

Results

Overview of Reviewed Studies

Figure 1 provides a summary of the selection process and illustrates our systematic review of the literature. The preliminary title review produced 7728 results, out of which 7616 underwent the abstract review, and 511 completed the full-text review. Following the assessment, 16 studies were excluded because they scored less than 50% of the total possible score for each respective quality test. Our final review included 39 studies.

Figure 1. PRISMA flow diagram.

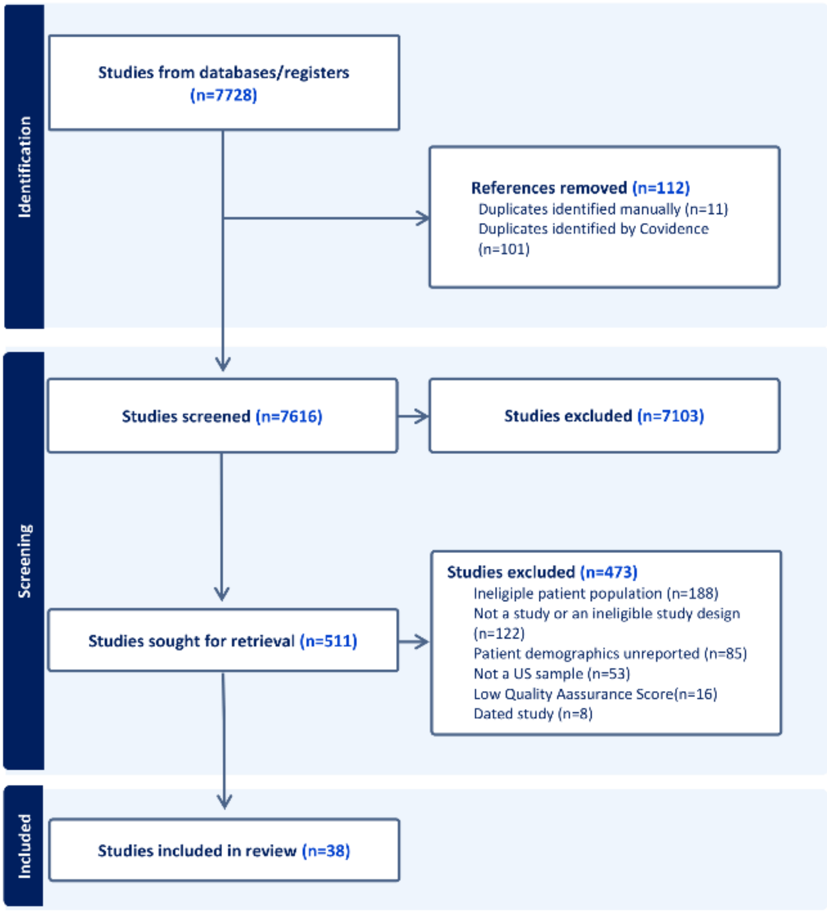


Table 1 details our data extraction, elucidating article information according to the first author and publication date, study population age, study design, technology used, and TAM domains. Since 25% of the population had to be at least 60 years old for eligibility, it is important to highlight that studies reported the mean age, included samples exclusively of older adults (eg, Medicare enrollees), and provided the percentage of participants within specific age ranges. The study design of eligible studies was cross-sectional (n=8), qualitative (n=7), and mixed methods (n=11). We will detail lessons learned from the eligible studies according to the TAM domains below.

Technology Usefulness

Digital health tools were useful for connecting patients to providers [25,26,35,38,54,61], assisting participants with improving health outcomes and care management [25,31,38-43,47-49,51-54,57,61], and reducing transportation burdens [17,23,30,34,43,47-50,53,55]. Uniquely, Anderson et al [23] reported on the role of telehealth videoconferencing in strengthening patients' social and peer connections. Providers also considered that digital health tools were useful for care management [24,29,32,36,40,49] and communicating with other providers [30,32,47]. Although the respective technology was considered useful, multiple studies reported that rural older adults needed additional supports to facilitate the usefulness of the technology, including behavioral goal setting [57] and assistive devices [49].

Despite the reported usefulness of digital health tools, some studies also reported that they were less than useful. Bernacchi et al [25] revealed that digital health technology was less useful for participants with limited experience with digital technology. Similarly, according to Kulcsar et al [41], nearly one in 5 patients scheduled for a telerheumatology visit was deemed unsuitable due to poor symptom understanding, symptom complexity, and the limitations of conducting a physical examination remotely.

Ease of Use

Articles that addressed the digital health tools' ease of use described the tools as uncomplicated [48], requiring an adequate amount of training to use [62], easy to access health information [39,44], easier than an in-person examination [42], easy to communicate with a health care professional [36,38,61], and intuitive [26]. Hybrid models that provided supplemental real-time instruction with the technology improved the ease of use [41,50]. Importantly, some digital health tools were easier to use than others. For example, Schooley et al [49] assessed that mental health evaluations via email and telephone were challenging because of barriers to observing nonverbal cues, yet were easier to accomplish via a videoconference. Additionally, Geller et al [35] reported that audio features improved the ease of using a computer-based intervention for older adults. Providers also reported that digital health tools improved the ease of managing appointments [24,45], accessing and reviewing patient records [24,45], and contacting patients [26]. Weiner et al [55] reported that a low "no-show" rate (3%)

evidenced the ease of using videoconferencing visits. Locke et al [43] reported that 90% of the providers reported that the technology was easy to use, despite the fact that more than half of the providers reported issues with scheduling, and 41% reported that patients were confused about using the technology.

Commonly, studies have reported on the challenges of using technology. Digital health tools were more complicated to use when faced with environmental barriers (eg, poor lighting) [23,28], technical or equipment issues [17,25,30-32,36,43,51,53,56], discomfort with the technology [17,25,30,56], reduced access to adequate internet [30,32,36,51], and distrust in the technology [17]. DeHart et al [30] and Hatch et al [37] observed a positive relationship between education levels and ease of use, which was particularly challenging for older adults with low levels of education. Although many studies reported the challenges of using digital tool technology, West et al [57] reported that supplemental instructions were critical for overcoming usage barriers.

Intention to Use the Technology

Most commonly, participants reported an intent to use the digital health tool technology through sustained use poststudy [34,38,41,47,51,53], future referrals or recommendations [28], high levels of digital tool uptake [43,50,55,58], and evidence of improved care management [25,43,47,50]. For example, Locke et al [43] reported that 96% of the participants preferred home video telehealth inhaler training rather than going to the clinic for in-person training. Bernacchi et al [25] detailed another example of intention to use videoconferencing technology through the patient's commitment to contact their health care provider despite challenges with equipment and broadband. Strowd et al [51] and Finley et al [34] reported that rural residents were more likely to consider telehealth in the future compared with urban residents. A study indicated that participants were more likely to continue with specific delivery modes of digital health tools, such as telephones, rather than web-based options [49].

When participants did not intend to use the digital health tool technology, it was often due to a preference for in-person visits [24,34,44,45,53,56] or a lack of buy-in about the technology [30,42]. Additionally, 2 studies identified age associations with intentions to use digital health tools. Older adults had reduced intentions of continuing their care digitally [49,51].

Discussion

Principal Findings

This review aimed to assess lessons learned about the acceptability of digital health tools among rural older adults. Following a systematic review approach, we organized our findings according to the TAM, focusing on the usefulness, ease of use, and intention to use digital health tools of rural older adults. The domains of the TAM aim to detail predictors of potential acceptance or rejection of the technology. Our findings revealed that digital health tools were, in most cases, useful for care management, reducing transportation burdens, and improving patient-provider communication. Two articles reported that digital health tools were less useful when the technology was misaligned with the participant's digital skills level or when the technology was unsuitable for the given care visit types. Several articles highlighted the ease of use of digital health tools for rural older adults, describing them as uncomplicated, intuitive, and effective for connecting with health care providers. Most articles that discussed the ease of using digital health tools focused on the ease of specific features of the tool (eg, audio capabilities) or the type of technology (eg, telephone vs digital). However, digital health tools were more difficult to use due to technical or equipment issues, discomfort with the technology, and limited access to broadband internet. Last, the findings on rural older adults' intention to use digital health tools were robust, as evidenced by participants' preference to continue using the technology after the study concluded and improved health outcomes. Comparatively, participants did not intend to use the digital health tools when they preferred an in-person visit or when they were not sold on the benefits of the digital health tool. Together, the TAM domains reveal that rural older adults and their providers largely consider digital health tools as acceptable modes of receiving care and, at times, a suitable alternative to in-person clinic visits. Despite barriers that reduced rural older adults' acceptance of digital health tools, many of these barriers were not associated with their age or rural residence.

Technology Usefulness: Lessons Learned

Digital health tools are useful for accessing health care and care management, but their effectiveness in improving health outcomes in older adults is mixed. Our review highlights that digital health tools were useful for mitigating burdens related to accessing health care and care management for both providers and patients. Articles reported that useful care management needs included scheduling, accessing health records, and patient-provider communication. The positive findings on remote care management and usefulness are specific to this review, and it is important to note that the effectiveness of remote care management and monitoring on health outcomes is mixed. In a review of remote care management of depression and anxiety in older adults, the findings on psychiatric outcomes were mixed, and no studies demonstrated a statistically significant effect of remote care management on health care use or cost [63]. Another review of mobile integrated health interventions for older adults revealed that these interventions reduced emergency department call volume and transports [64]. Thus signaling that digital health tools were useful for care management during emergency health events.

To build on this body of evidence, our review uniquely emphasizes the usefulness of digital/remote care management for rural older adults—a population that has complex care needs but often resides in a medically underserved area with reduced access to broadband internet and technology literacy programming. However, future systematic and meta-reviews are needed to assess the effectiveness of digital health tools with care management features on health outcomes and costs for rural older adults.

Ease of Use: Lessons Learned

Easy-to-use technology is associated with improved health outcomes; however, the design of technology may not be sufficient, and external support (eg, timely technical assistance) may be necessary. According to the TAM, digital tools that are perceived to be easy to use are more likely to be accepted by the intended audience. Based on this review, rural older adults and their providers frequently highlighted user-friendly features of the tools that improved ease of use. However, there was strong evidence that external factors—such as technical issues, equipment limitations, and discomfort with the technology—hindered usability. There is a positive relationship between the health of older adults, their social connectedness, access to high-quality health care resources, and the perceived ease of use of digital health tools [46,65]. Notably, this evidence signals that technology design alone does not improve the ease of using digital health tools. Specifically for older adults, in-person synchronous technical assistance, access to remote technical assistance, and early interventions from hospital administrators are reported facilitators for increasing the ease of use of digital health tools [15,25,29]. Overall, this evidence underscores the need for additional resources and external support to enhance the perceived ease of using digital health tools for rural older adults.

Intention to Use the Technology: Lessons Learned

Despite design flaws or technical difficulties, rural older adults generally intended to continue using technology beyond the observed period. Factors such as the preference for in-person care and a lack of buy-in about the technology influenced the participants' intent to use digital health tools. Yet, our findings revealed that participants described an intention to use digital health tools despite also reporting that the technology was not always useful or easy to use [25,34,41,43]. In a similar study on patient portal use by older adults, challenges were noted with log-ins and the user interface design, such as color and font [62]. Despite these issues, older adults expressed an intention to continue using the portals due to their other beneficial features. In summary, in light of the barriers to ease of use and usefulness, rural older adults often overcame them to continue using digital health tools [25].

Additional Considerations for Digital Tool Acceptability

In synthesizing lessons learned, our review identified several phenomena, not salient enough to categorize as a lesson, but worthy of continued discussion. Namely, our review reveals both differences and commonalities in user behavior and preferences between older adults in rural and urban areas. As reduced access to broadband internet is a common barrier for rural residents, this was not the most reported impediment to digital tool acceptability in this review, as expected. According to the Federal Communications Commission's data reported in 2021, 23%-50% of rural residents had poor access to broadband internet [66]. The rather limited mention of challenges associated with rural internet connectivity, in this review, is inconsistent with the existing literature, which indicates that poor internet connection is a key barrier to digital tool use acceptability for rural residents [15]. Also notable is that multiple studies reported a measure of digital tool acceptance among rural residents compared with urban residents. For example, Finley et al [34] reported urban-rural differences in intention to use, indicating that rural patients, compared with urban and suburban patients, had more favorable attitudes toward telecardiology. This higher acceptance of digital health tools by rural patients likely punctuates the growing reliance on and acceptance of digital health tools in the wake of dwindling local health care resources and rising health care costs. While these findings suggest that rural and urban residents share common barriers to technology acceptability, our conclusions do not suggest that "one-size-fits-all" interventions should be considered. Rather, additional qualitative examinations on the rural-urban differences in attitudes toward digital health tools and the acceptability of these tools are warranted.

This review synthesized a diverse representation of technology modes (eg, videoconference or phone call), highlighting the robust intervention designs implemented in rural settings. These findings strengthen the evidence base on digital tool acceptability among rural patients, with all modes being reported as acceptable. Yet, important distinctions emerged across the reported technology modes, making salient conclusions about the most acceptable modes speculative. For example, both Anderson et al [23] and Svistova et al [17] used videoconferencing; however, Anderson et al [23] employed a Veterans Affairs-supported platform, and Svistova et al [17] used Zoom. The journal articles provided limited details regarding platform-distinctive features, though such distinctions could plausibly impact acceptability and usability very differently. In the current review, technology modes were extracted as they were identified in the original article to ensure transparency. A more detailed analysis of the technology's distinctive features and their impact on acceptability warrants further investigation.

The data collected for this study includes both pre– and post–COVID-19 pandemic publications, which provide key insights into how rural older adults' acceptance of digital health tools evolved, resulting from the rapid uptake of telehealth due to COVID-19 precautions. Many of the studies that were published prior to the COVID-19 pandemic often emphasize the same sentiment of the digital divide between younger and older generations. The existing literature evidences that the pandemic exacerbated this divide, as prepandemic older adults were less likely to benefit from technological innovations [67]. Specifically, pre- and early-pandemic trends indicated that age and rural zip codes were inversely related to continuous digital tool use [68]. Multiple studies from our review that were published prepandemic identified that younger participants were more likely to engage in digital tool technologies [49,56]. Yet, in a peripandemic investigation, Bernacchi et al [25] describe that videoconferencing with a nurse increased access to care for older rural cancer patients. Similarly, postpandemic, individuals older than 65 years used telemedicine over 3 times more when compared with prepandemic, further emphasizing this technological shift [19].

The findings of this review have practice and research implications. To improve practice, our findings suggest that health care providers adopt hybrid care models—combining both digital and in-person visits. It is essential for providers to emphasize the benefits of digital health tools, such as reducing travel burdens and offering greater convenience

[30,39,43,47,49,50,53]. Additionally, future research should include interventions aimed at increasing technology literacy among rural older adults and provide additional synchronous and asynchronous supports to help rural older adults better understand the technology [17,25,30-32,36,37,51,56,57]. For caregivers who support rural older adults, digital health tools have the potential to reduce caregiving burdens. Tools such as remote monitoring enable doctors and nurses to keep patient health under observation while the patient remains in the comfort of their home [30,38,54]. This technology could reduce the burden on caregivers and limit accidents. Last, it is our hope that these findings will influence policy that increases funding for the development of digital health tools that can decrease health disparities within rural older adult populations.

Limitations

This study has several strengths and limitations. The synthesis of our findings was guided by the core domains of the TAM [69]. Previous studies have reported that the Perceived Usefulness, Perceived Ease of Use, and Intention to Use domains are important for determining user acceptability. They are the core of all TAM models used across the literature, thus chosen for this study. Yet, additional domains and extensions have been added over time, including the "Attitudes Toward Using" domain [70,71]. Although our synthesis is limited to conclusions derived from the core domains, these domains provide sufficient information to inform our overall aim of assessing user acceptability. Additionally, a trained librarian conducted a data search of publicly available databases (eg, PubMed). Despite our comprehensive search conducted by a trained librarian, some relevant studies may have been missed due to factors such as publication dates postreview, non–English language restrictions, or potential oversight in keyword selection. It is important to note that participants in health outcomes research are often younger than 65, have higher household incomes, greater technological literacy, and are more likely to reside in urban areas [27,33,72]. Therefore, our review's focus on rural older adults and digital tool use may highlight a potential participant bias in our sample, limiting the generalizability of our findings to the broader rural older adult population. As with all scholarly reviews, researcher bias could have impacted the results. However, because multiple reviewers assessed each manuscript and attended consensus meetings for each manuscript, this bias was reduced.

Conclusion

This study aimed to systematically review articles that incorporated digital health tools used by rural older adults in order to assess their acceptability and usage of the tools. Following the TAM, we highlighted the usefulness, ease of use, and intention to use digital health tools. In summary, digital health tools were valuable for rural older adults with complex care needs, helping mitigate access barriers and support care management tasks like scheduling and patient-provider communication. While rural older adults and providers found the tools user-friendly, external factors such as technical issues and equipment limitations impeded usability, signaling the need for additional support and resources. Despite challenges with ease of use, rural older adults expressed an intention to continue using digital health tools, recognizing their overall benefits in managing care, especially in underserved areas. As medical deserts widen in rural communities, and in response to the rapid uptake of telemedicine due to COVID-19 precautions, digital health tool reliance is likely to grow for rural residents. Understanding what this uniquely vulnerable population views as acceptable and what facilitates their uptake of digital health tools is critical for addressing health disparities and bridging the digital divide in rural communities.

Funding

MWL and ZMS were supported by a grant from the National Cancer Institute (Grant Number: K01CA262342). MWL is also supported by a Northwestern University Clinical and Translational Sciences Institute grant (Grant Numbers: NUCATS; UL1TR001422, Principal Investigator: Richard D'Aquilla), an Institutional Research Grant (IRG-21-144-27) from the American Cancer Society (Principal Investigator: Leonidas Platanias), and funds from the Northwestern University Center for Community Health (grant number: not applicable). Opinions and comments expressed in this paper belong to the authors and do not necessarily reflect those of the National Institutes of Health.

Authors' Contributions

ZMS was responsible for formal analysis, methodology, project administration, visualization, writing the original draft, and reviewing and editing subsequent drafts. EQ was responsible for the formal analysis, writing the original draft, and reviewing and editing subsequent drafts. JP was responsible for writing the original draft and reviewing and editing subsequent drafts.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA checklist.
[PDF File (Adobe PDF File), 140 KB - aging_v9i1e70012_app1.pdf]

References

1. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022;49(4):517-530 [FREE Full text] [doi: 10.1016/j.pop.2022.04.002] [Medline: 36357058]
2. HHS announces $16.3 million to expand telehealth care in the title X family planning program. Office of Population Affairs. 2022. URL: https://opa.hhs.gov/about/news/grant-award-announcements/hhs-announces-16-million-expand-telehealth-care-title-x [accessed 2025-12-16]
3. Haque SN. Telehealth beyond COVID-19. Psychiatr Serv 2021;72(1):100-103. [doi: 10.1176/appi.ps.202000368] [Medline: 32811284]
4. Koonin LM, Hoots B, Tsang CA, Leroy Z, Farris K, Jolly T, et al. Trends in the use of telehealth during the emergence of the COVID-19 pandemic - United States, January-March 2020. MMWR Morb Mortal Wkly Rep 2020;69(43):1595-1599 [FREE Full text] [doi: 10.15585/mmwr.mm6943a3] [Medline: 33119561]
5. Asante D, McLachlan CS, Pickles D, Isaac V. Understanding unmet care needs of rural older adults with chronic health conditions: a qualitative study. Int J Environ Res Public Health 2023;20(4) [FREE Full text] [doi: 10.3390/ijerph20043298] [Medline: 36833993]
6. Skoufalos A, Clarke JL, Ellis DR, Shepard VL, Rula EY. Rural aging in America: proceedings of the 2017 connectivity summit. Popul Health Manag 2017;20(S2):S1-S10 [FREE Full text] [doi: 10.1089/pop.2017.0177] [Medline: 29251548]
7. Eastman MR, Kalesnikava VA, Mezuk B. Experiences of care coordination among older adults in the United States: evidence from the health and retirement study. Patient Educ Couns 2022;105(7):2429-2435 [FREE Full text] [doi: 10.1016/j.pec.2022.03.015] [Medline: 35331572]
8. Probst J, Eberth JM, Crouch E. Structural urbanism contributes to poorer health outcomes for rural America. Health Aff (Millwood) 2019;38(12):1976-1984. [doi: 10.1377/hlthaff.2019.00914] [Medline: 31794301]
9. Cyr ME, Etchin AG, Guthrie BJ, Benneyan JC. Access to specialty healthcare in urban versus rural US populations: a systematic literature review. BMC Health Serv Res 2019;19(1):974 [FREE Full text] [doi: 10.1186/s12913-019-4815-5] [Medline: 31852493]
10. Sekhon H, Sekhon K, Launay C, Afililo M, Innocente N, Vahia I, et al. Telemedicine and the rural dementia population: a systematic review. Maturitas 2021;143:105-114. [doi: 10.1016/j.maturitas.2020.09.001] [Medline: 33308615]
11. Hodge H, Carson D, Carson D, Newman L, Garrett J. Using internet technologies in rural communities to access services: the views of older people and service providers. J Rural Stud 2017;54:469-478 [FREE Full text] [doi: 10.1016/j.jrurstud.2016.06.016]
12. Call VRA, Erickson LD, Dailey NK, Hicken BL, Rupper R, Yorgason JB, et al. Attitudes toward telemedicine in urban, rural, and highly rural communities. Telemed J E Health 2015;21(8):644-651. [doi: 10.1089/tmj.2014.0125] [Medline: 25839334]
13. Kruse C, Fohn J, Wilson N, Nunez Patlan E, Zipp S, Mileski M. Utilization barriers and medical outcomes commensurate with the use of telehealth among older adults: systematic review. JMIR Med Inform 2020;8(8):e20359 [FREE Full text] [doi: 10.2196/20359] [Medline: 32784177]
14. Abdallah L, Stolee P, Lopez KJ, Whate A, Boger J, Tong C. The impact of COVID-19 on older adults' perceptions of virtual care: qualitative study. JMIR Aging 2022;5(4):e38546 [FREE Full text] [doi: 10.2196/38546] [Medline: 36054599]
15. Hunter I, Lockhart C, Rao V, Tootell B, Wong S. Enabling rural telehealth for older adults in underserved rural communities: focus group study. JMIR Form Res 2022;6(11):e35864 [FREE Full text] [doi: 10.2196/35864] [Medline: 36331533]
16. Cho Y, Yang R, Gong Y, Jiang Y. Use of electronic communication with clinicians among cancer survivors: health information national trend survey in 2019 and 2020. Telemed J E Health 2023;29(6):866-874 [FREE Full text] [doi: 10.1089/tmj.2022.0203] [Medline: 36355055]
17. Svistova J, Harris C, Fogarty B, Kulp C, Lee A. Use of telehealth amid the COVID-19 pandemic: experiences of mental health providers serving rural youth and elderly in Pennsylvania. Adm Policy Ment Health 2022;49(4):530-538 [FREE Full text] [doi: 10.1007/s10488-021-01181-z] [Medline: 34846613]
18. Rush KL, Singh S, Seaton CL, Burton L, Li E, Jones C, et al. Telehealth use for enhancing the health of rural older adults: a systematic mixed studies review. Gerontologist 2022;62(10):e564-e577. [doi: 10.1093/geront/gnab141] [Medline: 34661675]
19. Lama Y, Davidoff AJ, Vanderpool RC, Jensen RE. Telehealth availability and use of related technologies among medicare-enrolled cancer survivors: cross-sectional findings from the onset of the COVID-19 Pandemic. J Med Internet Res 2022;24(1):e34616 [FREE Full text] [doi: 10.2196/34616] [Medline: 34978531]
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
22. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989;13(3):319-340. [doi: 10.2307/249008]

23. Anderson J, Godwin KM, Petersen NJ, Willson P, Kent TA. A pilot test of videoconferencing to improve access to a stroke risk-reduction programme for veterans. J Telemed Telecare 2013;19(3):153-159. [doi: 10.1177/1357633X13479703] [Medline: 23625913]
24. Barton AJ, Amura CR, Willems EL, Medina R, Centi S, Hernandez T, et al. Patient and provider perceptions of COVID-19-driven telehealth use from nurse-led care models in rural, frontier, and urban colorado communities. J Patient Exp 2023;10:23743735231151546 [FREE Full text] [doi: 10.1177/23743735231151546] [Medline: 36741820]
25. Bernacchi V, LeBaron VT, Hinton ID, DeGuzman PB. Rural cancer survivors' perceptions of a nurse-led telehealth intervention to manage cancer-related distress. Oncol Nurs Forum 2023;50(2):173-184. [doi: 10.1188/23.ONF.173-184] [Medline: 37677802]
26. Bonsignore L, Bloom N, Steinhauser K, Nichols R, Allen T, Twaddle M, et al. Evaluating the feasibility and acceptability of a telehealth program in a rural palliative care population: tapcloud for palliative care. J Pain Symptom Manage 2018;56(1):7-14 [FREE Full text] [doi: 10.1016/j.jpainsymman.2018.03.013] [Medline: 29551433]
27. Browning SV, Clark RC, Poff RM, Todd D. Telehealth monitoring: a smart investment for home care patients with heart failure? Home Healthc Nurse 2011;29(6):368-374. [doi: 10.1097/NHH.0b013e31821b7186] [Medline: 21633228]
28. Collie K, Kreshka MA, Ferrier S, Parsons R, Graddy K, Avram S, et al. Videoconferencing for delivery of breast cancer support groups to women living in rural communities: a pilot study. Psychooncology 2007;16(8):778-782. [doi: 10.1002/pon.1145] [Medline: 17253594]
29. Cummings DM, Morrissey S, Barondes MJ, Rogers L, Gustke S. Screening for diabetic retinopathy in rural areas: the potential of telemedicine. J Rural Health 2001;17(1):25-31. [doi: 10.1111/j.1748-0361.2001.tb00251.x] [Medline: 11354719]
30. DeHart D, King LB, Iachini AL, Browne T, Reitmeier M. Benefits and challenges of implementing telehealth in rural settings: a mixed-methods study of behavioral medicine providers. Health Soc Work 2022;47(1):7-18. [doi: 10.1093/hsw/hlab036] [Medline: 34910158]
31. Demaerschalk BM, Bobrow BJ, Raman R, Kiernan TJ, Aguilar MI, Ingall TJ, STRokE DOC AZ TIME Investigators. Stroke team remote evaluation using a digital observation camera in Arizona: the initial mayo clinic experience trial. Stroke 2010;41(6):1251-1258 [FREE Full text] [doi: 10.1161/STROKEAHA.109.574509] [Medline: 20431081]
32. DeVido J, Glezer A, Branagan L, Lau A, Bourgeois JA. Telepsychiatry for inpatient consultations at a separate campus of an academic medical center. Telemed J E Health 2016;22(7):572-576. [doi: 10.1089/tmj.2015.0125] [Medline: 26701608]
33. Donahue KE, Tillman J, Halladay JR, Cené CW, Hinderliter A, Cummings DM, et al. Lessons learned from implementing health coaching in the heart healthy lenoir hypertension study. Prog Community Health Partnersh 2016;10(4):559-567. [doi: 10.1353/cpr.2016.0064] [Medline: 28569681]
34. Finley BA, Palitsky R, Charteris E, Pacheco C, Kapoor D. Outpatient telecardiology perceptions among rural, suburban, and urban veterans utilizing in-person cardiology versus telecardiology services: a mixed methods analysis. J Rural Health 2021;37(4):812-820. [doi: 10.1111/jrh.12586] [Medline: 34002404]
35. Geller BM, Skelly JM, Dorwaldt AL, Howe KD, Dana GS, Flynn BS. Increasing patient/physician communications about colorectal cancer screening in rural primary care practices. Med Care 2008;46(9 Suppl 1):S36-S43 [FREE Full text] [doi: 10.1097/MLR.0b013e31817c60ea] [Medline: 18725831]
36. Gutierrez J, Moeckli J, Holcombe A, O'Shea AM, Bailey G, Rewerts K, et al. Implementing a telehospitalist program between veterans health administration hospitals: outcomes, acceptance, and barriers to implementation. J Hosp Med 2021;16(3):156-163 [FREE Full text] [doi: 10.12788/jhm.3570] [Medline: 33617436]
37. Hatch MN, Martinez RN, Etingen B, Cotner B, Hogan TP, Wickremasinghe IM, et al. Characterization of telehealth use in veterans with spinal cord injuries and disorders. PM R 2021;13(10):1094-1103. [doi: 10.1002/pmrj.12515] [Medline: 33098620]
38. Hicks LL, Fleming DA, Desaulnier A. The application of remote monitoring to improve health outcomes to a rural area. Telemed J E Health 2009;15(7):664-671. [doi: 10.1089/tmj.2009.0009] [Medline: 19694598]
39. Holloway B, Coon P, Kersten DW, Ciemins EL. Telehealth in rural Montana: promoting realistic independent self-management of diabetes. Diabetes Spectrum 2011;24(1):50-54. [doi: 10.2337/diaspect.24.1.50]
40. Khairat S, Haithcoat T, Liu S, Zaman T, Edson B, Gianforcaro R, et al. Advancing health equity and access using telemedicine: a geospatial assessment. J Am Med Inform Assoc 2019;26(8-9):796-805 [FREE Full text] [doi: 10.1093/jamia/ocz108] [Medline: 31340022]
41. Kulcsar Z, Albert D, Ercolano E, Mecchella JN. Telerheumatology: a technology appropriate for virtually all. Semin Arthritis Rheum 2016;46(3):380-385. [doi: 10.1016/j.semarthrit.2016.05.013] [Medline: 27395561]
42. Liu Y, Zupan NJ, Swearingen R, Jacobson N, Carlson JN, Mahoney JE, et al. Identification of barriers, facilitators and system-based implementation strategies to increase teleophthalmology use for diabetic eye screening in a rural US primary care clinic: a qualitative study. BMJ Open 2019;9(2):e022594 [FREE Full text] [doi: 10.1136/bmjopen-2018-022594] [Medline: 30782868]
43. Locke ER, Thomas RM, Woo DM, Nguyen EHK, Tamanaha BK, Press VG, et al. Using video telehealth to facilitate inhaler training in rural patients with obstructive lung disease. Telemed J E Health 2019;25(3):230-236 [FREE Full text] [doi: 10.1089/tmj.2017.0330] [Medline: 30016216]

44. McIlhenny CV, Guzic BL, Knee DR, Wendekier CM, Demuth BR, Roberts JB. Using technology to deliver healthcare education to rural patients. Rural Remote Health 2011;11(4):1798. [Medline: 21995854]
45. Owolo E, Petitt Z, Rowe D, Luo E, Bishop B, Poehlein E, et al. Sociodemographic trends in telemedicine visit completion in spine patients during the COVID-19 pandemic. Spine (Phila Pa 1976) 2023;48(21):1500-1507. [doi: 10.1097/BRS.0000000000004617] [Medline: 37235789]
46. Robinson KS, Morin PC, Shupe JAC, Izquierdo R, Ploutz-Snyder R, Meyer S, et al. Use of three computer training methods in elderly underserved rural patients enrolled in a diabetes telemedicine program. Comput Inform Nurs 2010;28(3):172-177. [doi: 10.1097/NCN.0b013e3181d785d5] [Medline: 20431360]
47. Rodriguez KL, Burkitt KH, Bayliss NK, Skoko JE, Switzer GE, Zickmund SL, et al. Veteran, primary care provider, and specialist satisfaction with electronic consultation. JMIR Med Inform 2015;3(1):e5 [FREE Full text] [doi: 10.2196/medinform.3725] [Medline: 25589233]
48. Schlittenhardt M, Smith SC, Ward-Smith P. Tele-continence care: a novel approach for providers. Urol Nurs 2016;36(5):217-223. [Medline: 29240333]
49. Schooley BL, Horan TA, Lee PW, West PA. Rural veteran access to healthcare services: investigating the role of information and communication technologies in overcoming spatial barriers. Perspect Health Inf Manag 2010;7(Spring):1f [FREE Full text] [Medline: 20697468]
50. Silvestrini M, Indresano J, Zeliadt SB, Chen JA. "There's a huge benefit just to know that someone cares:" a qualitative examination of rural veterans' experiences with TelePain. BMC Health Serv Res 2021;21(1):1111 [FREE Full text] [doi: 10.1186/s12913-021-07133-5] [Medline: 34656133]
51. Strowd RE, Strauss L, Graham R, Dodenhoff K, Schreiber A, Thomson S, et al. Rapid implementation of outpatient teleneurology in rural appalachia: barriers and disparities. Neurol Clin Pract 2021;11(3):232-241 [FREE Full text] [doi: 10.1212/CPJ.0000000000000906] [Medline: 34484890]
52. Switzer JA, Hall CE, Close B, Nichols FT, Gross H, Bruno A, et al. A telestroke network enhances recruitment into acute stroke clinical trials. Stroke 2010;41(3):566-569. [doi: 10.1161/STROKEAHA.109.566844] [Medline: 20056929]
53. Virmani T, Lotia M, Glover A, Pillai L, Kemp AS, Iyer A, et al. Feasibility of telemedicine research visits in people with Parkinson's disease residing in medically underserved areas. J Clin Transl Sci 2022;6(1):e133 [FREE Full text] [doi: 10.1017/cts.2022.459] [Medline: 36590358]
54. Waymouth M, Siconolfi D, Friedman EM, Saliba D, Ahluwalia SC, Shih RA. Barriers and facilitators to home- and community-based services access for persons with dementia and their caregivers. J Gerontol B Psychol Sci Soc Sci 2023;78(6):1085-1097 [FREE Full text] [doi: 10.1093/geronb/gbad039] [Medline: 36896936]
55. Weiner MF, Rossetti HC, Harrah K. Videoconference diagnosis and management of choctaw indian dementia patients. Alzheimers Dement 2011;7(6):562-566 [FREE Full text] [doi: 10.1016/j.jalz.2011.02.006] [Medline: 22055972]
56. West VL, Milio N. Organizational and environmental factors affecting the utilization of telemedicine in rural home healthcare. Home Health Care Serv Q 2004;23(4):49-67. [doi: 10.1300/J027v23n04_04] [Medline: 15778152]
57. West SP, Lagua C, Trief PM, Izquierdo R, Weinstock RS. Goal setting using telemedicine in rural underserved older adults with diabetes: experiences from the informatics for diabetes education and telemedicine project. Telemed J E Health 2010;16(4):405-416. [doi: 10.1089/tmj.2009.0136] [Medline: 20507198]
58. Zulman DM, Wong EP, Slightam C, Gregory A, Jacobs JC, Kimerling R, et al. Making connections: nationwide implementation of video telehealth tablets to address access barriers in veterans. JAMIA Open 2019;2(3):323-329 [FREE Full text] [doi: 10.1093/jamiaopen/ooz024] [Medline: 32766533]
59. Munn Z, Aromataris E, Tufanaru C, Stern C, Porritt K, Farrow J, et al. The development of software to support multiple systematic review types: the Joanna Briggs institute system for the unified management, assessment and review of information (JBI SUMARI). Int J Evid Based Healthc 2019;17(1):36-43. [doi: 10.1097/XEB.0000000000000152] [Medline: 30239357]
60. Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The mixed methods appraisal tool (MMAT) version 2018 for information professionals and researchers. Educ Inf 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
61. Depatie A, Bigbee JL. Rural older adult readiness to adopt mobile health technology: a descriptive study. Online J Rural Nurs Health Care 2013;15(1):150-184. [doi: 10.14574/ojrnhc.v15i1.346]
62. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. J Med Internet Res 2019;21(4):e11604 [FREE Full text] [doi: 10.2196/11604] [Medline: 30958272]
63. Lim CT, Rosenfeld LC, Nissen NJ, Wang PS, Patel NC, Powers BW, et al. Remote care management for older adult populations with elevated prevalence of depression or anxiety and comorbid chronic medical illness: a systematic review. J Acad Consult Liaison Psychiatry 2022;63(3):198-212 [FREE Full text] [doi: 10.1016/j.jaclp.2022.02.005] [Medline: 35189427]
64. Louras N, Reading Turchioe M, Shafran Topaz L, Ellison M, Abudu-Solo J, Blutinger E, et al. Mobile integrated health interventions for older adults: a systematic review. Innov Aging 2023;7(3):igad017 [FREE Full text] [doi: 10.1093/geroni/igad017] [Medline: 37090165]

65. Kainiemi E, Saukkonen P, Virtanen L, Vehko T, Kyytsönen M, Aaltonen M, et al. Perceived benefits of digital health and social services among older adults: a population-based cross-sectional survey. Digit Health 2023;9:20552076231173559 [FREE Full text] [doi: 10.1177/20552076231173559] [Medline: 37312955]
66. Busby J, Tanberk J. FCC underestimates Americans unserved by broadband internet by 50%. BroadBand Now Team. 2020. URL: https://broadbandnow.com/research/fcc-underestimates-unserved-by-50-percent [accessed 2025-12-16]
67. Cheung K, Chau AKC, Woo J, Lai ET. The age-based digital divide in an increasingly digital world: a focus group investigation during the COVID-19 pandemic. Arch Gerontol Geriatr 2023;115:105225. [doi: 10.1016/j.archger.2023.105225] [Medline: 37837792]
68. Cousins MM, Van Til M, Steppe E, Ng S, Ellimoottil C, Sun Y, et al. Age, race, insurance type, and digital divide index are associated with video visit completion for patients seen for oncologic care in a large hospital system during the COVID-19 pandemic. PLoS One 2022;17(11):e0277617 [FREE Full text] [doi: 10.1371/journal.pone.0277617] [Medline: 36395112]
69. Davis FD, Venkatesh V. A critical assessment of potential measurement biases in the technology acceptance model: three experiments. Int J Hum Comput Stud 1996;45(1):19-45. [doi: 10.1006/ijhc.1996.0040]
70. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manag Sci 2000;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
71. Venkatesh V, Bala H. Technology acceptance model 3 and a research agenda on interventions. Decis Sci 2008;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
72. Saphner T, Marek A, Homa J, Robinson L, Glandt N. Clinical trial participation assessed by age, sex, race, ethnicity, and socioeconomic status. Contemp Clin Trials 2021;103:106315 [FREE Full text] [doi: 10.1016/j.cct.2021.106315] [Medline: 33626412]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
TAM: technology acceptance model

Edited by M Mardini; submitted 30.Dec.2024; peer-reviewed by NA Puccinelli-Ortega, LR Guo; comments to author 22.Jul.2025; revised version received 09.Sep.2025; accepted 10.Nov.2025; published 04.Feb.2026.

Please cite as:
Siegel ZM, Quinkert E, Pai J, Miller CH, Lewis MW
Lessons Learned About Digital Health Tool Acceptability Among Rural Older Adults: Systematic Review Guided by the Technology Acceptance Model
JMIR Aging 2026;9:e70012
URL: https://aging.jmir.org/2026/1/e70012
doi:10.2196/70012
PMID:

Using Indoor Movement Complexity in Smart Homes to Detect Frailty in Older Adults: Multiple-Methods Case Series Study

Katherine Wuestney[1], PhD; Diane Cook[2], MSci, PhD; Catherine Van Son[1], MSN, PhD; Roschelle Fritz[3], MSN, PhD

[1]College of Nursing, Washington State University, Pullman, WA, United States

[2]School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA, United States

[3]Betty Irene Moore School of Nursing, UC Davis Health System, 2570 48th St, Sacramento, CA, United States

Corresponding Author:
Roschelle Fritz, MSN, PhD
Betty Irene Moore School of Nursing, UC Davis Health System, 2570 48th St, Sacramento, CA, United States

Abstract

Background: The theory of complexity in aging indicates that the complexity of sensor-derived physiological and behavioral signals reflects an older adult's adaptive capacity and, in turn, their frailty. Smart homes with ambient sensors offer a unique opportunity to longitudinally explore the complexity of older adults' indoor movement in a real-world setting. Here, we introduce a computational method to estimate behavior complexity from sensor data. We further conduct a multiple-methods case series to explore the relationship between entropy-measured smart home data complexity and older adult frailty.

Objective: This study aims to explore the relationship between entropy-measured ambient sensor data complexity and frailty in independent community-dwelling older adults.

Methods: The nature of older adults' indoor movement complexity is measured by quantifying the entropy of smart home data. Overall, 11 cases with persons aged 65 years and older were drawn from an ongoing smart home study to illustrate the method. We assessed weekly frailty for these cases using the Clinical Frailty Scale. For corresponding time ranges, we measured the complexity of smart home data using a fixed-width sliding window and an entropy-based complexity index (Rényi Complexity Index) built on a Universal Sequence Map (USM-Rényi). Descriptive statistics and graphical analysis were used to describe intraindividual frailty and sensor complexity change.

Results: The complexity of sensor-observed indoor movement does change over time in older adults as quantified by the computational method. In some individuals, these changes track with health transitions and frailty progression. The trends and monotonicity of complexity trajectories varied between cases. Overall, 3 of the cases demonstrated a negative association between frailty and complexity, while the association was not as clear for the other cases.

Conclusions: The complexity of older adults' smart home data is highly diverse. Changes in health and frailty influence indoor movement complexity. Although the findings suggest a relationship between frailty and complexity, confounding factors, such as home layout, visitors, external events, and technology disruptions, may influence sensor signals.

(JMIR Aging 2026;9:e77322) doi:10.2196/77322

KEYWORDS

entropy; complexity; smart home; movement; frailty

Introduction

Background

Frailty is a critical public health challenge among older adults globally. It is characterized as a clinically identifiable state of diminished physiologic reserve and heightened vulnerability to stressors and affects. An estimated 10%‑15% of community-dwelling individuals aged 65 years and older experience frailty, with the prevalence escalating to 51% among those aged 90 years and older [1]. This multifactorial syndrome, encompassing multiple impairments such as physical weakness, exhaustion, slow gait, low activity levels, and unintentional weight loss, elevates risks for adverse outcomes like falls, hospitalizations, functional decline, and mortality [1,2]. When unaddressed, frailty imposes substantial economic burdens, with frail older adult women incurring as much as 184% the health care cost of nonfrail older adult women [3,4]. Poor outcomes related to frailty strain health care systems and diminish quality of life [3-5].

Most older adults prefer to remain in their own homes and communities as they age. Despite this desire, age-related frailty and its sequelae remain a threat to their independence and quality of life. Older adults who are frail are more likely to present with atypical, nonspecific symptoms of acute illness, which include immobility, instability, incontinence, weakness, and delirium [6]. This can put them at risk for poorer outcomes if such atypical signs are treated as the primary problem rather than merely the manifestation of underlying, seemingly unrelated illnesses. Thus, while frailty is a significant issue, it also functions as a gateway to a wide array of other salient health issues for older adults [7].

Two urgent challenges face older adults who wish to age in place: (1) needing validated methods to detect incipient frailty at home and (2) determining the best way to analyze these data for predicting frailty that focuses on the efforts of early intervention strategies [8]. This study addresses both challenges by exploring complexity as a feature of frailty in smart home sensor data. In the context of aging, the term "complexity" is often used to describe difficult problems that must be mitigated, making care more daunting [9]. While complex behavior is often seen as a challenge in caregiving, from a systems theory perspective, too little complexity may indicate diminished physiological adaptability. The theory of complexity in aging asserts that complexity is a direct indicator of the health of physiologic systems and aging reduces this complexity, resulting in frailty [10].

Sensors are ubiquitous in our world, and this reality is accompanied by an increasing interest in discovering indicators of human health using these sensors. These indicators serve a similar function as conventional biological and imaging biomarkers with less reliance on expensive lab equipment, visits to remote sites for time-consuming tests, or physically invasive procedures [11]. Digital biomarkers are valuable components of geriatric telehealth and precision medicine since these technologies support continuous, longitudinal, remotely delivered measurement of intraindividual changes in older adults' health [12]. Among this class of markers are behavior markers created from continuously collected sensor data, which open substantial opportunities to explore the complex dynamics of aging in an ecologically valid, real-world setting.

In this study, we enlist digital biomarkers to explore the relationships between behavior and frailty. This is increasingly important because the number of persons aged 80 years and older is expected to triple between 2020 and 2050 [13]. With the rise in age-related frailty and incidence of chronic conditions, meeting the health needs of older adults is increasingly burdensome. A review of unobtrusive frailty digital biomarkers concluded that passive infrared motion sensors, especially as part of a smart home, are the most promising type of embedded ambient sensor for detecting frailty [14]. Smart homes were promoted for their potential to uniquely inform individual responses to disease or treatment.

Older adults prefer digital biomarker technologies that minimally impose on their lifestyles [15]. From this perspective, digital behavior markers derived from completely passive monitoring (eg, ambient sensors embedded in residential environments) offer advantages over those measured via semipassive or active monitoring (eg, wearable sensors that must be routinely charged and positioned) [14]. Smart homes represent a passive biomarker technology that consists of ambient sensors to monitor movement and door interactions, combined with a computing infrastructure to collect, organize, and store the data [16]. The resulting time series data can be analyzed to understand the smart home resident's health status.

Prior Work

The theory of complexity in aging hypothesizes that measuring the complexity of a person's sensor-derived signals can indicate the underlying state of an older adult's adaptive capacity [17,18]. We analyze a person's behavioral signal complexity as an indicator of their adaptive capacity or functional reserve, referring to the capacity of their physiological and behavioral systems to maintain or regain function when perturbed. In a complex-systems view of aging, this reserve depends on the multiscale dynamics that are present between the system components [19]. These dynamics support homeostasis, the process that maintains internal stability while adapting to change.

In earlier work, researchers have investigated the use of multiscale entropy (MSE) to quantify complexity across time scales for physiological data. Bizovska et al [20] used MSE and Shannon entropy to analyze gait complexity as a mechanism for predicting fall risk in older adults. Castiglia et al [21] investigated the selection of MSE parameters that yield the best predictive probability in differentiating subjects with Parkinson disease from healthy subjects based on trunk acceleration patterns. Gao et al [22] use distribution entropy, which calculates the complexity of signal pattern distribution within a phase space representation, to determine whether pulse rate complexity is associated with corresponding cognitive decline in older adults.

Frailty is hypothesized to be an emergent state that arises from a critically dysregulated complex system [9]. In other words, the system dynamics may erode with aging and disease, causing complexity to decline and frailty vulnerability to increase. Evidence from cross-sectional studies suggests this process can be observed as a change in complexity in a diverse range of physiological and behavioral signals. For example, lower blood pressure interbeat interval complexity, when the beat-to-beat pattern becomes simpler and more uniform, is associated with greater frailty and dementia risk [23]. Similarly, lower moment-to-moment center-of-pressure complexity, such as simpler, more regular sway, during balance tasks is associated with increased future incidence of falls [24]. Reduced complexity of spontaneous brain activity, measured via the blood oxygenation signal, is associated with slower gait speed [25], and lower physical activity complexity and variance are associated with greater self-reported frailty [26] and mortality risk [27]. These examples suggest that reduced signal complexity may be a generalizable marker of physiologic decline.

Prior studies have measured the complexity of smart home data [28,29]. These earlier studies included complexity as one of a set of variables input to machine-learning models that were trained for specific tasks such as detecting visitors or predicting in-home movement [30,31]. Little is known regarding how within-person complexity, in isolation from other variables, evolves in relation to health outcomes over the long term. In a study by Schutz et al [28], Shannon entropy of refrigerator use was one of the strongest predictors of frailty (r=−0.25). However, this analysis did not explore the evolution of complexity over time. A study by Takahashi et al [32] examined activities over a 2-year period and found that increased activity diversity manifested an inverse relationship with frailty. The findings support our hypothesis, but they are based on survey data rather than analysis of passively observed activity patterns.

Two prior studies applied complexity measures to data collected by the Center for Advanced Studies in Adaptive Systems (CASAS), the same data collection infrastructure used in this study. Specifically, Wang et al [30] estimated complexity using compression-based estimators to establish a theoretical limit on the predictability of indoor human mobility. In an earlier study by Gopalratnam and Cook [33], CASAS smart home data were analyzed with a Lempel-Ziv compression-based incremental parser to predict the resident's next interaction with the home. Although the smart home sites and analysis goals differed from this study, the prior work established the use of such behavior analyses from smart home data.

The common approach to predicting frailty leverages sources such as electronic health records and manually collected clinical data [34]. However, wearable sensors are increasingly accessible and offer a mechanism for passively sensing and detecting frailty [8]. Many frailty studies that analyze wearable data focus on predicting physical frailty components such as slowness and inactivity. These studies extract gait parameters such as cadence and indicators of time spent walking and standing [35,36]. One study instead analyzed Fitbit data that were collected while individuals performed an upper extremity function test [37]. While the primary component of these analyses is accelerometry, Merchant et al [38] combine these parameters with heart rate to analyze scripted movements such as sit-to-stand, walk, and climb stairs.

Wearable sensors have demonstrated the ability to sense and quantify changes in movement parameters that are associated with frailty. We focus here on monitoring activity and detecting frailty using ambient sensors in smart homes. Ambient sensors impose no user burden. Sensors collect data for multiple years on a charge, which results in continuous, uninterrupted monitoring of in-home behavior as a person's health status changes. Using wearable sensors, consistent multiday wear is challenging, and adherence varies with demographics and cognition [39]. While wearable sensors provide direct access to heart rate and gait parameters, the smart home sensors contribute context-rich information about location traces, sleep and wake routines, and activity patterns that are not easily modeled from wearable data [40]. Because we want to monitor uninterrupted longitudinal behavior patterns, we focus this analysis on data collected in smart home settings.

To address this knowledge gap, we present an exploratory case series investigating how the complexity of older adults' indoor movement patterns, as captured by the CASAS smart home, changes over time in relation to changes in their health status. Considering that this relationship between complexity and frailty has been observed across a diverse range of seemingly unrelated physiologic and behavioral signals, we hypothesize that changes in the complexity of time series obtained from smart home sensors are similarly associated with changes in health status and frailty of the older adult smart home occupant.

The case series design prioritizes investigation of intraindividual interpretation and allows us to integrate each participant's clinical narrative into the analysis. To promote replication of methods and application to new data, we make the analysis and visualization tools publicly available for the community to use in the calculation of sensor-derived behavior complexity.

Methods

Overview

We performed a multiple-methods exploratory case series, combining participant narrative and qualitative nursing data with complexity analysis of smart home sensor time series data to contextualize intraindividual changes in complexity of indoor movement. A case series analysis was chosen because the method is useful for exploring intricate, real-world issues in novel ways, especially when triangulating data from different sources to discover differences and similarities across similar cases. The method fosters a more nuanced, valid, and actionable understanding of the cases under study.

Participants

We used secondary data from sensors installed in the homes of community-dwelling older adults between October 2016 and December 2022 as part of the ongoing clinician-in-the-loop smart home research study [41]. To be included in the clinician-in-the-loop study, participants had to be aged 60 years and older, have at least 1 chronic condition, and had to be proficient in English. For this case study series, we applied the additional criteria of living alone without pets for the entire duration of the data collection and collected a minimum of 9 months of smart home data. Cases were further excluded if a majority of the days and sensors were missing. The resulting sample consisted of 11 cases, representing a balance between stable participants and those who exhibited frailty transitions. For this case series, each participant is considered as 1 case. Among these participants, cases 8 through 10 exhibited constant frailty scores, while the others experienced frailty that fluctuated throughout the data collection.

Participant cases included in the present analysis lived in independent living apartments in continuing care retirement communities. Most cases' ages were in the range of 80 to 89 years, although 2 were aged 70 - 79 years and 1 was aged 90 - 99 years. All included cases identified as non-Hispanic White, and 7 of the cases identified as women. Information summarizing participants, their chronic health conditions, and their home characteristics is provided in Table 1.

Table . Demographics and data characteristics for each participant case.

Case	Age[a] (years)	Sex	Home type	Sensors	Days	Window size[b]	Chronic conditions
1	80 - 89	Female	1-bedroom apartment	15	310	25,159	CV[c], NM[d], Pain[e]
2	70 - 79	Female	1-bedroom apartment	13	416	63,419	CV, Pulm[f], Pain
3	80 - 89	Female	3-bedroom duplex	20	366	19,714	CV, Pulm, NM
4	80 - 89	Female	1-bedroom apartment	13	349	33,358	CV, Pulm, NM
5	80 - 89	Male	1-bedroom apartment	15	629	60,800	NM, Pain
6	80 - 89	Female	1-bedroom apartment	12	571	28,951	CV, Pulm, NM, Pain
7	80 - 89	Female	1-bedroom apartment	14	385	28,182	Pain, CI[g]
8	70 - 79	Male	2-bedroom duplex	22	330	49,303	CV, NM, Pain
9	80 - 89	Female	Studio apartment	11	354	35,310	CV, Pain, CI
10	90 - 99	Male	1-bedroom apartment	12	264	18,951	CV, NM, Pain
11	80 - 89	Male	1-bedroom apartment	13	286	34,903	CV, Pain

[a]To preserve privacy, age is given as a range.

[b]Sliding window size was determined by the maximum biweekly count of sensor messages (excluding OFF and CLOSE) observed in the participant's data.

[c]CV: cardiovascular.

[d] NM: neuromuscular.

[e] Pain: chronic pain.

[f]Pulm: pulmonary.

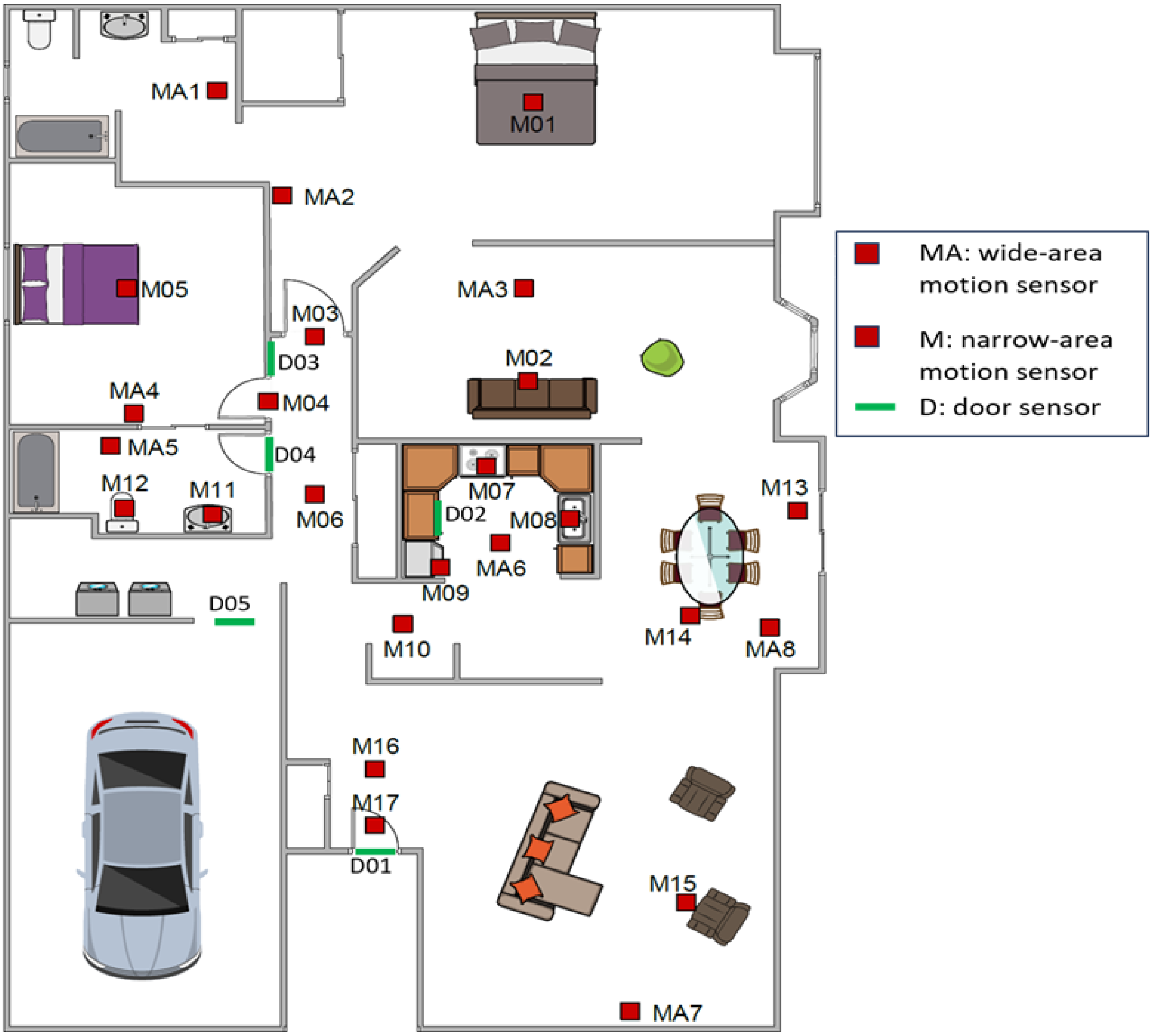
[g]CI: cognitive impairment.

CASAS Smart Home

The CASAS smart home contains passive infrared motion detectors, light, magnetic door use, and temperature sensors. These sensors were installed strategically throughout each house to capture activity in critical locations (Figure 1). At least 1 motion detector with a 360^o view was installed in each room. Additional motion detectors with a narrower field of view (approximately 1 m in diameter) were positioned in areas of high use, such as the bed, sinks, toilet, and frequented furniture (eg, preferred living room chair). Because floor plans, furniture layouts, and daily routines differed across each home, the number of sensors installed for included cases ranged from 11 to 22.

CASAS sensors send messages containing their readings to a middleware layer resident on a Raspberry Pi [16]. Architecture components communicate using a Zigbee wireless mesh. The middleware publish and subscribe manager allows hardware components to publish and receive messages. And annotates sensor readings with the corresponding sensor identifier and timestamp. All collected data are encrypted and securely transmitted to a password-protected server for storage and analysis.

Figure 1. Location of sensors in a Center for Advanced Studies in Adaptive Systems smart home.



We examined data collected from the passive infrared and door-use sensors. Each sensor samples the environment at 1.25 Hz. Rather than report the state at a constant frequency, the sensors record data when a change in state is sensed (eg, a door is opened, motion is detected). Once triggered, the sensor sends a message reflecting the new state to a central relay, which labels each message with the sensor identifier and timestamp, then transmits the data to a secured database. The resulting dataset is a timestamped series of binary messages (“ON” or “OFF” for motion sensors, “OPEN” or “CLOSE” for door sensors) indicating the time and location of the sensor reading in the home. Because an ON message from a motion sensor is followed by an OFF message (marking the end of movement within the sensor’s field of view or lack of activity for 1.25 s), both ON and OFF messages artificially inflated the regularity of the data sequence. Following previous literature measuring the entropy of smart home data [30,42], we excluded all OFF messages from motion sensors and all CLOSE messages from door sensors. Example deidentified CASAS datasets are available online [43].

Clinical Data

For each participant, nurse researchers conducted an initial comprehensive geriatric assessment, including functional status in activities of daily living (ADL) and instrumental activities of daily living (IADL), current health diagnoses, health history, medications, fall history, psychosocial supports and family presence, assistive device use, review of body systems, and personal demographic history. Participants then received weekly follow-up telehealth calls from a nurse researcher to assess for any changes in health or function from baseline. Weekly nursing data included, but was not limited to, vital signs, pain, sleep quality, psychosocial well-being (including the presence of visitors), changes in ADL and IADL status, and a brief review of physiologic systems and daily routines [41].

Although frailty was not measured as part of the primary data collection, the clinical data collected during the study provided information to retrospectively estimate weekly frailty using the Clinical Frailty Scale (CFS) [44]. The CFS is a 9-point scale (1=very fit to 9=terminal illness) designed to guide a clinician in assessing a holistic picture of a person’s frailty status using elements of a comprehensive geriatric assessment, including overall activity level, functional dependence, and management and control of chronic condition symptoms [45]. Two CFS-trained nurse scientists reviewed the clinical data for each participant and assigned a frailty score for each week of data collection (Table 2). Changes in ADL and IADL independence, use of a new assistive device, and descriptions of increasing fatigue or “slowing down” were the most common health changes associated with an upward shift in the participant’s CFS score.

Table . Example Clinical Frailty Scale codebook with scores for case 4.

Week[a]	Date (2017)	CFS[b] score	Rationale
45	July 24	5	No change
46	July 31	5	No change
47	August 7	5	Decreased activity, increased weeping lower legs
48	August 14	6	“I have to be careful not to fall”
49	August 21	6	Considering assisted living but hiring in-home help
50	August 28	6	No change
51	September 4	6	Losing weight, legs improving
52	September 11	6	Legs continue weeping due to heart failure
53	September 18	6	Began using pursed-lip breathing, moving less
54	September 25	6	Doctor’s visit, medication change
55	October 2	6	Legs improving, taking diuretic

[a]Weeks 1 - 44 (CFS score: mean 4.9, SD 0.33; range 4 - 5); weeks 56 - 60 (CFS score: mean 6.4, SD 0.89; range 6 - 8).
[b]CFS: Clinical Frailty Scale.

Data Preprocessing

Because smart home data were collected in real-world settings over extended periods, we needed to address missing and noisy data. We screened each participant’s sensor data for evidence of sensor malfunctioning, extended absences, and other issues. Periods associated with participant absence for more than one night (eg, vacation or hospitalization) were excluded from the analysis. Additionally, any periods where all sensors did not report readings, regardless of explanation, were excluded. Periods with no messages from a given sensor were cross-referenced with battery data from that sensor to confirm whether the absence was due to a change in behavior or a sensor malfunction. Sensors missing >50% of the observation time over one or more consecutive days were excluded. Sensors missing more than 50% of the observation period were either excluded or they were included, and the time associated with that sensor’s absence was excluded.

The varying size of the homes and the corresponding number and density of sensors impact the scale of Rényi Complexity Index (RényiCI) values we observe in each home. A cross-sectional study would require that sensors be grouped into larger, consistently sized sets or that the values be normalized. For this study, we are interested in within-home RényiCI changes, so no adjustments are made to the per-home RényiCI scales. Because the sensors report binary state (motion ON or OFF, door OPEN or CLOSED), the raw sensor values are not normalized.

Some of the participants included in this study were enrolled during the onset of the COVID-19 pandemic, which had a dramatic global impact on daily activities. For those participants, if the majority of a participant’s data were collected after the pandemic onset, data from before March 16, 2020, were excluded. Similarly, for participants with most data collected before the pandemic, we excluded data from March 16, 2020, onward.

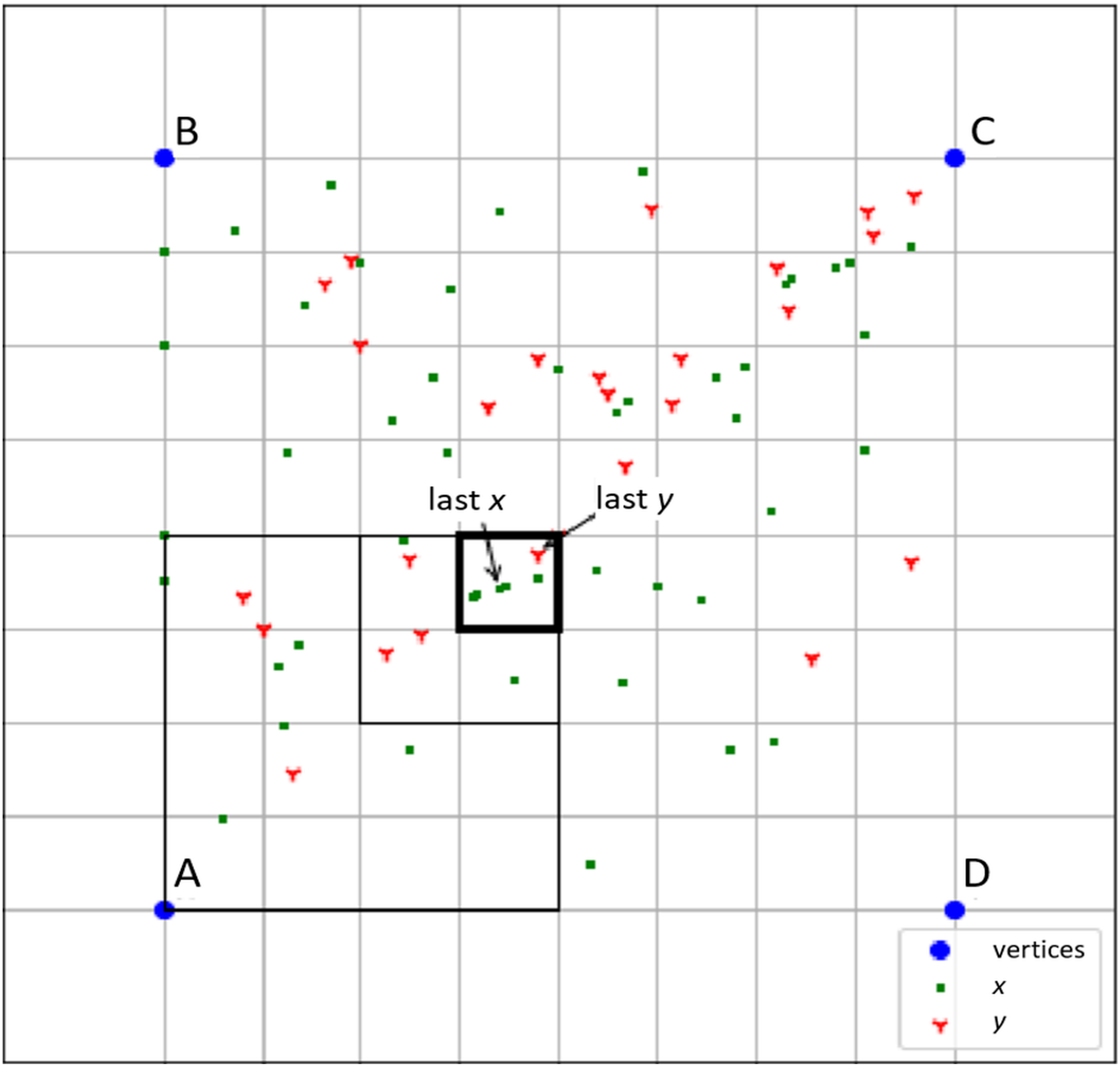
Data cleaning included the removal of sensor data from analysis for sensors sending “error” signals, which can occur when low battery health or technical issues occur during installation. Only 2 homes were affected by this: case 1 (dining room area sensor, hallway, bathroom sink, and door for the primary bedroom) and case 5 (entry door, refrigerator, and bathroom area sensor). After data cleaning procedures were applied, the series of timestamped sensor messages was coded based on the sensor identifiers, resulting in a time series of discrete (categorical) sensor states. These discrete-valued series were then used to compute the complexity of sensor state transitions over time.

Complexity Measurement

Understanding patterns in human behavior, especially those that signal changes in health or frailty, requires tools that quantify how predictable or irregular those behaviors are over time. One such method is based on entropy, a way of measuring complexity or unpredictability in a sequence of events. Almeida and Vinga [46] introduced a technique to calculate this complexity using a Universal Sequence Map (USM). This approach turns a sequence of events (eg, daily activities recorded in a smart home) into a set of coordinates in a multidimensional space. These coordinates reflect how often and in what sequence specific symbols (eg, messages from home sensors) occur relative to one another. Once the sequence has been mapped to this space, the method estimates how densely packed these points are in space using the kernel density estimation statistical technique. The resulting density provides insight into whether the behavior is highly repetitive (low complexity) or highly varied (high complexity). Highly repetitive behavior may, for example, reflect a person moving primarily between the living room and bathroom each day. A more complex behavior will vary the daily sequence and perhaps more frequently introduce additional areas, including the guest room, the garage, and the front door to leave the home.

Figure 2 illustrates the process of creating a USM. Unique sensor readings are converted to symbols (A-D). USM coordinates are calculated by assigning each unit symbol in the sequence to a position in a multidimensional space. The positions are defined so that each symbol is equally distant from the others, ensuring that no symbol is biased in how the space is structured. The number of dimensions of the space is chosen so that each distinct symbol can be uniquely represented using binary digits. The sequence is processed forward (considering prior context) and backward (considering subsequent context), and the 2 resulting maps are merged to capture bidirectional structure in the behavior.

Figure 2. Plot of 2 sequences in a Universal Sequence Map. The last 4 symbols of sequence x are *ACCA*, and the last 4 of sequence y are *CCCA*. The highlighted subquadrant contains the coordinate of the sequences' last symbol, *A*.



The resulting space creates a unified framework to measure the complexity of sequences from the resulting coordinates. To quantify the complexity of these mapped sequences, Vinga and Almeida [47] introduced a method that computes Rényi entropy, a generalization of Shannon entropy, from the density of the USM coordinate distribution. This approach is particularly effective for relatively short sequences, such as those representing daily behavior in smart homes.

Since the idea of entropy was introduced in information theory, many variations have been introduced to measure complexity in different contexts. These measures vary by the type and quantity of data they process, their sensitivity to noise, and their assumptions about the underlying state space distribution. An ideal measure of sensor-based time series complexity is one that tends toward a minimum value for both deterministic and random sequences while handling varying alphabet sizes and being sensitive to changes in complexity over short sample lengths. Rényi entropy of USMs was selected for our analysis as a method that meets these constraints.

A key strength of this method is its flexibility: it can emphasize either common or rare patterns, depending on how the parameters are configured. Importantly, the frequency of any subsequence of any length can be estimated by analyzing how dense different regions of the USM space are. The kernel size (ie, the size of the region considered) controls the length of the subsequences being emphasized. We use this principle to estimate Rényi entropy at multiple scales, where each scale corresponds to a different behavioral timespan or sequence length. This flexibility enables a multiscale view of behavioral complexity, which we refer to as the RényiCI. An in-depth tutorial and code are provided online [48].

Statistical Analysis

Because this case series investigates how the complexity of motion sensor transitions, representing indoor movement trajectories, evolves over time, we computed RényiCI for each participant using a sliding window approach with a fixed window size, *n*. The actual RényiCI values will shift with the number of sensors in the space and the window size; thus, the values should be examined for change within a single home across multiple time points. Higher RényiCI values indicate more complex behavior, while lower values suggest simpler, more predictable patterns. The sliding window method evaluates the time series in overlapping segments: starting with the first *n* data points, it computes summary statistics, shifts the window forward by a set number of steps, and repeats the process.

To ensure each window captured both routine cyclic behaviors (eg, weekly housekeeping) and short-term variations, we defined each participant’s window size as the maximum number of sensor messages observed within any 2-week period (Table 3). The window was advanced using a step size equal to one-quarter of the window size.

Table . Sliding window statistics for each case. Runs test results were omitted as all resulted in *P* values <.001.

Sliding window			RényiCI statistics					ρ[a]	*P* value
Case	Count	Days, median (IQR; max)	Mean (SD)	CoV[b, c]	Median (IQR)	KS[d]	*P* value		
1	70	16.6 (15.5 to 17.2; 18.9)	−45.95 (0.26)	.006	−45.89 (−46.07 to −45.79)	0.10	.41	0.11	.48
2	66	24.7 (21.7 to 26.1; 30.8)	−38.85 (0.57)	.015	−38.65 (−39.09 to −38.47)	0.21	<.001	0.29	.06
3	80	16.3 (15.4 to 16.9; 18)	−61.97 (0.23)	.004	−61.94 (−62.18 to −61.78)	0.11	.29	0.12	.36
4	73	18.1 (16.8 to 19.2; 21)	−38.91 (0.27)	.007	−38.93 (−39.13 to −38.74)	0.09	.60	0.01	.96
5	121	18.1 (17.4 to 18.8; 30.7)	−44.66 (0.19)	.004	−44.68 (−44.81 to −44.51)	0.09	.23	−0.68	<.001
6	121	16 (15.5 to 16.6; 18.4)	−36.32 (0.10)	.003	−36.3 (−36.4 to −36.25)	0.09	.34	0.26	<.001
7	82	16.8 (15.6 to 17.7; 20.5)	−42.74 (0.14)	.003	−42.76 (−42.84 to −42.66)	0.07	.77	0.42	<.001
8	52	17.1 (16.1 to 18.1; 20.1)	−70.13 (0.20)	.003	−70.18 (−70.26 to −70.01)	0.12	.38	—[e]	—
9	73	18.3 (17.7 to 19; 20.6)	−33.07 (0.13)	.004	−33.07 (−33.15 to −32.99)	0.05	.98	—	—
10	49	18.9 (18 to 19.5; 22.6)	−36.64 (0.25)	.007	−36.64 (−36.84 to −36.43)	0.12	.47	—	—
11	65	15.5 (15 to 16.2; 18)	−39.58 (0.24)	.006	−39.58 (−39.78 to −39.35)	0.10	.53	0.24	.09

[a]Spearman rank correlation.

[b]CoV: coefficient of variance.

[c]Coefficient of variance was computed as the SD/mean.

[d]KS: Kolmogorov-Smirnov distance.

[e]Correlation is not provided because CFS is a constant.

To examine how the complexity of patterns relates to frailty status, we visualized RényiCI values using time series plots and categorical scatter (jitter) plots. Because RényiCI values can vary in scale depending on the number of sensors and the window size, we applied normalization within each case to enable comparison. To assess temporal fluctuations in complexity, we also computed the first-order difference of the normalized RényiCI sequence: ΔRenyiCI`t=RenyiCI`t-RenyiCI`t-1. Here, ΔRenyiCI`t represents the change in normalized complexity between consecutive windows.

To evaluate whether these complexity estimates varied systematically over time (in comparison to random changes in complexity), we applied Kolmogorov-Smirnov (KS) tests and runs tests to each participant's sequence of RényiCI values under the null hypothesis of randomness. The KS test checks whether the complexity values follow a normal distribution, as would be expected with random data. The runs test looks at the order of values in the sequence, rather than just the distribution, to determine if they appear in nonrandom patterns. Computation of USM-based RényiCI values was conducted in Python (version 3.9; Python Software Foundation) using our pyusm library [48]. This open-source package is publicly available and includes tools for computing USM, USM-Rényi, and generating 2D USM visualizations.

Finally, to resolve ambiguous quantitative results, sequential explanatory techniques were used. Quantitative results were reviewed alongside frailty scores assigned to each week of nursing narrative documentation, which included written text about participants' physical and functional health recorded during weekly phone calls and monthly home visits. RényiCI complexity values were compared to recorded CFS scores. Lower complexity values combined with higher CFS scores meant the participant was frailer.

Ethical Considerations

The Washington State University Institutional Review Board approved the presented secondary analysis (protocol 18764) and parent study (protocol 15412). All participants provided informed consent, and their data were deidentified and securely managed for analysis. Participation was voluntary and without compensation.

Results

Distributional Characteristics and Statistical Testing

Figure 3 presents time series plots of normalized RényiCI, frailty scores, and first-order differences in normalized RényiCI for each case. In the plots, time is measured in observation days. Summary statistics of overall RényiCI, KS, and Spearman rank correlation values are reported in Table 3. The shape of the RényiCI distributions varied notably across cases. Case 9 exhibited the only unimodal, symmetric distribution (Figure 4), while the remaining cases showed skewness or kurtosis. Cases 1 and 2 were strongly left-skewed, while cases 3, 5, 10, and 11 displayed low kurtosis. Cases 3 and 11 also showed bimodal distributions. Despite this heterogeneity, only case 2 showed a statistically significant deviation from a random normal distribution (*P*=.006) based on the KS test of normality.

Figure 3. Plots of complexity, frailty, and complexity change as a function of time. Lower values reflect less complexity. Δ_4RényiCI' represents the value difference between sliding windows at times *t* and *t-4*. CFS: Clinical Frailty Scale; RényiCI: Rényi Complexity Index.

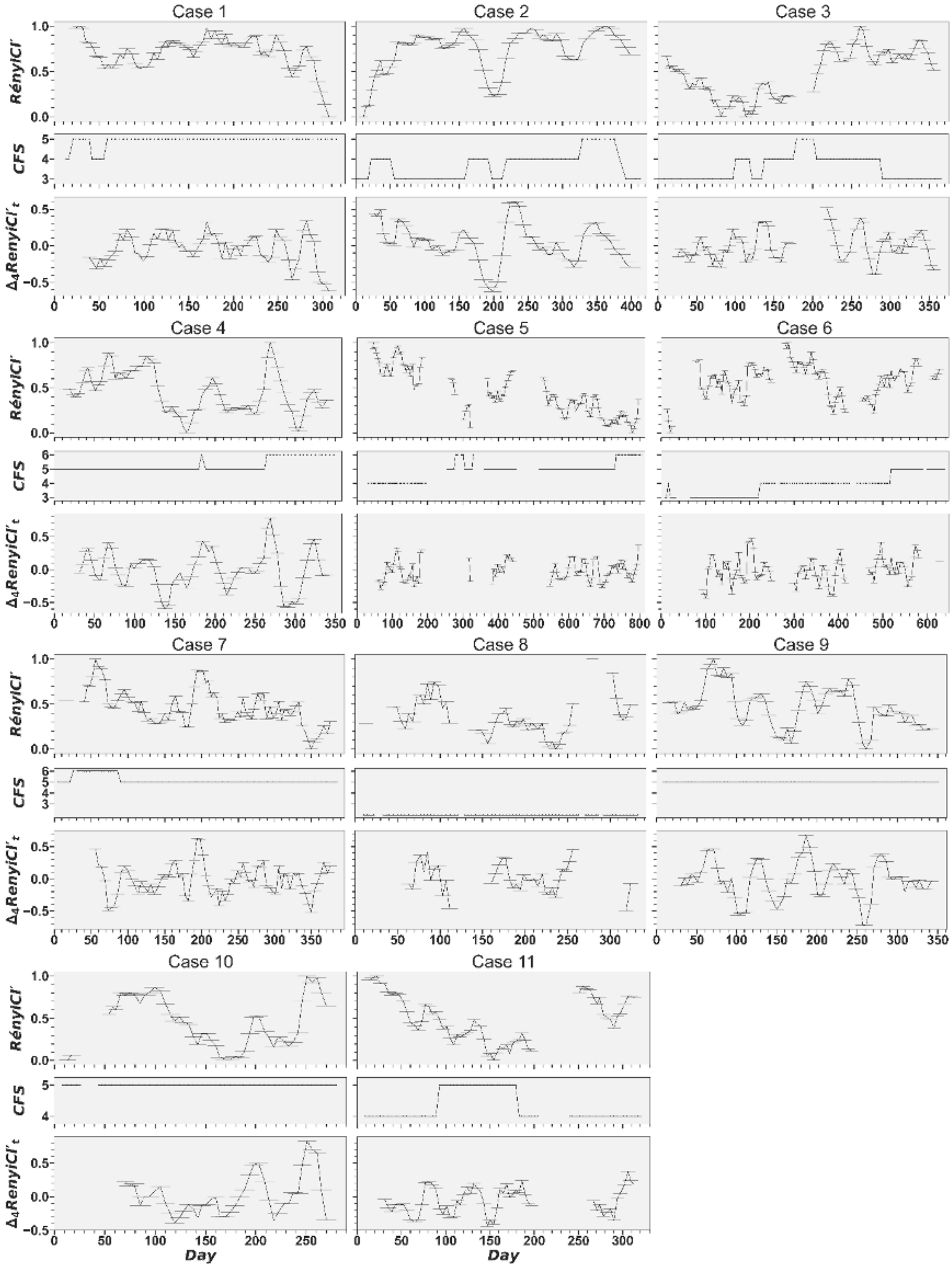
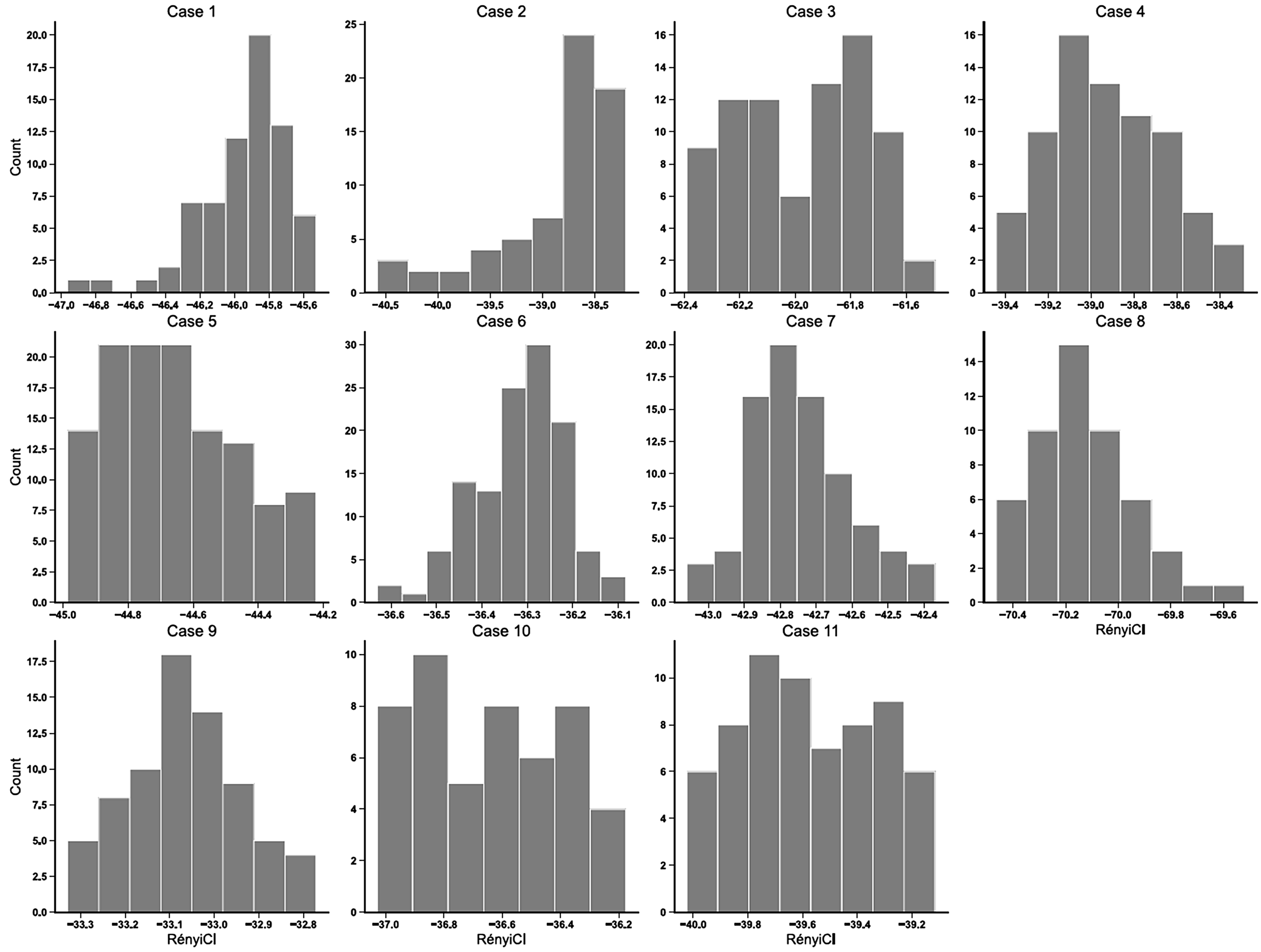


Figure 4. Histograms of RényiCI values.



Overall, case participants ranged from "fit" (CFS=2) to "living with moderate frailty" (CFS=6), although the trajectories of frailty within each participant varied considerably (Figure 3). For example, cases 2, 3, 7, and 11 all experienced periods of elevated frailty but all recovered and returned to baseline by the end of observation. Cases 4, 5, and 6 were the only cases that increased in frailty over time, ending frailer than their baseline. Most cases exhibited 2 - 3 transitions in frailty over time, with the extreme being case 2 with 7 frailty transitions.

While the goal of this study is not to directly infer CFS from RényiCI values, we note that Table 3 shows a significant correlation for all cases that have variable frailty scores. While they are significant, the correlations are mostly quite small. The overall correlation for all combined values is $\rho=-.055$ ($P<.001$). These results indicate that while a relationship between behavioral complexity and frailty can be observed, other factors must be considered when assessing a person's frailty from smart home sensor readings.

As Figure 3 demonstrates, trajectories of sensor complexity were similarly varied. RényiCI values for cases 1, 5, 7, and 9 exhibited downward trends over time, while cases 3 and 6 demonstrated an overall positive trend. Case 11 showed a concave shape with a general downward trend in sensor RényiCI for the first half of the data, followed by a general upward trend. The remaining cases exhibited nonmonotonic fluctuations. In each case, RényiCI values and frailty trajectories aligned with frailty scores assigned by the CFS-trained researchers during qualitative processing of clinical data.

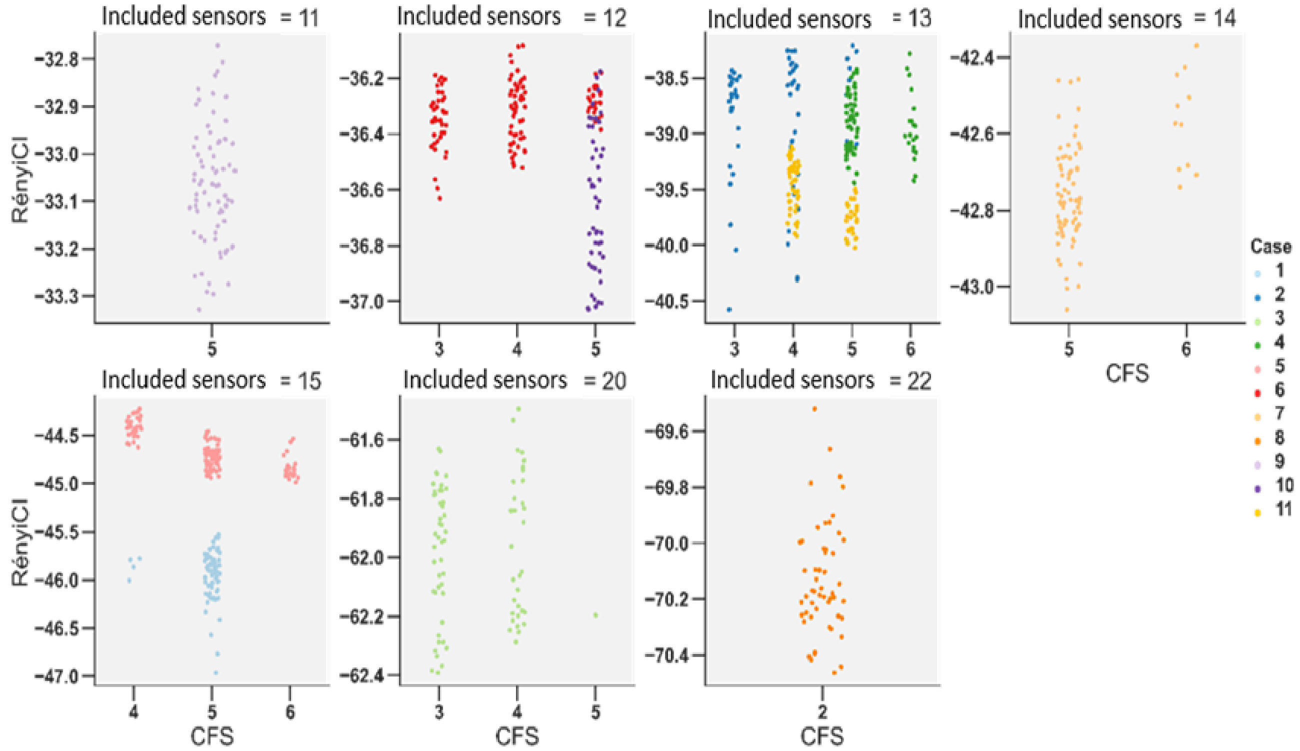
Frailty-Complexity Associations

The relationship between frailty level and sensor data complexity also varied from case to case. Figure 5 shows jitter plots of RényiCI values by CFS score grouped by the number of sensors installed in the home. Cases 5 and 11 exhibited a negative trend between complexity and frailty, while cases 3 and 7 demonstrated a mostly positive trend. The range and mean of USM-Rényi values shift farther from 0 as the number of sensors in the home increases. The range of RényiCI for the home with the fewest sensors (11 sensors) spanned approximately −34 to −32, while in the home with the most sensors (22 sensors), the range extended from about −71 to −69.

Initial runs tests applied to the full sequence of RényiCI values were statistically significant ($P<.001$) for all participants, suggesting nonrandom temporal ordering. To reduce potential autocorrelation introduced by overlapping windows, we repeated the runs tests on a downsampled sequence using every fourth window. Under this condition, only cases 3 ($P<.001$) and 5 ($P=.002$) remained statistically significant.

To account for these differences, Table 3 also includes the coefficient of variation (CoV) that normalizes the RényiCI SD by the mean for each case. Cases 6, 7, and 8 exhibited the least amount of relative variability in RényiCI (CoV=0.003), while case 2 exhibited the highest (CoV=0.015).

Figure 5. Categorical scatter (jitter) plot of RényiCI values by frailty, grouped by sensor count and case. CFS: Clinical Frailty Scale.



The cases with statistically significant nonoverlapping runs tests, cases 3 and 5, also represented the clearest long-term monotonic trends. Case 3 demonstrated an overall increase in complexity between the start and end of her data, while case 5 demonstrated an overall decreasing trend. The only RényiCI distribution with a statistically significant KS test, case 2, had some of the most extreme variation among the cases, with an extremely left-skewed distribution and a coefficient of variation 5 times greater than the smallest coefficient of variation among the cases. To explore possible explanations for the diverse patterns of frailty and complexity trajectories observed, we compared the frailty and complexity trajectories of the cases with contextual information derived from the nursing assessment records.

Data selection for cases 2 and 3 started at the beginning of the COVID-19 pandemic shutdowns. The horizontal bars in Figure 3 represent sliding window durations (in days). For both cases, the shortest windows are at the beginning of the time series. The shorter sliding window durations during this period likely reflect increased in-home activity, with participants generating more sensor events due to spending more time indoors during COVID-19 lockdowns. However, where RényiCI increased steadily over the coming months for case 2, RényiCI decreased steadily for case 3 (Figure 3).

Cases 2 and 3 also showed pronounced shifts in RényiCI midway through the observation period. For both cases, this period roughly correlates to a time period of hazardous air quality caused by continued wildfire smoke over the course of about a month. However, the steep dip in RényiCI for case 2 occurring between days 186 and 212 is short, while for case 3, the sharp increase in RényiCI around day 231 appears to be a vertical shift in their average complexity that continues for the rest of her data.

Case Narratives

Case 2

A woman in her 70s with congestive heart failure and mild asthma. She was independent at baseline (CFS=3) and in stable health, which is consistent with the increasing RényiCI values plotted in Figure 3 at the beginning of data collection. This participant experienced 3 episodes of worsening fatigue and shortness of breath (days 20 - 54, 160 - 215, 258 - 397), which impacted her ADLs and IADLs and contributed to transient increases in frailty. During the first episode, the nurse's report indicated that "walking has been much more taxing on her this past week. She will walk around the building today but runs out of energy very quickly... her fatigue level has increased significantly over the past week." During the second episode, the nurse reported "the last 3 days she noticed … more soa [sic, shortness of air] and tired[ness]." During the third episode, the nurse recorded a direct quote from the participant indicating that she had "absolute fatigue beyond anything I've experienced." Two of these periods coincided with substantial troughs in RényiCI, suggesting alignment between behavioral simplification and functional decline. The primary movement patterns manifested in the CFS score and measured by RényiCI (ie, mechanisms of interest) were less in-home overall activity and less time spent out of the home on walks and social activities (frequency and duration). This case also had the highest variability in complexity and was the only one with a statistically significant KS test result. Possible gerontological clinical actions informed by these results include recommendations to follow up with the cardiologist, referral to a pulmonologist, and referral to senior services to determine whether the patient qualifies for in-home care support services.

Case 3

A woman in her 80s with cardiovascular disease and allergy-induced asthma. Initially independent, she experienced progressive health decline, including 2 hospitalizations for acute hypertension and dyspnea. Her frailty peaked after the second hospitalization during a period of wildfire smoke, when she relied full-time on a walker for ambulation. Nursing records include statements during this timeframe like “no energy, has not left the house since Thursday [4 d]” and “overall health is declining.” Following physical therapy, she recovered and reported no activity limitations by day 295. The nursing record indicated that she “went shopping” and had “several visitors over” across multiple days of the week. Her RényiCI trajectory reflected this pattern. As shown in Figure 3, the RényiCI values initially show a steady decrease aligned with the health issues. After she received treatment and improved her ambulation and functional independence, the RényiCI values showed a steady rise in complexity. Notably, this case showed a strong monotonic increase in RényiCI and passed the runs test even under downsampling. This result provides evidence that the pattern of increasing frailty, followed by improvement after treatment, is distinct and nonrandom. The mechanism of interest impacting her RényiCI trajectory was a renewed increase in time spent out of the home (frequency and duration) concurrent with an increase in the number of visitors. The case exemplifies how RényiCI trajectories could help clinical gerontologists understand treatment efficacy through novel remote patient monitoring tools that include sensor monitoring and associated behavior patterns.

Case 5

A man in his 80s with Parkinson disease. He began with mild frailty (CFS=4) and was independent but slowed by symptoms. Over time, he required increasing assistance with ADLs and IADLs. Nurses recorded that he began to require assistance “getting compression sock on in the morning and off at night” and “needing help with laundry and housekeeping” and that his daughter began assisting with bill paying. He experienced multiple hospitalizations and rehabilitation stays and ultimately progressed to moderate frailty (CFS=6). The moderate frailty score was based on the nurse reporting “unsteady gait” and that he “has cracked ribs from a fall last week” and his “symptoms of PD [are] increasing, [and] noticeable upon observation.” His RényiCI trajectory followed a corresponding decline, with complexity peaking early and then falling across successive rehabilitation episodes. This case also exhibited a significant runs test and a clear downward trend in complexity. The mechanism of interest in this case was more overall time spent in his recliner chair, more nighttime sleeping in the recliner chair, and the decreased time spent out of the home (frequency and duration). This case illustrates how RényiCI trajectories may support automated smart home monitoring aimed at detecting increasing frailty upstream so interventions can be implemented.

Cases 4 and 10

Both cases involved sustained or increased caregiving over time. In case 4, RényiCI peaked just before caregiving began and declined thereafter. Case 10, who had consistent caregiving throughout, showed generally lower complexity than case 6, who lived alone with the same number of sensors. These comparisons suggest that increased caregiving frequency does not necessarily lead to increased behavioral complexity as measured by RényiCI. Older adults with consistent professional caregiving are likely to experience slower rates of decline due to the intentionality of caregiving, which aims to extend independence through building physical, functional, and cognitive strength. Findings could inform care planning and resource allocation.

Discussion

Principal Results

This study introduces and applies a novel entropy-based algorithm, the RényiCI, to quantify behavioral complexity from smart home sensor data in older adults. Using a USM framework with multiscale Rényi entropy, our method captures subtle temporal dynamics in sensor-derived movement sequences. In this exploratory case series, within-person indoor-movement complexity, as exhibited by RényiCI values, fluctuated over time. In several cases, these fluctuations coincided with frailty changes.

Across 11 participants, we observed diverse complexity trajectories, ranging from steady increases, steady declines, and concave patterns to nonmonotonic fluctuations. Case-level analysis revealed that greater fluctuations in complexity were frequently aligned with periods of functional decline or recovery. Notably, 2 cases (3 and 5) exhibited statistically significant nonrandom patterns in complexity over time, confirmed by runs tests on downsampled data, and showed clear monotonic trends in behavior complexity that matched health trajectories. Only one case (2) showed a RényiCI distribution that deviated significantly from normality, corresponding with extreme within-person variability and periods of worsening frailty. Scatter plots further revealed heterogeneous associations between complexity and frailty, with both positive and negative trends across cases. Importantly, increased caregiver presence was not associated with greater behavioral complexity, suggesting that RényiCI may reflect intrinsic changes in individual functional capacity rather than external support.

Changes in CFS scores fluctuated in alignment with changes in RényiCI values for some cases, like 2, 3, 6, and 7. These cases may suggest that changes in frailty do impact the person’s behavioral routine and regularity. However, in cases 8 through 10, we observed changes in RényiCI values despite the lack of change in frailty scores. This observation highlights the fact that our findings provide 1 set of indicators of changes in frailty, but should not be analyzed in isolation. Other factors, such as visitors, seasonal effects, and external events, can also impact behavioral routines. These should be controlled for when examining frailty as a function of changes in RényiCI.

While observed repetitive behavior may correlate with frailty, the relationship is not one-to-one. Reduced complexity of movement often, but not always, aligns with frailty progression. Repetitive behavior can signal frailty because it reflects narrowed activity routines, reduced introduction of new routine elements, and corresponding reduced adaptability. At the same time, we note that complexity is multifactorial. Other influences, such as visitors in the home, home layout, and external events, also affected the entropy measures. The results showed nonmonotonic relationships in those cases. To interpret complexity, it is therefore best to consider an individual over time rather than compare cross-sectionally. Moreover, interventions aimed at slowing the impact of frailty on maintaining independence, like a smart home that projects RényiCI trajectories, would be more helpful for older adults living alone. Mechanisms of interest become difficult to automatically recognize in multiresident homes where ambient sensors detect movement from all residents.

Limitations

Several factors impacted the interpretation and generalizability of our findings. First, entropy-based measures like RényiCI are inherently sensitive to sample length and the number of sensors deployed in a participant's home. To prioritize intraindividual validity, we customized the sliding window size for each participant using a fixed number of sensor messages (n), rather than a fixed time duration. This approach allowed for consistent comparisons within individuals but introduced variability in the time span covered by each window, both within and across cases, limiting our ability to analyze complexity as a direct function of chronological time. Future work could develop correction factors for RényiCI to account for sample length, enabling the detection of periodic, seasonal, or event-driven patterns in indoor behavior.

Relatedly, interindividual comparisons were constrained by differences in sensor configurations across homes. Participants varied in the number and placement of sensors, affecting both the density of event data and the scale of RényiCI values. Standardizing sensor deployments in future studies would facilitate more robust cross-participant comparisons and support investigation into whether home-level sensor complexity systematically relates to frailty markers at the population level.

The impact of the COVID-19 pandemic further complicates interpretation. Several participants were enrolled during or shortly after the onset of pandemic-related lockdowns, which led to changes in daily routines, increased time spent indoors, and potentially long-term shifts in behavior and social support. These behavioral changes may have altered both the complexity of movement and its relationship to frailty. Additionally, one period of the study coincided with prolonged hazardous air quality due to regional wildfires, which may have further restricted participants' movement and contributed to abrupt changes in sensor complexity. Such exogenous events likely altered daily routines independent of health. We therefore interpret Rényi changes within homes and in the presence of annotated event periods. We also provide event-excluded sensitivities to reduce confounding.

Sensor noise and dropout also presented challenges. While preprocessing steps excluded known periods of sensor failure or participant absence, subtle forms of sensor drift or inconsistent message delivery could still introduce noise into the RényiCI estimates. Further improvements to sensor reliability and the integration of sensor health metrics into complexity analysis pipelines would strengthen future research.

In terms of statistical methods, the runs test was useful in identifying nonrandom patterns in behavioral complexity over time, but it is not well-suited to detecting more complex temporal structures such as oscillatory or nonlinear trends. Future research may benefit from time series models drawn from signal processing or machine learning that can more precisely characterize evolving behavioral dynamics.

Frailty measurement also posed a limitation. Because frailty was not a primary outcome in the parent study, we relied on retrospective CFS scoring based on weekly nursing reports. This limits temporal precision and may miss subtle fluctuations in functional status. Larger-scale studies using prospectively collected frailty data, including both clinician-reported and self-reported measures, could reveal more detailed associations between complexity and health.

Finally, this sample was racially and culturally homogeneous, limiting the generalizability of our findings. RényiCI analyses should be interpreted as a within-home monitoring signal rather than a cross-sectional diagnostic tool. As efforts to diversify smart home research populations expand, it will be essential to explore whether the relationships between sensor-derived behavioral complexity and frailty differ across racial, cultural, and socioeconomic groups. Inclusive, representative samples are critical to ensuring that digital biomarkers are both effective and equitable.

Conclusions

Detection of incipient frailty in community-dwelling older adults is a key component to supporting their independence. The findings in this study demonstrate that RényiCI, as a passive and unobtrusive complexity metric, offers a promising tool for monitoring functional health changes in aging populations and may help enable early detection of frailty in real-world settings. The PyUSM software package developed for this analysis is publicly available and supports future application of this method in diverse behavioral monitoring contexts. These findings support the potential of entropy-based digital behavior markers to unobtrusively monitor intraindividual health changes and capture early signs of frailty in aging-in-place.

Future enhancements of this analysis may reveal additional factors that influence change in indoor movement complexity and inform how the complexity of smart home data may inform clinical practice. For example, significant departures of RényiCI values from a person's complexity baseline may trigger a nurse call or follow-up when integrated into a remote monitoring or telemonitoring system. In routine care, weekly summaries of the analysis would support triage and help care providers select appropriate actions. Additionally, when performing a functional assessment of an individual, a summary of the complexity trend augments traditional frailty analysis to improve assessment and treatment options. Future work will also emphasize analytical validity (repeatability and robustness across sensors and windowing), clinical validity (prospective prediction of frailty transitions), and clinical use (impact on downstream outcomes such as unplanned care, falls, and functional decline) for diverse homes and populations.

Additionally, future work should focus on integrating RényiCI in machine learning predictive modeling as a high-level feature to assist with identifying meaningful digital biomarkers [49]. Other temporal activities associated with frailty (eg, walking speed, ADL, and IADL behaviors) could also be integrated to optimize frailty classifications. Machine learning integration of features from RényiCI values that signal possible increasing frailty will support nurses and caregivers in providing timely interventions, thereby potentially extending independence and optimizing older adults' outcomes.

Acknowledgments

The authors would like to thank Jason Minor, Bryan Minor, and Brian Thomas for their assistance with the smart homes and Shandeigh Nikki Berry, Samantha Denison, and Ellen Hinderlie for their data collection contributions.

Funding

This work is supported in part by the National Institute of Nursing Research under grant R01NR016732, Touchmark Foundation, and the Washington State University College of Nursing's Linblad Scholarship funds.

Data Availability

Sensor data are available at CASAS DataDownloader [50]. Deidentified medical information about frailty cases included in this study may be made available for authorized research use upon request.

Conflicts of Interest

None declared.

References

1. Kim DH, Rockwood K. Frailty in older adults. N Engl J Med 2024 Aug 8;391(6):538-548. [doi: 10.1056/NEJMra2301292] [Medline: 39115063]
2. Kurnat-Thoma EL, Murray MT, Juneau P. Frailty and determinants of health among older adults in the United States 2011-2016. J Aging Health 2022 Mar;34(2):233-244. [doi: 10.1177/08982643211040706] [Medline: 34470533]
3. Jang J, Kim A, Choi M, et al. Association of Frailty Index at 66 years of age with health care costs and utilization over 10 years in Korea: retrospective cohort study. JMIR Public Health Surveill 2025 Jan 27;11:e50026. [doi: 10.2196/50026] [Medline: 39874179]
4. Ensrud KE, Schousboe JT, Kats AM, Taylor BC, Boyd CM, Langsetmo L. Incremental health care costs of self-reported functional impairments and phenotypic frailty in community-dwelling older adults: a prospective cohort study. Ann Intern Med 2023 Apr;176(4):463-471. [doi: 10.7326/M22-2626] [Medline: 37011386]
5. Crocker TF, Brown L, Clegg A, et al. Quality of life is substantially worse for community-dwelling older people living with frailty: systematic review and meta-analysis. Qual Life Res 2019 Aug;28(8):2041-2056. [doi: 10.1007/s11136-019-02149-1] [Medline: 30875008]
6. Lin YC, Yan HT. Frailty phenotypes and their association with health consequences: a comparison of different measures. Aging Clin Exp Res 2024 Dec 3;36(1):233. [doi: 10.1007/s40520-024-02887-4] [Medline: 39625598]
7. Fletcher J, Reid N, Hubbard RE, et al. Frailty Index, not age, predicts treatment outcomes and adverse events for older adults with cancer. J Frailty Aging 2024;13(4):487-494. [doi: 10.14283/jfa.2024.22] [Medline: 39574272]
8. Yixiao C, Hui S, Quhong S, Xiaoxi Z, Jirong Y. A review of utility of wearable sensor technologies for older person frailty assessment. Exp Gerontol 2025 Feb;200:112668. [doi: 10.1016/j.exger.2024.112668] [Medline: 39733783]
9. Woo J, Yu R, Tsoi K, Meng H. Variability in repeated blood pressure measurements as a marker of frailty. J Nutr Health Aging 2018;22(9):1122-1127. [doi: 10.1007/s12603-018-1082-9] [Medline: 30379313]
10. Fried LP, Cohen AA, Xue QL, Walston J, Bandeen-Roche K, Varadhan R. The physical frailty syndrome as a transition from homeostatic symphony to cacophony. Nat Aging 2021 Jan;1(1):36-46. [doi: 10.1038/s43587-020-00017-z] [Medline: 34476409]
11. Coravos A, Khozin S, Mandl KD. Developing and adopting safe and effective digital biomarkers to improve patient outcomes. NPJ Digit Med 2019;2(1):1-5. [doi: 10.1038/s41746-019-0090-4] [Medline: 30868107]
12. Beattie Z, Miller LM, Almirola C, et al. The collaborative aging research using technology initiative: an open, sharable, technology-agnostic platform for the research community. Digit Biomark 2020;4(Suppl 1):100-118. [doi: 10.1159/000512208] [Medline: 33442584]
13. Ageing and health. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-12-11]

14. Cobo A, Villalba-Mora E, Pérez-Rodríguez R, Ferre X, Rodríguez-Mañas L. Unobtrusive sensors for the assessment of older adult's frailty: a scoping review. Sensors (Basel) 2021 Apr 23;21(9):2983. [doi: 10.3390/s21092983] [Medline: 33922852]
15. Bian C, Ye B, Hoonakker A, Mihailidis A. Attitudes and perspectives of older adults on technologies for assessing frailty in home settings: a focus group study. BMC Geriatr 2021 May 8;21(1):298. [doi: 10.1186/s12877-021-02252-4] [Medline: 33964887]
16. Cook DJ, Crandall AS, Thomas BL, Krishnan NC. CASAS: a smart home in a box. Computer (Long Beach Calif) 2013 Jul;46(7):62-69. [doi: 10.1109/MC.2012.328] [Medline: 24415794]
17. Lipsitz LA. Dynamics of stability: the physiologic basis of functional health and frailty. J Gerontol A Biol Sci Med Sci 2002 Mar;57(3):B115-B125. [doi: 10.1093/gerona/57.3.b115] [Medline: 11867648]
18. Hao M, Zhang H, Li Y, et al. Using physiological system networks to elaborate resilience across frailty states. J Gerontol A Biol Sci Med Sci 2024 Oct 1;79(10):1-7. [doi: 10.1093/gerona/glad243] [Medline: 37824088]
19. Cohen AA, Ferrucci L, Fülöp T, et al. A complex systems approach to aging biology. Nat Aging 2022 Jul;2(7):580-591. [doi: 10.1038/s43587-022-00252-6] [Medline: 37117782]
20. Bizovska L, Svoboda Z, Vuillerme N, Janura M. Multiscale and Shannon entropies during gait as fall risk predictors-a prospective study. Gait Posture 2017 Feb;52:5-10. [doi: 10.1016/j.gaitpost.2016.11.009] [Medline: 27842283]
21. Castiglia SF, Trabassi D, Conte C, et al. Multiscale entropy algorithms to analyze complexity and variability of trunk accelerations time series in subjects with Parkinson's disease. Sensors (Basel) 2023 May 22;23(10):4983. [doi: 10.3390/s23104983] [Medline: 37430896]
22. Gao C, Lim ASP, Haghayegh S, et al. Reduced complexity of pulse rate is associated with faster cognitive decline in older adults. J Am Heart Assoc 2025 May 20;14(10):e041448. [doi: 10.1161/JAHA.125.041448] [Medline: 40331928]
23. Ma Y, Zhou J, Kavousi M, et al. Lower complexity and higher variability in beat - to - beat systolic blood pressure are associated with elevated long - term risk of dementia: the Rotterdam study. Alzheimer's & Dementia 2021 Jul;17(7):1134-1144. [doi: 10.1002/alz.12288] [Medline: 33860609]
24. Zhou J, Habtemariam D, Iloputaife I, Lipsitz LA, Manor B. The complexity of standing postural sway associates with future falls in community-dwelling older adults: the MOBILIZE Boston study. Sci Rep 2017 Jun 7;7(1):2924. [doi: 10.1038/s41598-017-03422-4] [Medline: 28592844]
25. Zhou J, Poole V, Wooten T, et al. Multiscale dynamics of spontaneous brain activity is associated with walking speed in older adults. J Gerontol A Biol Sci Med Sci 2020 Jul 13;75(8):1566-1571. [doi: 10.1093/gerona/glz231] [Medline: 31585008]
26. Rector JL, Gijzel SMW, van de Leemput IA, van Meulen FB, Olde Rikkert MGM, Melis RJF. Dynamical indicators of resilience from physiological time series in geriatric inpatients: lessons learned. Exp Gerontol 2021 Jul 1;149:111341. [doi: 10.1016/j.exger.2021.111341] [Medline: 33838217]
27. Raichlen DA, Klimentidis YC, Hsu CH, Alexander GE. Fractal complexity of daily physical activity patterns differs with age over the life span and is associated with mortality in older adults. J Gerontol A Biol Sci Med Sci 2019 Aug 16;74(9):1461-1467. [doi: 10.1093/gerona/gly247] [Medline: 30371743]
28. Schütz N, Knobel SEJ, Botros A, et al. A systems approach towards remote health-monitoring in older adults: introducing a zero-interaction digital exhaust. NPJ Digit Med 2022 Aug 16;5(1):116. [doi: 10.1038/s41746-022-00657-y] [Medline: 35974156]
29. Cook DJ, Schmitter-Edgecombe M. Fusing ambient and mobile sensor features into a behaviorome for predicting clinical health scores. IEEE Access 2021;9:65033-65043. [doi: 10.1109/access.2021.3076362] [Medline: 34017671]
30. Wang T, Cook DJ, Fischer TR. The indoor predictability of human mobility: estimating mobility with smart home sensors. IEEE Trans Emerg Top Comput 2023;11(1):182-193. [doi: 10.1109/tetc.2022.3188939] [Medline: 37457914]
31. Howedi A, Lotfi A, Pourabdollah A. Employing entropy measures to identify visitors in multi-occupancy environments. J Ambient Intell Human Comput 2022 Feb;13(2):1093-1106. [doi: 10.1007/s12652-020-02824-z]
32. Takahashi J, Kawai H, Ejiri M, et al. Activity diversity is associated with the prevention of frailty in community-dwelling older adults: the Otassha study. Front Public Health 2023;11:1113255. [doi: 10.3389/fpubh.2023.1113255] [Medline: 37033071]
33. Gopalratnam K, Cook D. Online sequential prediction via incremental parsing: the active LeZi algorithm. IEEE Intell Syst 2007;22(1):52-58. [doi: 10.1109/MIS.2007.15]
34. Bai C, Mardini MT. Advances of artificial intelligence in predicting frailty using real-world data: a scoping review. Ageing Res Rev 2024 Nov;101:102529. [doi: 10.1016/j.arr.2024.102529] [Medline: 39369796]
35. Park C, Mishra R, Golledge J, Najafi B. Digital biomarkers of physical frailty and frailty phenotypes using sensor-based physical activity and machine learning. Sensors (Basel) 2021 Aug 5;21(16):5289. [doi: 10.3390/s21165289] [Medline: 34450734]
36. Osuka Y, Chan LLY, Brodie MA, Okubo Y, Lord SR. A wrist-worn wearable device can identify frailty in middle-aged and older adults: the UK Biobank study. J Am Med Dir Assoc 2024 Oct;25(10):105196. [doi: 10.1016/j.jamda.2024.105196] [Medline: 39128825]
37. Hosseinalizadeh M, Asghari M, Toosizadeh N. Sensor-based frailty assessment using fitbit. Sensors (Basel) 2024 Dec 7;24(23):7827. [doi: 10.3390/s24237827] [Medline: 39686364]

38. Merchant RA, Loke B, Chan YH. Ability of heart rate recovery and gait kinetics in a single wearable to predict frailty: quasiexperimental pilot study. JMIR Form Res 2024 Oct 3;8:e58110. [doi: 10.2196/58110] [Medline: 39361400]
39. Ding H, Ho K, Searls E, et al. Assessment of wearable device adherence for monitoring physical activity in older adults: pilot cohort study. JMIR Aging 2024 Oct 25;7:e60209. [doi: 10.2196/60209] [Medline: 39454101]
40. Tannou T, Lihoreau T, Gagnon-Roy M, Grondin M, Bier N. Effectiveness of smart living environments to support older adults to age in place in their community: an umbrella review protocol. BMJ Open 2022 Jan 25;12(1):e054235. [doi: 10.1136/bmjopen-2021-054235] [Medline: 35078843]
41. Fritz R, Wuestney K, Dermody G, Cook DJ. Nurse-in-the-loop smart home detection of health events associated with diagnosed chronic conditions: a case-event series. Int J Nurs Stud Adv 2022 Dec;4:100081. [doi: 10.1016/j.ijnsa.2022.100081] [Medline: 35642184]
42. Wang Y, Yalcin A, VandeWeerd C. An entropy-based approach to the study of human mobility and behavior in private homes. PLOS ONE 2020;15(12):e0243503. [doi: 10.1371/journal.pone.0243503] [Medline: 33301515]
43. CASAS datasets. Center for Advanced Studies in Adaptive Systems (CASAS). URL: https://casas.wsu.edu/datasets/ [accessed 2025-11-04]
44. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005 Aug 30;173(5):489-495. [doi: 10.1503/cmaj.050051] [Medline: 16129869]
45. Tamamura Y, Hachiuma C, Matsuura M, Shiba S, Nishikimi T. Frailty and energy intake deficiency reduce the efficiency of activities of daily living in patients with musculoskeletal disorders: a retrospective cohort study. Nutrients 2025 Apr 12;17(8):1334. [doi: 10.3390/nu17081334] [Medline: 40284199]
46. Almeida JS, Vinga S. Universal sequence map (USM) of arbitrary discrete sequences. BMC Bioinformatics 2002;3(1):6. [doi: 10.1186/1471-2105-3-6] [Medline: 11895567]
47. Vinga S, Almeida JS. Rényi continuous entropy of DNA sequences. J Theor Biol 2004 Dec 7;231(3):377-388. [doi: 10.1016/j.jtbi.2004.06.030] [Medline: 15501469]
48. Wuestney K. Demo of Universal Sequence Maps (USM) in Python. Documentation of pyusm Package. 2022. URL: https://katherine983.github.io/pyusm/demo_usm.html [accessed 2025-09-25]
49. Kong LN, Yang L, Lyu Q, Liu DX, Yang J. Risk prediction models for frailty in older adults: a systematic review and critical appraisal. Int J Nurs Stud 2025 Jul;167:105068. [doi: 10.1016/j.ijnurstu.2025.105068] [Medline: 40184783]
50. CASAS DataDownloader. URL: https://data.casas.wsu.edu/ [accessed 2025-12-23]

Abbreviations

ADL: activity of daily living
CASAS: Center for Advanced Studies in Adaptive Systems
CFS: Clinical Frailty Scale
CoV: coefficient of variance
IADL: instrumental activity of daily living
KS: Kolmogorov-Smirnov
MSE: multiscale entropy
RényiCI: Rényi Complexity Index
USM: Universal Sequence Map

Edited by R Yang; submitted 11.May.2025; peer-reviewed by J Wang, R Buenrostro; revised version received 22.Nov.2025; accepted 27.Nov.2025; published 02.Jan.2026.

Please cite as:
Wuestney K, Cook D, Van Son C, Fritz R
Using Indoor Movement Complexity in Smart Homes to Detect Frailty in Older Adults: Multiple-Methods Case Series Study
JMIR Aging 2026;9:e77322
URL: https://aging.jmir.org/2026/1/e77322
doi:10.2196/77322

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org