JMIR Aging

Journal Impact Factor (JIF) (2023): 5.0
Volume 5 (2022), Issue 3 ISSN 2561-7605 Editor in Chief: Yun Jiang, PhD, MS, RN, FAMIA; Jinjiao Wang, PhD, RN, MPhil

Contents

Original Papers

Mitigating Feelings of Loneliness and Depression by Means of Web-Based or Print-Based Physical Activity Interventions: Pooled Analysis of 2 Community-Based Intervention Trials (e36515)
Sonia Lippke, Tiara Ratz, Franziska Keller, Dennis Juljugin, Manuela Peters, Claudia Pischke, Claudia Voelcker-Rehage. . . . 4

Objectively Measured Physical Activity Levels and Associated Factors in Older US Women During the COVID-19 Pandemic: Cross-sectional Study (e38172)
Renoa Choudhury, Joon-Hyuk Park, Ladda Thiamwong, Rui Xie, Jeffrey Stout. . . . 26

Implementing Affordable Socially Assistive Pet Robots in Care Homes Before and During the COVID-19 Pandemic: Stratified Cluster Randomized Controlled Trial and Mixed Methods Study (e38864)
Hannah Bradwell, Katie Edwards, Rhona Winnington, Serge Thill, Victoria Allgar, Ray Jones. . . . 42

Mobile Videoconferencing for Occupational Therapists' Assessments of Patients' Home Environments Prior to Hospital Discharge: Mixed Methods Feasibility and Comparative Study (e24376)
Karine Latulippe, Dominique Giroux, Manon Guay, Dahlia Kairy, Claude Vincent, Katia Boivin, Ernesto Morales, Natasa Obradovic, Véronique Provencher. . . . 61

Advance Planning for Technology Use in Dementia Care: Development, Design, and Feasibility of a Novel Self-administered Decision-Making Tool (e39335)
Clara Berridge, Natalie Turner, Liu Liu, Sierramatice Karras, Amy Chen, Karen Fredriksen-Goldsen, George Demiris. . . . 92

A Web-Based Platform (CareVirtue) to Support Caregivers of People Living With Alzheimer Disease and Related Dementias: Mixed Methods Feasibility Study (e36975)
Justin Boutilier, Priya Loganathar, Anna Linden, Eleanore Scheer, Sofia Noejovich, Christian Elliott, Matthew Zuraw, Nicole Werner. . . . 106

Identifying Caregiver Availability Using Medical Notes With Rule-Based Natural Language Processing: Retrospective Cohort Study (e40241)
Elham Mahmoudi, Wenbo Wu, Cyrus Najarian, James Aikens, Julie Bynum, V Vydiswaran. . . . 129

A Social Media–Based Intervention for Chinese American Caregivers of Persons With Dementia: Protocol Development (e40171)
Y Hong, Kang Shen, Huixing Lu, Hsiaoyin Chen, Yang Gong, Van Ta Park, Hae-Ra Han. . . . 140

App-Based Evaluation of Older People's Fall Risk Using the mHealth App Lindera Mobility Analysis: Exploratory Study (e36872)
Nicole Strutz, Hanna Brodowski, Joern Kiselev, Anika Heimann-Steinert, Ursula Müller-Werdan. . . . 149

XSL•FO
RenderX

An Unsupervised Data-Driven Anomaly Detection Approach for Adverse Health Conditions in People Living With Dementia: Cohort Study (e38211)
Nivedita Bijlani, Ramin Nilforooshan, Samaneh Kouchaki. 159

Understanding the Experience of Geriatric Care Professionals in Using Telemedicine to Care for Older Patients in Response to the COVID-19 Pandemic: Mixed Methods Study (e34952)
Wenwen Chen, Ashley Flanagan, Pria Nippak, Michael Nicin, Samir Sinha. 180

Implementing Technology Literacy Programs in Retirement Homes and Residential Care Facilities: Conceptual Framework (e34997)
Karen Li, Nathan Nagallo, Erica McDonald, Colin Whaley, Kelly Grindrod, Karla Boluk. 196

Examining Structural Disparities in US Nursing Homes: National Survey of Health Information Technology Maturity (e37482)
Gregory Alexander, Jianfang Liu, Kimberly Powell, Patricia Stone. 208

The Effects of a Digital Well-being Intervention on Older Adults: Retrospective Analysis of Real-world User Data (e39851)
Eliane Boucher, Ryan Honomichl, Haley Ward, Tyler Powell, Sarah Stoeckl, Acacia Parks. 218

Scaling Up Citizen Workshops in Public Libraries to Disseminate and Discuss Primary Care Research Results: Quasi-Experimental Study (e39016)
José Massougbodji, Hervé Zomahoun, Evehouenou Adisso, Jasmine Sawadogo, Valérie Borde, Cynthia Cameron, Hélène Moisan, Jean-Sébastien Paquette, Zamzam Akbaraly, Lëa-Kim Châteauneuf, Geneviève David, Bibliothèques2 Group, France Légaré. 231

A Smartwatch Step-Counting App for Older Adults: Development and Evaluation Study (e33845)
George Boateng, Curtis Petersen, David Kotz, Karen Fortuna, Rebecca Masutani, John Batsis. 245

Validation of a Remote and Fully Automated Story Recall Task to Assess for Early Cognitive Impairment in Older Adults: Longitudinal Case-Control Observational Study (e37090)
Caroline Skirrow, Marton Meszaros, Udeepa Meepegama, Raphael Lenain, Kathryn Papp, Jack Weston, Emil Fristed. 256

Evaluating Web-Based Automatic Transcription for Alzheimer Speech Data: Transcript Comparison and Machine Learning Analysis (e33460)
Thomas Soroski, Thiago da Cunha Vasco, Sally Newton-Mason, Saffrin Granby, Caitlin Lewis, Anuj Harisinghani, Matteo Rizzo, Cristina Conati, Gabriel Murray, Giuseppe Carenini, Thalia Field, Hyeju Jang. 272

Shared Decision-Making Training for Home Care Teams to Engage Frail Older Adults and Caregivers in Housing Decisions: Stepped-Wedge Cluster Randomized Trial (e39386)
Évèhouénou Adisso, Monica Taljaard, Dawn Stacey, Nathalie Brière, Hervé Zomahoun, Pierre Durand, Louis-Paul Rivest, France Légaré. . .
3 4 9

Reviews

A Conceptual Model of Experiences With Digital Technologies in Aging in Place: Qualitative Systematic Review and Meta-synthesis (e34872)
Mareike Hechinger, Diana Hentschel, Christine Aumer, Christian Rester. 74

mHealth Interventions to Support Caregivers of Older Adults: Equity-Focused Systematic Review (e33085)
Anna Garnett, Melissa Northwood, Justine Ting, Ruheena Sangrar. 286

XSL•FO
RenderX

Online Peer Support for People With Parkinson Disease: Narrative Synthesis Systematic Review (e35425)
Esther Gerritzen, Abigail Lee, Orii McDermott, Neil Coulson, Martin Orrell. 310

Short Paper

Automatically Identifying Twitter Users for Interventions to Support Dementia Family Caregivers: Annotated Data Set and Benchmark Classification Models (e39547)
Ari Klein, Arjun Magge, Karen O'Connor, Graciela Gonzalez-Hernandez. 123

Viewpoint

Capturing Cognitive Aging in Vivo: Application of a Neuropsychological Framework for Emerging Digital Tools (e38130)
Katherine Hackett, Tania Giovannetti. 327

Corrigenda and Addendas

Correction: Web-Based Delivery of the Caregiving Essentials Course for Informal Caregivers of Older Adults in Ontario: Mixed Methods Evaluation Study (e42215)
Shelley Rottenberg, Allison Williams. 344

Correction: The Effect of Cognitive Function Health Care Using Artificial Intelligence Robots for Older Adults: Systematic Review and Meta-analysis (e42312)
Hocheol Lee, Min Chung, Hyeji Kim, Eun Nam. 345

Correction: A Web-Based Platform (CareVirtue) to Support Caregivers of People Living With Alzheimer Disease and Related Dementias: Mixed Methods Feasibility Study (e41912)
Justin Boutilier, Priya Loganathar, Anna Linden, Eleanore Scheer, Sofia Noejovich, Christian Elliott, Matthew Zuraw, Nicole Werner. 347

XSL•FO
RenderX

Original Paper

Mitigating Feelings of Loneliness and Depression by Means of Web-Based or Print-Based Physical Activity Interventions: Pooled Analysis of 2 Community-Based Intervention Trials

Sonia Lippke[1], PhD; Tiara Ratz[2], PhD; Franziska Maria Keller[1], MSc; Dennis Juljugin[3], BSc; Manuela Peters[4], MSc; Claudia Pischke[5*], PhD; Claudia Voelcker-Rehage[6*], PhD

[1]Psychology & Methods, Jacobs University Bremen, Bremen, Germany

[2]Department of Reproductive Endocrinology, University Hospital Zurich (USZ), University of Zurich (UZH), Zurich, Switzerland

[3]Psychology, Universitaet Bremen, Bremen, Germany

[4]Leibniz Institute for Prevention Research and Epidemiology–BIPS, Bremen, Germany

[5]Institute of Medical Sociology, Centre for Health and Society, Medical Faculty, Heinrich Heine University Duesseldorf, Duesseldorf, Germany

[6]Department of Neuromotor Behavior and Exercise, Institute of Sport and Exercise Sciences, University of Muenster, Muenster, Germany

[*]these authors contributed equally

Corresponding Author:
Sonia Lippke, PhD
Psychology & Methods
Jacobs University Bremen
Campus Ring 1
Bremen, 28759
Germany
Phone: 49 421200 ext 4730
Email: s.lippke@jacobs-university.de

Abstract

Background: Physical activity (PA) is associated with benefits, such as fewer depressive symptoms and loneliness. Web- and print-based PA interventions can help older individuals accordingly.

Objective: We aimed to test the following research questions: Do PA interventions delivered in a web- or print-based mode improve self-reported PA stage of change, social-cognitive determinants of PA, loneliness, and symptoms of depression? Is subjective age a mediator and stage of change a moderator of this effect?

Methods: Overall, 831 adults aged ≥60 years were recruited and either allocated to a print-based or web-based intervention group or assigned to a wait-list control group (WLCG) in 2 community-based PA intervention trials over 10 weeks. Missing value imputation using an expectation-maximization algorithm was applied. Frequency analyses, multivariate analyses of variance, and moderated mediation analyses were conducted.

Results: The web-based intervention outperformed (47/59, 80% of initially inactive individuals being adopters, and 396/411, 96.4% of initially active individuals being maintainers of the recommended PA behavior) the print-based intervention (20/25, 80% of adopters, and 63/69, 91% of maintainers) and the WLCG (5/7, 71% of adopters; 141/150, 94% of maintainers). The pattern regarding adopters was statistically significant (web vs print Z=–1.94; P=.02; WLCG vs web Z=3.8367; P=.01). The pattern was replicated with stages (χ^2_4=79.1; P<.001; contingency coefficient 0.314; P<.001); in the WLCG, 40.1% (63/157) of the study participants moved to or remained in action stage. This number was higher in the groups receiving web-based (357/470, 76%) or print-based interventions (64/94, 68.1%). A significant difference was observed favoring the 2 intervention groups over and above the WLCG ($F_{19, 701}$=4.778; P<.001; η^2=0.098) and a significant interaction of time and group ($F_{19, 701}$=2.778; P<.001; η^2=0.070) for predictors of behavior. The effects of the interventions on subjective age, loneliness, and depression revealed that both between-group effects ($F_{3, 717}$=8.668; P<.001; η^2=0.018) and the interaction between group and time were significant ($F_{3, 717}$=6.101; P<.001; η^2=0.025). In a moderated mediation model, both interventions had a significant direct effect on depression in comparison with the WLCG (web-based: c' path −0.86, 95% CI −1.58 to −0.13, SE 0.38; print-based: c' path −1.96, 95% CI −2.99 to −0.92, SE 0.53). Furthermore, subjective age was positively related to depression (b path 0.14, 95% CI 0.05-0.23; SE

XSL•FO
RenderX

0.05). An indirect effect of the intervention on depression via subjective age was only present for participants who were in actor stage and received the web-based intervention (*ab* path −0.14, 95% CI −0.34 to −0.01; SE 0.09).

Conclusions: Web-based interventions appear to be as effective as print-based interventions. Both modes might help older individuals remain or become active and experience fewer depression symptoms, especially if they feel younger.

Trial Registration: German Registry of Clinical Trials DRKS00010052 (PROMOTE 1); https://tinyurl.com/nnzarpsu and DRKS00016073 (PROMOTE 2); https://tinyurl.com/4fhcvkwy

International Registered Report Identifier (IRRID): RR2-10.2196/15168

(JMIR Aging 2022;5(3):e36515) doi:10.2196/36515

KEYWORDS

physical activity; older adults; intervention; loneliness; depression; eHealth; mobile health; mHealth

Introduction

Web-Based and Printed Intervention Material

Loneliness is a key element, along with lifestyle factors such as physical activity (PA), which is interrelated with health and well-being [1-3]. Although the concept of loneliness has a long history, many concerns exist that modern times increase social isolation among older people [3,4]. Since the beginning of the COVID-19 (SARS-CoV-2) pandemic, social isolation and loneliness have received heightened attention [1-4]. Reasons for elevated concerns related to loneliness during the COVID-19 pandemic were seen because of the required social distancing (ie, because of distancing rules, citizens were not allowed to be in close physical contact with others or generally to come together). In addition, people limited their personal contacts and followed stay-at-home orders and face mask mandates in public [2-4]. Steps are needed to bridge the gap between the necessary actions for public health, individual health, and well-being. Such bridging can be done by means of web-based and print-based interventions in comparison with no support (ie, a wait-list control group [WLCG]).

Dickens et al [5] performed a systematic review of interventions targeting social isolation among older adults. They found that 86% of the interventions aimed at supporting activities (social activities and PAs) were effective [5]. Specifically, these activities were comprised of group and psychosocial accomplishments and included besides exercise and PA also arts, discussion rounds, therapeutic writing, group therapy, reading to children, lectures, assistance with organizing social behavior, outings, mutual help networks as well as leisure and cultural events with different durations [5]. Activities, especially social components and PA, are key to preventing or overcoming social isolation and loneliness [4,6]. However, only 25% of internet training interventions have revealed a successful reduction in social isolation among older adults [5].

In conclusion, the evidence demonstrates that interventions fostering physical exercise and PA can improve mood; increase physical, social, and cognitive activities; and decrease social isolation [7]. However, little is known about the delivery mode of the intervention (eg, the comparison of internet-based training interventions and traditional print-based interventions), as well as the mechanisms that may explain possible differential effects. Therefore, this study addressed this open question. The research question was whether more adults reported changing their PA behavior if they received the web-based or print-based intervention than those who received the control condition.

A previous study by Boekhout et al [8] revealed the benefits of a printed delivery method compared with a web-based version. Specifically, the authors found higher participation and lower attrition rates in this group [8]. Golsteijn et al [9] compared printed materials with web-based materials in terms of cost-effectiveness and cost-utility to promote PA among adults aged ≥50 years. The results revealed that the print-based material was most cost-effective in terms of increasing PA and could also contribute to better overall health at the population level [9]. However, little is known about the effectiveness in terms of loneliness and social-cognitive predictors, what actually explains the effects of the intervention, and in whom and how the intervention works [10]. Depression can be an important component of mental health [1], whereas a central factor of well-being and successful aging is feeling subjectively fit [7]. Thus, this study investigated whether a web-based or print-based PA intervention improved outcomes such as social-cognitive predictors of PA behavior change, subjective age, feelings of loneliness, and symptoms of depression in comparison with a control group. Furthermore, we examined whether mediating and moderating mechanisms exist. Conceivable mechanisms will be outlined in the following sections to set the stage for this study.

Potential of PA Interventions

PA is imperative for health and well-being at any age and is increasing in importance with older age [11,12]. Approximately half of the population will be aged >60 years by 2030 [13]. Consequently, it is important to improve the health of this population. Regular PA, particularly cardiovascular training (also called endurance training), is considered to have enormous potential for maintaining and improving the health and well-being of older adults [14,15]. Following the World Health Organization (WHO) recommendations, PA should be conducted for ≥150 minutes per week with moderate to vigorous PA (MVPA) in bouts of at least 10 minutes to improve and maintain health [13].

Cross-sectionally, more PA is related to better health and vice versa, and the improvement of a healthy lifestyle has been demonstrated to pay off in terms of increasing or recovering health [16,17]. Olson and McAuley [18] demonstrated the effectiveness of an intervention, including walking exercise (endurance training) and theory-based group workshops, aimed

XSL•FO
RenderX

at improving the PA level of older adults in the short and long term.

An important limitation of PA programs for older adults is that they are often unattractive to older adults [19]. This might be overcome by addressing the individual-level characteristics and (technology-based) preferences of participants [20]. Concerning the overall uptake of PA and sports offerings, demographic and socioeconomic individual-level differences have been shown to be relevant in past studies [21-23]. Those who are already actively involved in PA and sports are more likely to participate [23-25]. However, some people may also experience that being active operates as a barrier to adopting activities such as a new physical exercise program [26], showing that previous behavior is yet another predictor that depends on individual circumstances. A study of older adults' specific requirements for PA class meetings also revealed sex-specific differences. For example, men, in contrast to women, were more critical of group activities [27]. Further identification of how and why interventions work can help the development and organization of attractive future health interventions [19,20]. A theoretical framework that might explain the differences based on baseline characteristics such as previous experience is described in the following section.

Theory-Based Interventions and Social-Cognitive Predictors of PA Behavior Change

Research comparing the effectiveness of theory-based and non–theory-based health behavior change interventions has demonstrated a higher potential for theory-based approaches to effectively promote PA [28-30], although not consistently [31]. However, overall, it should be noted that health behavior change interventions to improve PA are very heterogeneous with regard to theoretical approaches, designs, and effectiveness. In addition, some interventions have only been found to produce small to moderate effects [32-34]. For example, an aggregated effect of Cohen *d*=0.27 was determined by Rhodes et al [34] in a high-level overview of published reviews of the literature, which has been interpreted as small but meaningful. This shows that theoretical frameworks that take further relevant parameters and pathways into account are needed for the design of PA interventions.

Social-cognitive variables are imperative for predicting active behavior change [35]. Knowledge of such variables enables the design of interventions. For example, key social-cognitive variables are described in the Health Action Process Approach (HAPA) [35]. The HAPA is a theory that organizes different social-cognitive variables into a meaningful structure [35]. The HAPA has two layers: a continuum layer with social-cognitive variables and a stage layer with the stages of change. The HAPA assumes three different stages of change: the nonintenders stage with its processes that lead to a behavioral intention, the intenders stage with postintentional volition processes that lead to the actual health behavior, and the action stage where the goal behavior is performed.

Within different stages, different patterns of social-cognitive predictors may emerge. In the nonintenders stage, a person develops the intention to act. In this phase, risk perception is seen as a distal antecedent (eg, "I am at risk for cardiovascular disease"). Risk perception in itself is insufficient to enable a person to form an intention. Rather, it sets the stage for a contemplation process and further elaboration of thoughts on consequences and competencies. Similarly, positive outcome expectancies (eg, "If I exercise five times per week, I will reduce my cardiovascular risk") are chiefly seen as important in the motivation phase when a person balances the pros and cons of certain behavioral outcomes. Furthermore, one needs to believe in one's capability to perform the desired action (perceived self-efficacy; eg, "I am capable of adhering to my exercise schedule despite the temptation to watch TV"). Perceived self-efficacy operates in concert with positive outcome expectancies, both of which contribute substantially to the formation of intention. Both beliefs are needed to form intentions to adopt difficult behaviors, such as regular physical exercise.

After forming an intention, the volitional phase is entered. Once a person is inclined to adopt a particular health behavior, the *good intention* must be transformed into detailed instructions on how to perform the desired action. As soon as an action is initiated, it must be maintained. This is not achieved through a single act of will but involves self-regulatory skills and strategies. Thus, the postintentional phase should be further broken down into more proximal factors, such as planning, action control, social support, and recovery self-efficacy.

Social support is a factor that reflects the barriers and resources part of the HAPA model: support represents a resource, and the lack of it can be a barrier to adopting or maintaining health behaviors. Instrumental, emotional, and informational social support can enable the adoption and continuation of behaviors [35]. The theoretical assumptions not only improve the prediction of behavior but also allow for designing of interventions more effectively by tailoring the intervention components to the needs of the recipient and, finally, enhance participation. The relevant factors are described in the following sections.

Tailored Web-Based Intervention

Tailoring is a key aspect of making interventions more effective, not only by considering the users' stage of change but also by matching the users' needs. For instance, such needs can be that participants prefer self-monitoring and activity tracking as components of their intervention (eg, by digital formats, as found by Powell et al [36]).

Digital modes have much more potential than paper-based intervention modes as they provide more options for personalization. At the same time, information can be delivered in different forms, including textual, visual, and audiovisual information to suit individual preferences and abilities [37]. However, when preferences are considered, older people in particular like print formats better [38] and accordingly might benefit more from it. However, this requires more systematic research.

The tailoring of interventions is a method that aims to meet the needs of all individuals more appropriately. However, meeting all these different needs is typically challenging. Therefore, it is necessary to evaluate whether all individuals benefit equally.

XSL•FO
RenderX

For instance, in a previous study [39], it was found that participants who were not sufficiently physically active before the study (nonintenders and intenders in comparison with actors) found the intervention useful. In another study [40], the printed method was more effective than the internet method in participants with a high baseline intention for PA (intenders). Thus, the question remains of whether the intervention is moderated by the stage of change in endurance activities. Other needs may be interrelated with sociodemographic characteristics such as age, which necessitates further elaboration, as will be described in the following sections.

Aging, Loneliness, and Subjective Age

Aging is typically stereotyped as feeling lonely. However, loneliness is not related to older age but the opposite—younger cohorts feel more lonely than older cohorts [1]. The aging population is at higher risk of other health-related challenges [1,3,4]. Aging processes and the health of older adults are highly important. Many older adults experience more health limitations and an increased burden, such as falling upon their caregivers [41]. In addition, older adults might have the highest risk of inactive lifestyles because of their reduced functioning [42].

Aging theories posit that older adults prefer to exercise with other individuals instead of exercising alone [43]. Accordingly, blended web-based and print interventions for older individuals promoting PA proved to be effective [44-46] as web-based materials would typically be used more for unaccompanied modes. However, whether print and web-based materials are beneficial for older adults' health (eg, symptoms of depression), well-being (or the opposite, eg, loneliness), health behavior, and its predictors requires further investigation.

Typically, calendrical and subjective ages are distinguished [47,48]. Calendrical age is determined by the date of birth [38]. In contrast, if a person is asked how old they feel, then the perceived or subjective age can be determined [47]. The latter is associated with health status and well-being, as well as with behavioral, cognitive, and biological processes, including frailty [47].

Although calendrical age cannot be changed, subjective age contains many options for interventions: individuals who feel younger are better off and more optimistic [48]. Previous studies have demonstrated that interventions can improve subjective age and general health status or even reverse frailty [47]. The question remains whether the effect of a PA intervention is mediated by subjective age; thus, this study investigated this in further detail.

Interrelations and Stage of Change

Loneliness and mental health issues, such as depression, are interrelated [1]. Fortuna et al [10] summarized that older adults can benefit from digital services to overcome their mental health limitations (such as loneliness and depression). Moreover, PA can help reduce depression and loneliness using mastery experience and self-regulation with regard to physical perceptions and repairing interpersonal skills and relationships [6]. Accordingly, the PA intervention group (IG) allocation should reduce the likelihood of depressive symptoms and loneliness.

However, much is still unresearched, such as whether internet-based services are as good as, or better than, traditional services for older individuals' mental health. Although the advantages are obvious, the effects on outcomes such as loneliness and symptoms of depression still need more systematic attention, which will be addressed in this study.

With respect to intervention studies, it is assumed that the assignment of participants to specific study arms with different forms of content would have an effect on symptoms of depression and that this effect is mediated by subjective age and moderated by the stage of change in endurance training (Figure 1).

Furthermore, the question remains as to whether the same relationship with loneliness as a dependent variable would be feasible (Figure 2). Accordingly, research is required regarding the key question of whether the intervention effects on loneliness and symptoms of depression depend on subjective age and whether the stage of change for endurance activities affects this effect. Thus, the hypotheses described in the following sections were investigated.

XSL•FO
RenderX

Figure 1. Proposed moderated mediation model for depression. CES-D: Center for Epidemiologic Studies Depression; T0: time point 0; T1: time point 1.

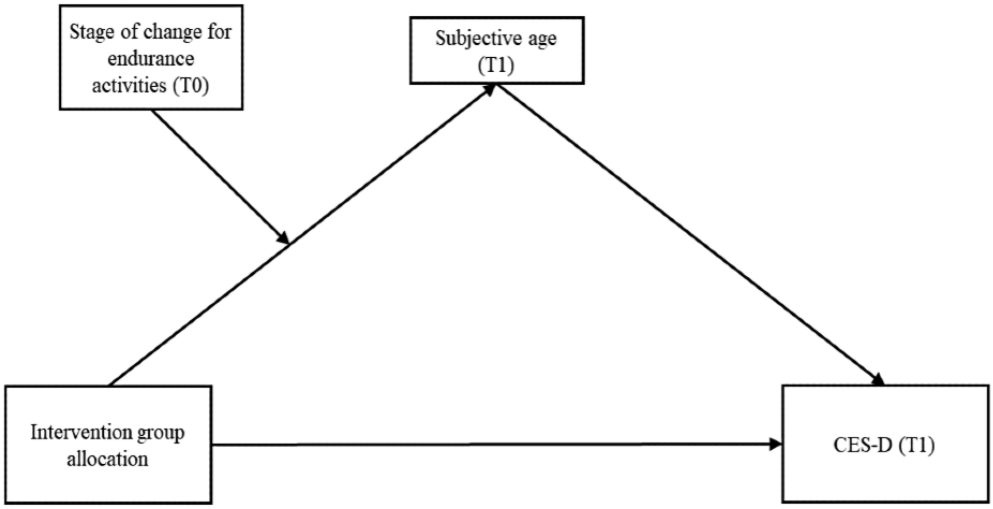
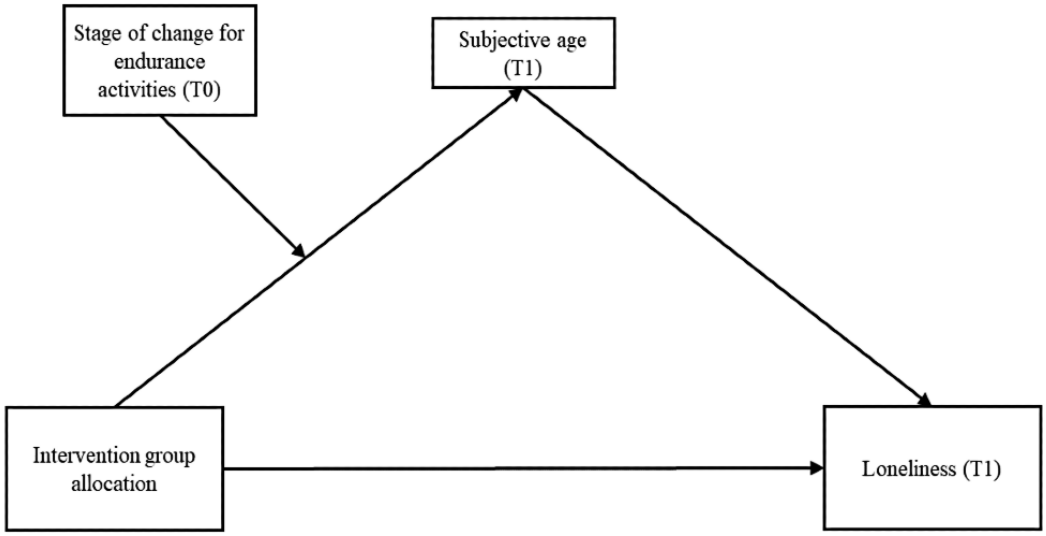


Figure 2. Proposed moderated mediation model for loneliness. T0: time point 0; T1: time point 1.



Goals of This Study

The goal of this study was to test the following research questions: do interventions delivered in a web or print mode improve self-reported PA stage of change, social-cognitive determinants of PA, feelings of loneliness, and symptoms of depression, and in this effect, does subjective age act as a mediator and stage of change act as a moderator?

The following hypotheses were tested:

1. The proportion of older adults who self-report a PA behavior change is higher in the web-based and print-based PA interventions than in the respective control conditions.
2. Compared with the control condition, the web-based and print-based PA interventions improve social-cognitive predictors of PA behavior change, subjective age, feelings of loneliness, and symptoms of depression.
3. The intervention's effect on feelings of loneliness is mediated by subjective age, and this mediation is moderated by the stage of change for endurance activities (moderated mediation).
4. The intervention effect on symptoms of depression (Center for Epidemiologic Studies Depression [CES-D] score) is mediated by subjective age, and this mediation is moderated by the stage of change in endurance activities (moderated mediation).

Methods

Overview

The PROMOTE study comprised 2 cohorts: PROMOTE 1 and 2. In this study, web- and print-based programs to promote PA in community-dwelling older adults were developed, analyzed, and evaluated according to multiple theoretical models and intervention effects using randomized intervention trials [49-54]. These were conducted as part of the interdisciplinary Physical Activity and Health Equity: Primary Prevention for Healthy Aging research network [55].

The first trial (2015-2018, PROMOTE 1) compared the effects of 2 web-based interventions with a wait-listed control condition, whereas the second trial (2018-2021, PROMOTE 2) compared adapted versions of the web-based interventions (the program

XSL•FO
RenderX

was adapted to initially inactive older adults) with a print-based PA intervention. For the analyses in this study, the groups receiving the web-based intervention in PROMOTE 1 and 2 were combined. The plan for pooling the data of the 2 intervention studies is described in the study protocol of PROMOTE 2 [51]. Accordingly, measures with the intervention design were taken to synchronize the different interventions from the beginning of PROMOTE 2: interventions did not significantly differ in their content and with the levels of recommended activity levels [51]. Measures were matched for PROMOTE 1 and 2 to pool data from both trial periods for the overarching analyses.

Ethics Approval

PROMOTE 1 was approved by the Ethics Committee of the Technical University of Chemnitz, Faculty of Behavioral and Social Sciences, on July 14, 2015, with the ethics approval number: V-099-17-HS-CVR-PROMOTE-03072015. Ethics approval for PROMOTE 2 was obtained from the Medical Association in Bremen on July 3, 2018, with the ethics approval number 635. The trials were conducted in accordance with the ethical principles of the American Psychological Association and the 1964 Declaration of Helsinki and its later amendments of comparable ethical standards. All participants were fully informed about the study and provided informed consent.

Recruitment

Detailed information regarding data collection (recruitment and randomization strategies) can be obtained from the studies by Muellmann et al [49] and Pischke et al [51]. Briefly, in 2016 (PROMOTE 1) and 2018 (PROMOTE 2), random samples of n=8299 older adults aged 65 to 75 years and n=3492 older adults aged ≥60 years and living independently (without assisted living) were selected by the residents' registration offices from municipalities in the Bremen metropolitan region and invited to participate by mail. In addition, both study phases were promoted in the local press, as well as via prior discussions with the research staff with an offer to enroll voluntarily. Eligibility for study participation, which is described in detail in the published study protocols [49,51], was determined through computer-assisted telephone interviews with trained study nurses.

The main inclusion criteria were being aged 65 to 75 years in PROMOTE 1 and ≥60 years in PROMOTE 2, as well as living independently, having basic knowledge of German, being able to walk without a walking aid, participation in study assessments and weekly group meetings without external assistance, and providing informed consent.

The exclusion criteria were as follows: a medical condition or diagnosis prohibiting PA, severe visual or other impairments, implanted cardiac devices, or occasional syncopal episodes leading to exclusion or cognitive impairment (Mini-Mental State Examination [MMSE] <25 in PROMOTE 1 and MMSE second edition: MMSE-2: Brief Version [MMSE-2:BV] score <13 in PROMOTE 2). Individuals were excluded from the study if they were planning a vacation during the intervention period, had certain medical conditions or severe health impairments, or did not have a mobile device or internet access. As the results of the first study indicated that predominantly already active individuals participated in the study, the exclusion criteria for PROMOTE 2 were modified, and individuals were excluded if they reported being regularly physically active for at least 2.5 hours per week for >1 year before the start of the study. Potential study participants for PROMOTE 2 were excluded if they had already participated in PROMOTE 1 [52].

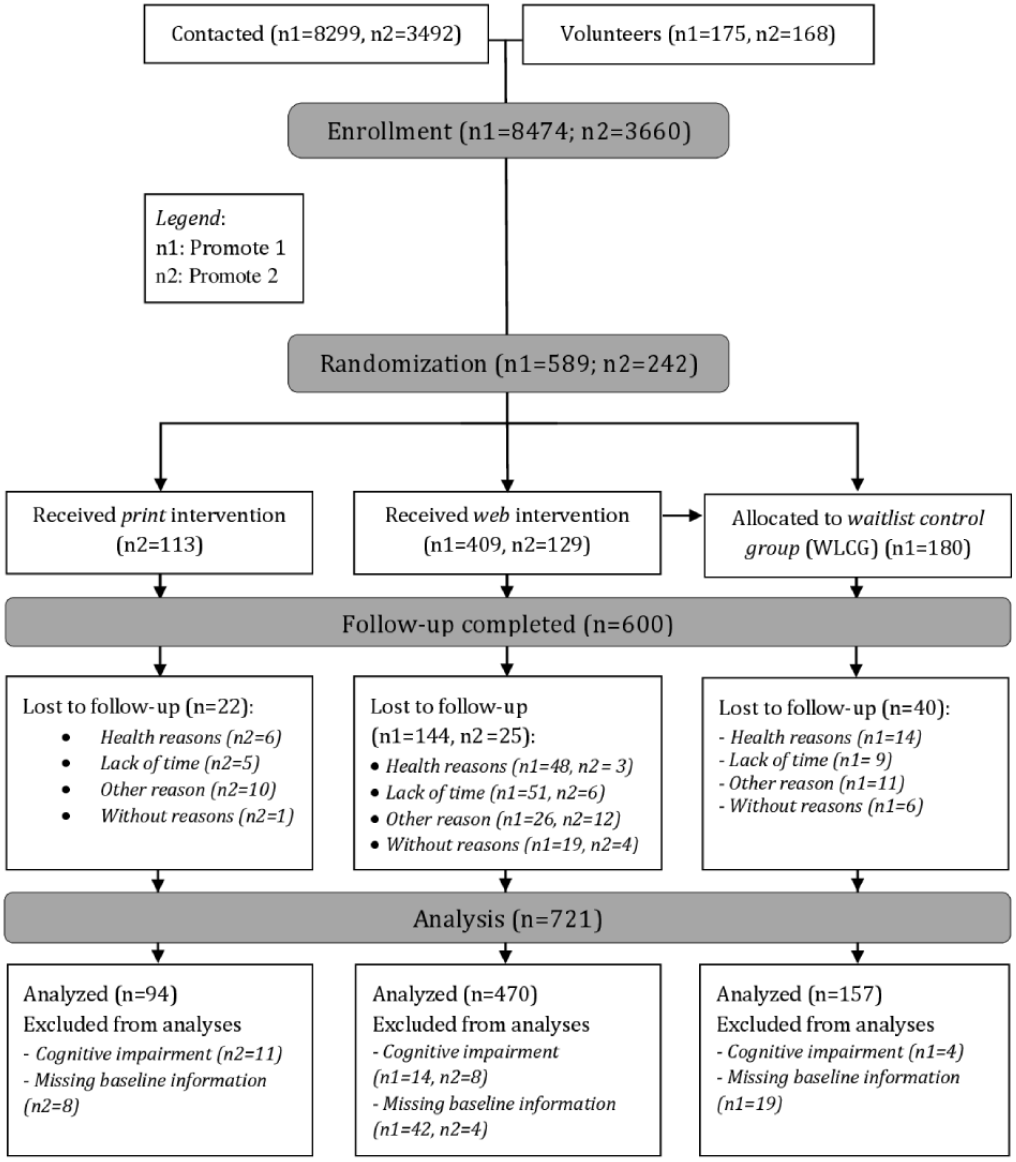
Finally, participants were randomly assigned to the following study arms:

1. In PROMOTE 1 (N=589), to either a web-based intervention with subjective PA self-monitoring (211/589, 35.8%), web-based intervention with subjective and objective PA self-monitoring (198/589, 33.6%), or WLCG (180/589, 30.6%) [49]
2. In PROMOTE 2 (N=242), to a print-based intervention with subjective PA self-monitoring via a printed PA pyramid (113/242, 46.7%) and web-based intervention with subjective PA self-monitoring via a web-based PA pyramid (129/242, 53.3%); approximately 29.5% (38/129) of the latter were randomly selected and received a PA tracker (objective PA self-monitoring) in addition [51]

In total, 831 individuals were randomized into one of these three groups: web-based PA intervention, print-based PA intervention, or WLCG. Further details are outlined in the flow chart in Figure 3.

XSL•FO
RenderX

Figure 3. Study flow chart.



Procedure and Interventions

The *PA interventions* in this study were developed on the basis of the self-regulation theory [56,57]; behavior change techniques (ie, goal setting, planning, social support, and feedback) [58]; and (in PROMOTE 2) in a cocreative process with individuals of the target group and stakeholders, such as exercise instructors, leaders of older adults' facilities, and managers of older adults' homes. The intervention aimed to improve overall self-monitoring capabilities regarding PA and enable transfer for the time after the intervention.

According to the PA recommendations of the WHO and the American College of Sports Medicine [13], older adults randomized to the IGs were advised to engage in physical exercises. The recommendations included suggestions to improve balance (twice per week), flexibility (twice per week), and strength (twice per week for the 8 major muscle groups). In addition, the participants were instructed to engage in at least 150 minutes per week of moderate endurance training or 75 minutes per week of vigorous training or a combination of both moderate and vigorous training intensities [59].

After randomization, a baseline assessment (time point 0 [T0]) was conducted. Following T0, the IGs in PROMOTE 1 received a *print-based intervention* in the form of a booklet, which was tailored to the individual baseline PA levels. Feedback was tailored to the baseline motivational stage (nonintention, intention, or action) to engage in the recommended PA. In addition, the material was tailored based on sex: pictures of men for male participants and pictures of women for female participants were used to model the recommended exercises. The web-based materials offered in the corresponding study arms in PROMOTE 1 and 2 included access to web-based materials that contained and displayed the same information on exercises for balance, flexibility, endurance, and strength, as for the print-based version. In addition, for PROMOTE 2, a print-based PA diary was developed in the form of an expert-driven approach and contained all exercises; provided the option to enter data on performed exercises; and, thus, visualize personal progress. The web-based intervention from

XSL•FO
RenderX

PROMOTE 1 was adapted based on feedback obtained during the first study period. An additional web-based option (PA tracker) to capture the daily step count was included for the randomly selected subgroups in both study arms (PROMOTE 1 and 2).

All IGs were encouraged to use the material and engage in the recommended PA over 10 weeks. These were accompanied by weekly group meetings conducted and moderated by trained staff members. During these meetings, questions concerning the program could be raised. In addition, theoretical inputs for healthy aging were provided. Moreover, physical exercises were performed together with feedback from the participants regarding their exercise practice. At the same time, social interactions among the participants and their contacts with the study team were facilitated. After 10 weeks of group meetings, a follow-up assessment for time point 1 (T1) was conducted 12 weeks after the baseline assessment. Several collected variables served as the basis for this study. More information on the interventions and procedures can be found in previous publications [49-51,53].

Used Instruments

Adherence was measured according to the WHO recommendations of ≥150 minutes per week of MVPA in bouts of at least 10 minutes. The daily minutes for PA in terms of MVPA were assessed by asking the study participants what activities they performed in bouts of at least 10 minutes. Minutes per week for MVPA in the bouts were derived by multiplying the daily average minutes in the 10-minute bouts by 7. Subsequently, this measure was dichotomized as meeting or not meeting the WHO recommendations. This resulted in a dichotomous variable with 1=does not meet 150 minutes of MVPA recommendation and 2=meets 150 minutes of MVPA at T0 and T1. To determine adherence over time, a categorical variable regarding the change in meeting the recommended 150 minutes of MVPA was computed by subtracting the baseline value from the 12-week follow-up value. The resulting variable indicated whether study participants *remained active or inactive* (0), fell back into *not meeting the recommendation anymore* (−1), or *became active* (1).

To assess the *stage of PA behavior*, participants were asked whether they had performed ≥150 minutes of endurance exercise per week (eg, fast walking, walking, biking, and swimming) at 2 measurement time points (T0 and T1). Participants were asked to respond on a 5-point Likert scale: 1=“No, and I do not intend to start” (precontemplation stage or nonintenders); 2=“No, but I am considering it” (contemplation stage or nonintenders); 3=“No, but I seriously intend to start” (preparation stage or intenders); 4=“Yes, but only for a brief period of time” (action stage or actors); and 5=“Yes, and for a long period of time” (maintenance stage or actors).

The stage item is based on items used by Lippke et al [60]. For this study, participants in the precontemplation and contemplation stages were categorized as nonintenders, participants in the preparation stage were categorized as intenders, and participants in the action and maintenance stage were categorized as actors.

Intention to engage in regular endurance and strength training was assessed with 2 items (“I intend to engage in strenuous endurance training for at least 75 minutes per week and strength- and balance training twice a week” and “I intend to engage in moderate endurance training for at least 150 minutes per week and strength- and balance training twice a week”).

These items were based on previous literature [60,61]. Both items were measured at T0 and T1 on a 7-point Likert scale ranging from 1=*completely disagree* to 7=*completely agree*. The 2 items were kept separate as they had discriminant validity (Spearman ρ at T0=0.410; Spearman ρ at T1=0.467; P<.001). The retest reliability was Spearman ρ T0 to T1 of 0.531 and 0.378, respectively (P<.001).

Outcome expectancies, as suggested by Lippke et al [62] and Schwarzer et al [63], were measured using 4 items in total at 2 measurement time points. A total of 2 items measured positive outcome expectancies (“If I engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise of strength and balance training twice per week, it is good for my health.” and “If I engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise of strength and balance training twice per week, it makes me feel better afterwards.”).

The remaining 2 items focused on negative outcome expectancies (“If I engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise of strength and balance training twice per week, it takes too long.” “If I engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise of strength and balance training twice per week, it is too costly.”). All 4 items were measured on a 7-point Likert scale ranging from 1=*completely disagree* to 7=*completely agree*.

The 2 items measuring positive outcome expectancies were kept separate as they had rather discriminant validity (Spearman ρ at T0=0.693; Spearman ρ at T1=0.703; P<.001). The retest reliability was Spearman ρ T0 to T1 of 0.425 and 0.508, respectively (P<.001). The 2 items measuring negative outcome expectancies were also kept separate as they had rather discriminant validity (Spearman ρ at T0=0.474; Spearman ρ at T1=0.443; P<.001). The retest reliability was Spearman ρ T0 to T1 of 0.339 and 0.441, respectively (P<.001).

Self-efficacy was measured with 5 items, in total, at both measurement time points T0 and T1 [61,64]. A single item was used to assess task self-efficacy (“I am confident that I can engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise and strength and balance training twice a week, even if it gets difficult.”).

A total of 2 items assessed maintenance self-efficacy (“I am confident that I can engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise and strength and balance training twice a week, even if it takes long, until it is a habit.” and “I am confident that I can engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise and strength and balance training twice a week, even if I am worried or face problems, e.g., scheduling difficulties.”).

XSL•FO
RenderX

In addition, recovery self-efficacy was measured by 2 items ("I am confident that I can engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise and strength and balance training twice a week, even if I postponed my plans several times." and "I am confident that I can engage in 150 minutes of moderately strenuous or 75 minutes of strenuous endurance exercise and strength and balance training twice a week, even If I suspended several times.").

All 5 items were measured on a 7-point Likert scale ranging from 1=*completely disagree* to 7=*completely agree*. For this study, the sum of all 5 items was calculated (Cronbach α at T0=.884; Cronbach α at T1=.897). The retest reliability was Spearman ρ at T0 to T1 of 0.475 (*P*<.001).

To further assess social-cognitive predictors, *planning* was measured using 6 items. The items were adapted for this study from those used in previous studies on PA [65].

Three items measured *action planning*: "For the next month, I have already planned where I will be physically active," "For the next month, I have already planned how I will be physically active," and "For the next month, I have already planned when and how often I will be physically active.".

Furthermore, three items assessed the construct of *coping planning*, respectively: "For the next month, I have already planned when I have to take care not to suspend," "For the next month, I have already planned what I can do in difficult situations to stick to my intentions," and "For the next month, I have already planned how I can remain physically active even if there are barriers."

All 6 items were measured on a 7-point Likert scale ranging from 1=*completely disagree* to 7=*completely agree*. For this study, the sum of all 6 items was calculated (Cronbach α at T0=.932; Cronbach α at T1.899). The retest reliability was Spearman ρ T0 to T1 of 0.492 (*P*<.001).

Habits were measured using two items at two measurement time points [66]: "Engaging in the recommended endurance, strength and balance training is something that has become my habit." and "Engaging in the recommended endurance, strength, and balance training is something that I do without thinking about it."

Both items were measured on a 7-point Likert scale from 1 *completely disagree* to 7 *completely agree* (Spearman ρ at T0=0.474; Spearman ρ at T1=0.443; *P*<.001). The retest reliability was Spearman ρ T0 to T1 of 0.339 and 0.441, respectively (*P*<.001).

Symptoms of *depression* were measured using the CES-D [67] scale both at T0 and T1. The scale comprises 20 items with a possible sum score range of 0 to 60. Each item was measured on a 5-point Likert scale ranging from 0=*rarely or never (<1 day)* to 1=*some or little of the time (1-2 days)*, 2=*often or a moderate amount of time (3-4 days)*, and 3=*most of the time (5-7 days)*.

For the purpose of analysis, the mean score of all 20 items was calculated for all participants (Cronbach α at T0=.605, Cronbach α at T1=.587). The retest reliability was Spearman ρ T0 to T1 of 0.759 (*P*<.001).

To assess perceived *loneliness* at T0 and T1, the item *I felt lonely* was used from the CES-D scale [67]. The item was measured on a 4-point Likert scale from 0=*rarely or never (<1 day)* to 1=*some or little of the time (1-2 days)*, 2=*often or a moderate amount of time (3-4 days)*, and 3=*most of the time (5-7 days)*. The retest reliability was Spearman ρ T0 to T1 of 0.597 (*P*<.001).

Sociodemographic data were collected using a questionnaire administered before the intervention (at the baseline level). The questionnaire was formulated according to the German Health Interview and Examination Survey for Adults [68]. The collected variables included date of birth, sex of the participants (male or female), height (in cm), and weight (in kg).

In addition to the date of birth, *perceived age* was measured using an open-ended question. The participants were asked, "How old do you feel?" Perceived age was assessed at T0 and T1. The retest reliability was Spearman ρ T0 to T1 of 0.826 (*P*<.001). Furthermore, country of birth, mother tongue, family status, living alone, number of children, qualification, educational level, and employment status were assessed.

Employment status was measured with a single item taken from a questionnaire assessing demographic and sociostructural data from German older adults and adapted for this study [69]. Qualification and educational level were measured with 2 items and aggregated based on the 2016 version (volume 17) of the International Standard of Education [70].

BMI was calculated using height and weight and categorized into underweight, normal weight, overweight, and obese, according to the WHO BMI classification for adults aged ≥20 years [71].

All the used instruments were validated before and are described in the study protocols [49,51], as well as in previous publications [50,52-54].

Analysis Sample

IG allocations from PROMOTE 1 and 2 were summarized within a pooled IG variable that included the following three categories: *a*=WLCG from PROMOTE 1 (reference), *b*=web-based IG from PROMOTE 1 and 2, and *c*=print-based IG from PROMOTE 2.

Only the participants who completed the baseline assessment (T0) were included in the analysis. In PROMOTE 1, participants' cognitive status was assessed using the MMSE [72] 1 week before the start of the intervention phase. In PROMOTE 2, participants' cognitive status was assessed using MMSE-2:BV [73] during the first weekly group meeting (ie, the start of the intervention phase). Participants who scored <25 points on the MMSE or <13 points on the MMSE-2:BV were excluded from the analysis. Amendments to the cutoff values for exclusion because of cognitive impairment have been discussed in previous publications [50,52].

After excluding individuals with cognitive impairment (37/831, 4.5%; Figure 3) and missing information on baseline demographic characteristics (73/831, 8.8%), the analysis sample included 721 older adults (see *Data Exclusion* section and Figure 3). To determine adherence, a variable regarding meeting the

XSL•FO
RenderX

recommended 150 minutes of MVPA was computed by subtracting the baseline measure from the T1 measure. The resulting variable thereby indicated whether study participants *remained active or inactive* (0), fell back into *not meeting recommendations anymore* (−1), or *became active* (1).

Statistical Analysis

Preparation

All analyses were performed using SPSS (version 27; IBM Corp). The Little missing completely at random test (P>.05) suggested that data were not missing completely at random (ie, it suggested that data were missing systematically). Assuming that existing data could be used to produce an estimate of the missing information (ie, assuming that data were missing at random) [74], single data imputation was implemented by using the expectation-maximization algorithm.

Test of Hypotheses

To assess whether the 2 IGs outperformed the WLCG (hypothesis 1) on self-reported behavior and stage of change, frequency analyses and chi-square tests, Z tests, and contingency coefficient tests were used to test the number of participants who adopted or maintained an active lifestyle.

Changes in social-cognitive predictors (hypothesis 2) were analyzed with repeated-measures multivariate analysis of variance (MANOVA) via mixed-effects generalized linear models with group and time as factors. For the F values, the Roy largest root was reported.

The primary aim of the moderated mediation analyses was to investigate whether the IG allocation (independent variable or predictor) had an effect on perceived symptoms of depression (CES-D score) at T1 (dependent variable or outcome), which is mediated by subjective age at T1 and moderated by the stage of change in endurance training at T0 (hypothesis 3; see Figure 1 for the proposed model).

The secondary aim of the moderated mediation analysis was to investigate the same relationship with loneliness as the dependent variable (hypothesis 4; see Figure 2 for the proposed model). These associations were investigated using moderated mediation models within the PROCESS macro (version 3.0; Hayes, The International Association of Applied Psychology mediation analysis).

The models were adjusted for the following baseline variables: loneliness and CES-D score, as well as subjective age, sex, age, educational status (International Standard of Education), family status, and BMI (all at T0). A bootstrapping approach of 10,000 samples and a specific seed (seed=1) was applied to ensure robust and replicable results. The effect sizes were represented by unstandardized regression coefficients. To calculate the heteroscedasticity-robust SE, the HC3-Option in the process function was used. Accordingly, the assumption of homoscedasticity could be avoided.

Data Exclusion

The analyses were conducted following the intention to treat principle; that is, participants were included in primary analyses according to their original group allocation and disregarding study completion. This was managed by missing value imputation using an expectation-maximization algorithm.

In addition, according to the study protocol, participants with cognitive impairments were excluded (37/831, 4.5%; Figure 3). In addition, missing baseline demographic information was not imputed; thus, participants with missing information on sex, age, educational status, family status, or BMI were excluded from the analyses (73/831, 9%).

Results

Hypothesis 1

To test whether the 2 interventions outperformed the WLCG (hypothesis 1), the study participants who adopted or maintained an active lifestyle were analyzed, as recommended by the WHO. First, those who did *not meet the recommendation regarding PA at T0* based on self-reported adherence were investigated; of those individuals, more individuals became adherent if they were exposed to the web-based intervention (47/59, 80%; Table 1) or received the print-based intervention (20/25, 80%) than those who were not treated (WLCG; 5/7, 71%). At a descriptive level, the numbers indicate the favoring of the IG over the control group.

Second, those who *met the recommendation regarding PA before the study* were investigated; of these, 600 individuals self-reported to be adherent at T1. More individuals remained adherent if they were exposed to the web-based intervention (396/411, 96.4%) or not treated (WLCG; 141/150, 94%) than those receiving the print-based intervention (63/69, 91%).

The difference in the proportion of adopters between the web-based and print-based interventions was statistically significant (Z=-1.94; P=.02); as well as the differences between the IGs and the WLCG were significant (WLCG vs print Z=2.3967 and WLCG vs web Z=3.8367; both P=.01).

This finding was replicated by the stages of change in endurance training. In Table 2, the number of study participants in the 3 intervention conditions moving from the nonintenders, intenders, or actor stage to another stage or remaining in the former stage is reported.

In all stage groups, the percentage of individuals moving a stage forward (from nonintentional stage to intentional or action, and from intentional to action) or maintaining the stage when starting as an actor was higher in the web-based or print group than in the WLCG (Table 2). In contrast, in the WLCG, the percentage of individuals remaining in the nonintentional or intentional stage was higher than that in the IGs (Table 2). The pattern in Table 2 was statistically significant (χ^2_4=79.1; P<.001; contingency coefficient 0.314; P<.001). Z tests were performed to test whether group differences in the stage of change movements were statistically significant, which was the case for initial intenders who moved to the actor stage (WLCG vs web-based Z=−4.2325; P=.01 and WLCG vs print-based Z=−5.349; P=.01) and initial actors who remained in the actor stage (WLCG vs print Z=−3.1853; P<.01).

Summarizing the findings regarding hypothesis 1 that the web-based and print-based interventions outperformed the

XSL•FO
RenderX

control condition in terms of PA behavior change, we can conclude that our results suggest this direction. The web-based intervention seemed to work better in terms of the prevention of remaining in or a relapse into *not meeting recommendation*s (Table 1) and adopting or remaining in the intender or actor stage (Table 2).

Table 1. Numbers and percentages of study participants assigned to 1 of 3 experimental groups regarding who met or did not meet the recommended physical activity level at T0[a] and T1[b].

At T0	At T1, n (%)		Total, N
	Not meeting the recommendation	Meeting the recommendation	
Not meeting the recommendation			
WLCG[c]	2 (28.6)	5 (71.4)	7
Web-based	12 (20.3)	47 (79.7)	59
Print-based	5 (20)	20 (80)	25
Total	19 (20.9)	72 (79.1)	91
Meeting the recommendation			
WLCG	9 (6)	141 (94)	150
Web-based	15 (3.6)	396 (96.4)	411
Print-based	6 (8.7)	63 (91.3)	69
Total	30 (4.8)	600 (95.2)	630

[a]T0: time point 0.

[b]T1: time point 1.

[c]WLCG: wait-list control group.

Table 2. Cross-tabulation of nonintenders, intenders, and actors at T0[a] moving to or remaining nonintenders, intenders, and actors at T1[b] depending on the experimental group they were in.

T0 and T1	Nonintenders, n (%)	Intenders, n (%)	Actors, n (%)	Total, N
Nonintenders				
WLCG[c]	36 (59)	16 (26.2)	9 (14.8)	61
Web-based	35 (16.7)	46 (22)	128 (61.2)	209
Print-based	9 (17.6)	13 (25.5)	29 (56.9)	51
Intenders				
WLCG	7 (23.3)	11 (36.7)	12 (40)	30
Web-based	5 (4.8)	11 (10.6)	88 (84.6)	104
Print-based	4 (16)	2 (8)	19 (76)	25
Actors				
WLCG	16 (24.2)	8 (12.1)	42 (63.6)	66
Web-based	9 (5.7)	7 (4.5)	141 (89.8)	157
Print-based	2 (11.1)	0 (0)	16 (88.9)	18
All stages together				
WLCG	59 (37.6)	35 (22.3)	63 (40.1)	157
Web-based	49 (10.4)	64 (13.6)	357 (76)	470
Print-based	15 (16)	15 (16)	64 (68.1)	94
Total	123 (17.1)	114 (15.8)	484 (67.1)	721

[a]T0: time point 0.

[b]T1: time point 1.

[c]WLCG: wait-list control group.

XSL•FO
RenderX

Hypothesis 2

A total of 2 MANOVAs with 19 predictors (Multimedia Appendix 1, Table S1) and 3 outcomes (Multimedia Appendix 1, Table S2) were calculated to test hypothesis 2.

The first MANOVA revealed a significant difference between the 3 groups, favoring the 2 IGs over and above the WLCG ($F_{19,\ 701}$=4.778; P<.001; η^2=0.098), as well as a significant interaction between time and group ($F_{19,\ 701}$=2.778; P<.001; η^2=0.070).

This effect was mainly based on intention, negative outcome expectancies, planning, and habit (see Multimedia Appendix 1, Table S1, for the means, SDs, and statistics). Figure 4 outlines the development, indicating that the WLCG dropped slightly with its intention over time, whereas the 2 IGs improved over time. Figure 4 shows that the WLCG retained its negative outcome expectancies, whereas the 2 IGs improved in terms of perceiving fewer negative outcomes. Figure 4 also demonstrates that the WLCG dropped with its self-efficacy, whereas self-efficacy remained stable in the web-based IG and increased in the print-based IG. Finally, Figure 4 shows that the WLCG remained stable in terms of habit strength, whereas the 2 IGs improved over time.

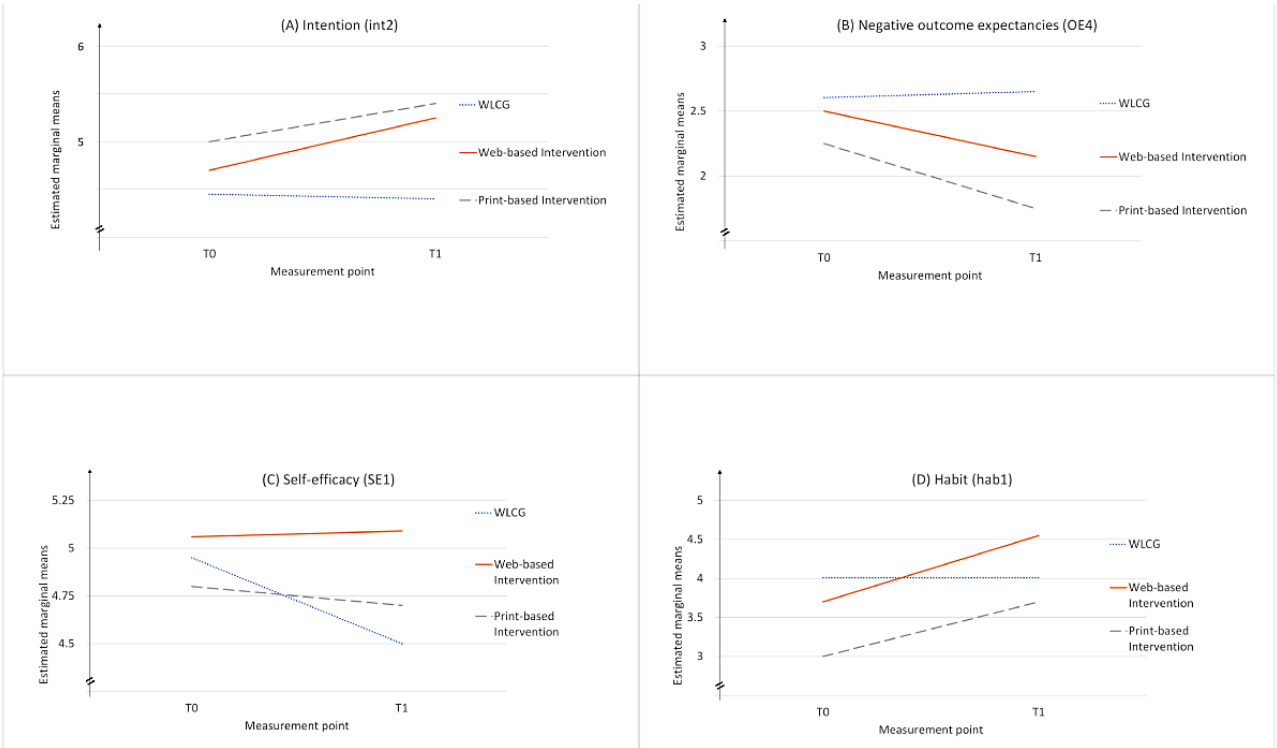
With the second MANOVA testing the outcomes, the effects of the interventions on subjective aging, loneliness, and symptoms of depression were tested (Multimedia Appendix 1, Table S2). Both the between-group effect ($F_{3,\ 717}$=8.668; P<.001; η^2=0.018) and the interaction of group and time were significant ($F_{3,\ 717}$=6.101; P<.001; η^2=0.025). The group effect was mainly based on subjective age and symptoms of depression, and the interaction effect was based on symptoms of depression with regard to group and time on loneliness (see Multimedia Appendix 2, Table S1-S3 with means, SDs, and statistics).

Figure 5 highlights that the effect of *loneliness* comes from regression to the mean, with the WLCG increasing in its loneliness, the web-based IG maintaining its previous level, and the print-based IG decreasing in its loneliness over time. With *subjective age*, all groups showed an increase over time (Figure 5). Differences from baseline values remained. Figure 5 also shows that all groups started off at almost the same level of depressive symptoms. However, over time, the WLCG increased with regard to the reported symptoms of depression, whereas the 2 IGs decreased, with an even better effect of the print-based intervention than that of the web-based intervention.

Overall, the effect sizes were rather small, ranging from η^2=0.098 to 0.018.

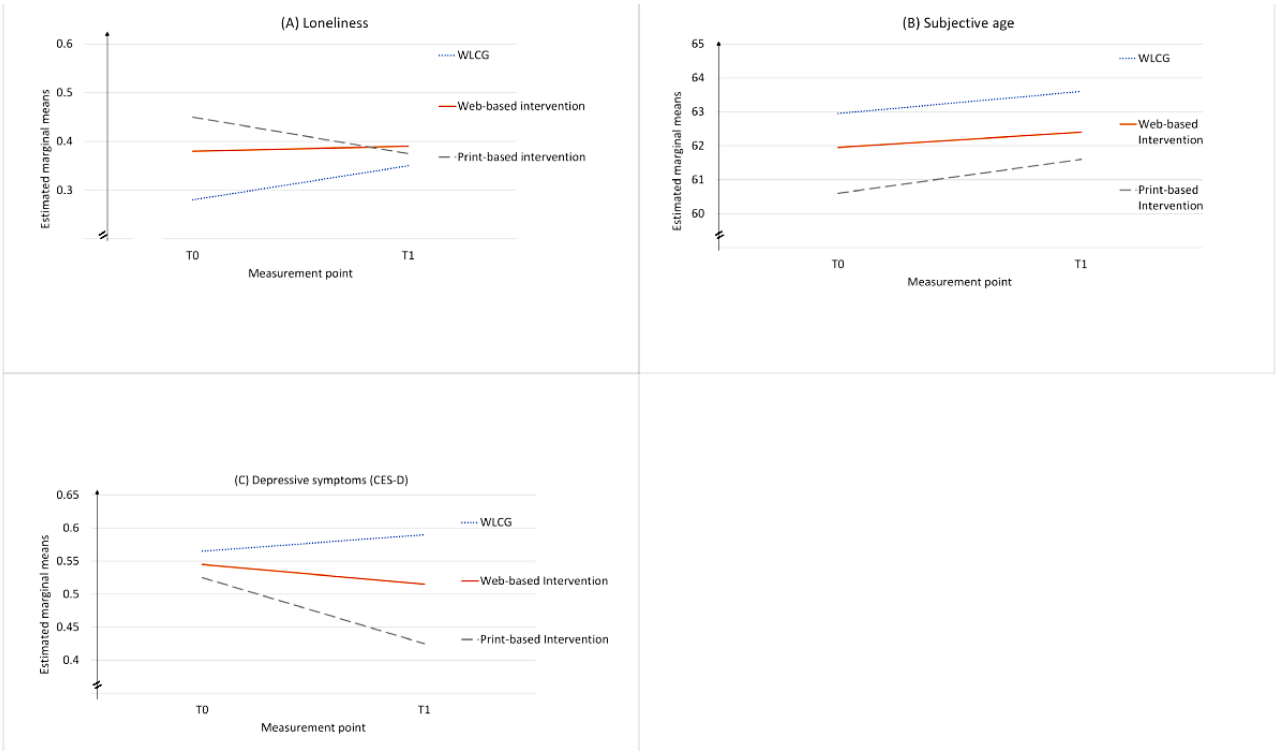
Summarizing the findings regarding hypothesis 2, we found overall support. The web-based and print-based interventions improved the social-cognitive predictors of PA behavior change, subjective aging, loneliness, and depression compared with the control condition. The web-based and print-based interventions were significantly different from the control condition.

Figure 4. (A) Intention, (B) negative outcome expectancies, (C) self-efficacy, and (D) habit. T0: time point 0; T1: time point 1; WLCG: wait-list control group.



XSL•FO
RenderX

Figure 5. (A) Loneliness, (B) subjective age, and (C) symptoms of depression. CES-D: Center for Epidemiologic Studies Depression; T0: time point 0; T1: time point 1; WLCG: wait-list control group.



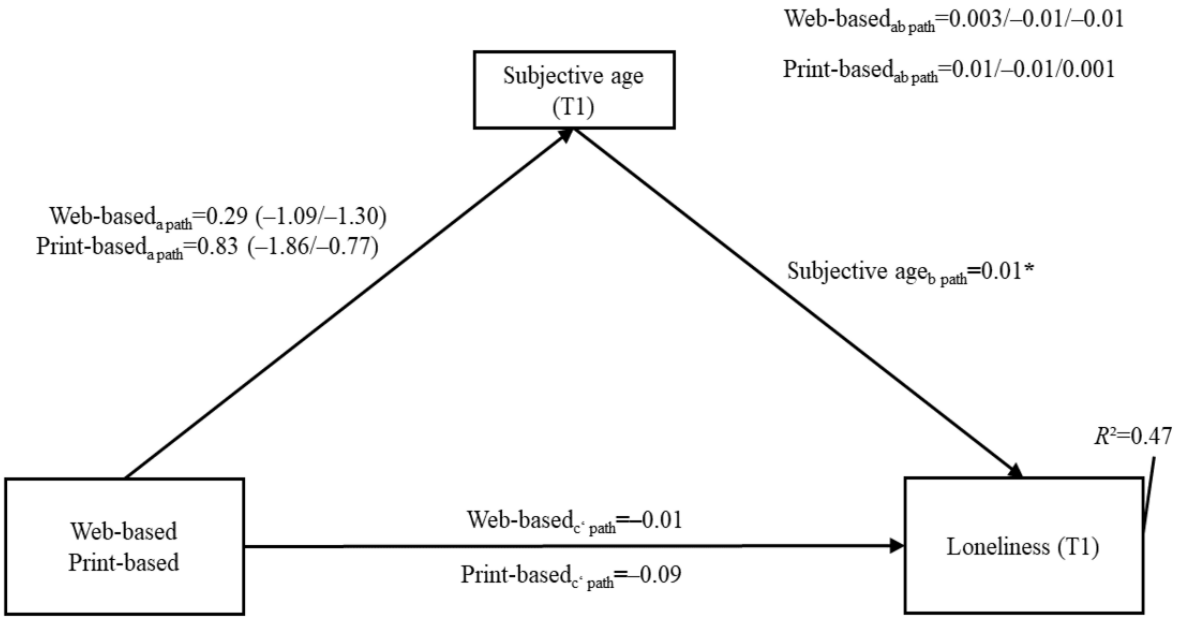
Hypothesis 3

To test the mechanisms and hypothesis 3, we conducted a moderated mediation analysis. The results are shown in Figure 6 and Tables S1 and S2 in Multimedia Appendix 2.

None of the interventions had a significant direct effect on loneliness compared with the WLCG (web-based: *c′* path −0.01, 95% CI −0.11 to 0.08, SE 0.05; print-based: *c′* path −0.09, 95% CI −0.24 to 0.06, SE 0.08). Regarding the mediator subjective age (subjective age), a significant positive relationship between subjective age and loneliness at T1 was found (subjective age: *b* path 0.01, 95% CI 0.001-0.02; SE 0.01). However, the effect size was very small, and no significant indirect effects of the interventions on loneliness through subjective age were revealed. To summarize, the results suggest that subjective age could not account for a significant proportion of the relationship between the IGs and loneliness. Furthermore, there was no significant interaction, suggesting that there was no moderation of the stage of change in endurance activities.

XSL•FO
RenderX

Figure 6. Moderated mediation model results for loneliness. The interaction between the intervention group and the stage of change in the a path is shown in parentheses. The first value represents the intention stage, and the second value represents the actor stage. The moderation of the ab path is shown on the upper right. The first value represents the nonintenders stage, the second value represents the intenders stage, and the third value represents the actor stage. The model was adjusted for the following baseline variables: loneliness, subjective age, sex, age, educational status (International Standard Classification of Education), family status, and BMI (all at time point 0). The wait-list control group was used as a reference. *Statistically significant value. T1: time point 1.



Hypothesis 4

To test hypothesis 4, another moderated mediation model was tested.

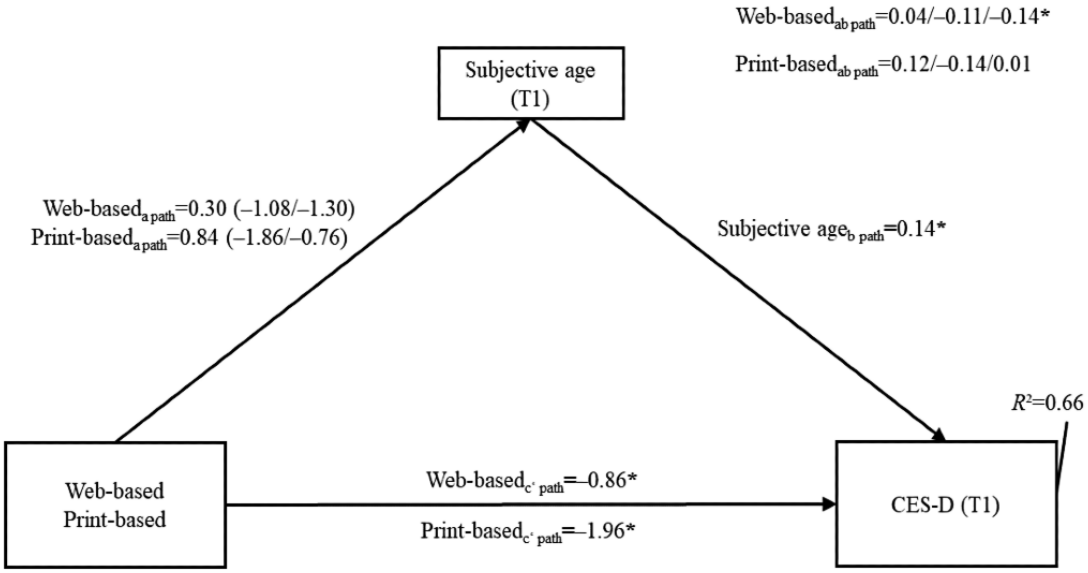
The results are shown in Figure 7 and Tables S3 and S4 in Multimedia Appendix 2.

Validating the previous analyses, both interventions had a significant direct effect on the symptoms of depression at T1 compared with the WLCG (web-based: *c′* path −0.86, 95% CI −1.58 to −0.13, SE 0.38; print-based: *c′* path −1.96, 95% CI −2.99 to −0.92, SE 0.53). Furthermore, subjective age at T1 was positively related to depressive symptoms (subjective age: *b* path 0.14, 95% CI 0.05-0.23; SE 0.05). An indirect relationship between the intervention and the symptoms of depression via subjective age was only present for participants who were both in the actor stage of change for endurance activities and received the web-based intervention (web-based: *ab* path −0.14, 95% CI −0.34 to −0.01; SE 0.09).

Summarizing the findings regarding hypothesis 4, only older adults in the actor stage of endurance training who received the web-based intervention were associated with lower symptoms of depression at T1, which was partially mediated by subjective age at T1. Thus, the existence of moderated mediation was confirmed, although all other mediation pathways were not significant, and the effect sizes were small.

XSL•FO
RenderX

Figure 7. Moderated mediation model results for depression. The interaction between the intervention group and the stage of change in the a path is shown in brackets. The first value represents the intention stage, and the second value represents the actor stage. The moderation of the ab path is shown in the upper right. The first value represents the nonintenders stage, the second value represents the intenders stage, and the third value represents the actor stage. The model was adjusted for the following baseline variable: depressive symptoms (CES-D), subjective age, sex, age, educational status (International Standard Classification of Education), family status, and BMI (all at time point 0). The wait-list control group was used as a reference. *Statistically significant values. CES-D: Center for Epidemiologic Studies Depression; T1: time point 1.



Discussion

Principal Findings

This study aimed to compare the effects of web- and print-based PA interventions on self-reported PA, stage of change, determinants of PA, loneliness, and depression. Moreover, the goal was to investigate whether subjective age is a mediator and whether the stage of change is a moderator of the effectiveness in 831 older individuals participating in the PROMOTE 1 or PROMOTE 2 study.

The main finding was that the print-based and web-based interventions both worked well and helped a higher proportion of individuals meet the recommendations for PA and move forward with their stage of change compared with the WLCG. Support for the effectiveness of the interventions was also found regarding the social-cognitive predictors of PA behavior. None of the interventions had a significant direct effect on loneliness compared with the control group. Thus, the main assumption that a PA intervention always helps reduce loneliness does not hold true. It seems more important to take mastery experience into account: the results of our moderated mediation analyses suggest that compared with the WLCG, receiving the web-based intervention was associated with lower symptoms of depression at T1 and that subjective age could explain a substantial proportion of variance. However, this holds true only for participants in the actor stage of change for endurance activities. The mechanisms are in accordance with the assumption that mastery experience and self-regulation—operationalized with subjective age—help the study participants who are already physically active at the baseline to reduce their symptoms because of maintained or improved physical perceptions and repairing or maintaining interpersonal skills and relationships [6]. In this group, the PA IG allocation reduced the likelihood of depressive symptoms and loneliness if they actually felt younger.

However, compared with the wait-list control, the intervention did not help older adults feel less lonely, perhaps because of 2 aspects. We revealed a floor effect (ie, generally low loneliness levels). Moreover, one should also keep in mind that the intervention was not designed to reduce loneliness but to increase PA; accordingly, the relationship between the intervention and feeling lonely was not strong enough to be of statistical importance. Nevertheless, the effects in the IGs underlined the importance of supporting active older adults to remain physically active to feel fit and subjectively young, as well as to lower symptoms of depression. In the following paragraphs, the hypotheses are reviewed in more detail, followed by more discussion.

Hypothesis 1, assuming that web-based and print-based interventions outperform the control condition in terms of self-reported PA behavior change, was confirmed by our data. The interventions seemed to work better in terms of preventing a relapse into *not meeting recommendations anymore* and moving study participants into the actor stage. Without any intervention, 71% (5/7) of the previously inactive participants became active at the recommended level. This percentage was higher in the web-based (47/59, 80%) and print-based groups (20/25, 80%). Without any intervention, 94% (141/150) of the previously active study participants remained active at the recommended level. This percentage was slightly higher in the web-based group (396/411, 94%).

When replicating this finding with the stages of behavior change, the effects clearly demonstrated the benefits of the interventions; in the untreated WLCG, individuals were more likely to remain in or relapse to the nonintentional stage than the individuals in the 2 IGs.

XSL•FO
RenderX

Our results also support hypothesis 2: participation in web-based and print-based PA interventions was associated with improvements in the social-cognitive predictors of PA behavior change, self-reported PA behavior change, subjective age, loneliness, and depressive symptoms compared with the WLCG. The web-based and print-based interventions were significantly different from the WLCG, which matched expectations.

Hypothesis 3 was not supported by our results, as we found that the results were not consistent with the assumption that individuals in the IG would benefit from fewer feelings of loneliness as they would experience a decrease in subjective age. Regarding hypothesis 4, only older adults who were in the actor stage regarding endurance training and who received the web-based intervention revealed an intervention effect on depressive symptoms (CES-D score), partially mediated by subjective age at T1. This subgroup reported a lower subjective age than that of the WLCG, which was further associated with reporting lower symptoms of depression at T1.

Limitations and Suggestions for Future Work

In this study, only subjective data were included as most of the variables such as social-cognitive variables, stage of change, loneliness, subjective age, and symptoms of depression could only be measured in this form. However, the behavior should be assessed using objective measures (such as an activity tracker). Accordingly, validation studies are required to better understand behavior changes more thoroughly.

In addition, the included study participants were not representative samples, as many individuals had already met the behavioral criteria, especially in the first study phase (PROMOTE 1). Despite the adaptation of the inclusion criteria for the second phase of the study, the target group of individuals with low activity was still not sufficiently reached, and their needs were not adequately addressed. This is a common problem in public health intervention research, which requires further effort (eg, to improve the recruitment of nonactive and low-motivation individuals). Therefore, future studies should attempt to overcome this problem. Therefore, to address the needs of these individuals more effectively, need-based assessments (ie, tailored to current circumstances and hindering factors) should be conducted in future studies.

Another limitation is access to, and availability of, web-based technology, particularly among the studied age groups. Thus, in this context, on the one hand, some of the respondents interested in the study invitation had to be excluded from participation as they did not meet the inclusion criterion of owning a PC or did not have access to the internet.

On the other hand, personal affinity for and acceptance of modern technology services, which in part becomes apparent over the course of the study, might affect the success of information technology interventions and, therefore, should not be overlooked. These eligibility criteria partly exclude disadvantaged groups with different psychological preconditions and developmental possibilities. In the future, such limitations should be overcome by making study participation more accessible to individuals with low computer or internet literacy and technology affinity.

Nevertheless, although it is helpful to compare rather *conventional* print alternatives with web-based interventions, it should also be considered that future senior populations are likely to have more digital competence. Web-based interventions offer inclusionary benefits that can be of particular interest when dealing with an older adult population (eg, text-to-speech options and variable font sizes). Therefore, the design, content, formatting, and acceptance of digital or internet interventions should be further researched and tailored when implementing interventions for older adults.

In addition, a critical point is selective dropout (eg, higher attrition in either the web-based IG, in the groups that received only print-based material, or the WLCG). This was particularly evident in the first phase of the study among the groups with high technology requirements (web material and activity trackers that had to be synchronized with the website). Accordingly, future projects should support those at risk of dropping out of the intervention to remain in the study.

A further shortcoming of this study is that the interventions were designed to improve PA and did not explicitly reduce loneliness or symptoms of depression directly. As the findings are promising, future studies should follow up on this as PA can also be a very effective tool to address these emotions and cognitions. Parts of the data have already been published, such as the effects of the interventions on the stages of change in PROMOTE 1 [50,53]. Consequently, when using these data for a systematic review or meta-analysis, this needs to be taken into account.

In general, whether older adults benefit from improved health (other indicators in addition to symptoms of depression) or well-being (other indicators in addition to loneliness), health behavior, and its predictors from print-based material or web-based interventions still requires more attention, and future studies should follow up on the mechanisms that explain how the interventions work.

Process Analysis

The results of the moderated mediation analysis revealed that there is not only a simple linear effect of the intervention on symptoms of depression but also that the effects of the intervention on symptoms of depression are modeled in a more complex way. Therefore, it can be hypothesized that interventions generally tend to affect target variables in a nonlinear fashion and that, in general, interventions and health should be understood as complex systems. With the actor stage as a moderator, the conceivable conclusion is that only people who were physically active before the intervention benefited from it with regard to the symptoms of depression. Therefore, the HAPA theory is a helpful tool in designing interventions as it can assist in identifying a person's current stage of change and aid them in moving to the actor stage.

Li et al [47] found that subjective age is related to various health benefits. In their study, subjective age acted as a mediator between the intervention and symptoms of depression. With this in mind, the following can be assumed: in interventions, a set of variables that can act as potential mediators exists, as these variables are linked to several health benefits. However,

XSL•FO
RenderX

it is easy to overlook these mediators when only focusing on linear and bivariate effects, when, in fact, a mediation might only apply to a subgroup of individuals. Future research needs to identify this set of variables and implement it effectively in intervention frameworks. Subjective age might be one of these variables, although there might be more underlying components.

Comparison With Prior Work

Loneliness and mental health in the aging population are important topics, and our data match previous work showing that physical exercise is a key factor in addressing these issues [4,6]. However, the effects of the intervention were stronger regarding behavior adherence and stage (contingency coefficient 0.314), as well as behavioral habits (η^2=0.025-0.026), than regarding loneliness and depression (η^2=0.005-0.024). The finding of stage-specific effects (only actors benefited) with regard to depression matches the previous finding that interventions improving PA can also improve mood [7].

To date, little research has been conducted on whether web-based or print-based interventions are more likely to result in successful behavior change regarding PA, which can simultaneously affect feelings of loneliness and depression. By matching the findings of Golsteijn et al [9], this study revealed some benefits of the print-based intervention for this age group [9]. However, in general, the web-based and print-based interventions were more effective than the WLCG.

Aspects assumed to affect participation in, and effectiveness of, interventions have been previously studied. With respect to the uptake of PA offerings and intervention modalities, sociodemographic differences have emerged in the past. Accordingly, women, those with higher levels of education, and those who are already physically active are more likely to participate [23,75]. Physical inactivity, being overweight, and having a low educational status were indicators of discontinuation of the intervention before it was completed [22]. Furthermore, it is known that younger individuals seem to prefer web-based services [8,76], whereas older adults or women appear more likely to favor print-based offerings [76].

This and former analyses of the PROMOTE 1 study [50] identified higher attrition rates in the IGs than in WLCG, which is in line with other similar intervention studies [77]. This may be explained by the fact that the participants were more motivated to stay in the study as they received the program afterward. However, those who obtained the intervention later also predominantly maintained their behavior over the course of the intervention. Cunningham et al [78] found that wait-list groups interrupt efforts to change and pointed to the importance of activities toward changing different psychological preconditions and developmental possibilities. Therefore, to reduce the dropout rate in the WLCG, individuals randomized to this group should be informed about the progress and expected time of participation. As previous studies have indicated that longer waiting times are associated with higher dropout rates, the study design should be adapted to decrease the waiting time for the WLCG [8,22,54].

In addition, in the first phase of the study, the problem of increased dropout or greater dissatisfaction in the study arms with a need for greater technical skills and information technology acceptance became relevant [50]. This has also been reported by other authors [79]. As a reaction, the study design for PROMOTE 2 was adapted such that participants could change the type of materials provided at a defined time point on a preference basis, which could improve the loss of participants because of this aspect, as well as the satisfaction with the intervention in a recognizable way [52].

Conclusions

In times of physical distancing, for instance, during the COVID-19 pandemic, alternative forms of support, such as print- and web-based PA exercise content for individual implementation at home, are essential and in high demand. As users are very heterogeneous, tailoring PA interventions according to their specific needs (including differing motivations to engage in PA) and previous experiences (captured by stages), as well as according to their individual technology-based preconditions, could be an effective approach to initiating behavior change with regard to PA. Particularly, for digital interventions, the varying availability and use preferences of digital devices should be considered.

Tools for individual use, including activity monitoring (such as exercise diaries) and exercise instructions, are highly relevant for location- and time-independent use. Finally, but importantly, the relevance is determined by the reduced mobility in old age and, thus, the possibility or need for exercise at home.

Successful aging in terms of helping older adults feel fit to perform PA should be considered more explicitly, which can help to ensure that a PA intervention actually translates into the reduction of symptoms of depression. At the same time, taking the stage of change-specific aspects into account can benefit the knowledge that, similar to this study, the interventions worked well in intenders and actors for successful aging and symptoms of depression but not in nonintenders. Nonintenders might need other support such as just-in-time adaptive interventions and more instant social support, not only through print-based and web-based modes. However, the delivered interventions appeared to be supportive of intenders and actors and improve the predictors of behavior. Ingredients of the intervention's behavior change techniques (ie, goal setting, planning, social support, and feedback [58]) paid off.

Acknowledgments

This study was supported by research funding from the Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung), with project numbers 01EL1822A, 01EL1822C, 01EL1822F, and 01EL1822I. The funder was not involved in the design of the study; in the collection, analysis, and interpretation of data; or in writing the manuscript.

XSL•FO
RenderX

The authors would like to thank Grace Wire, Ronja Bellinghausen, and Veronika Hochdahl for manuscript editing and proofreading and Robin Rinn for his help in improving an earlier version of this manuscript.

Authors' Contributions

All authors made substantial contributions to the study conception (SL, CP, and CV-R), design of the study (SL, TR, and DJ), acquisition and analysis (SL, TR, FMK, DJ, and MP), data interpretation (SL, TR, FMK, DJ, MP, CP, and CV-R), manuscript drafting (SL, TR, and FMK), and substantial revision of the manuscript (SL, TR, FMK, DJ, MP, CP, and CV-R). All authors approved the submitted manuscript and agreed to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved and the resolution documented in the literature.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Effect of the web and print interventions in comparison to the wait-list control group on social-cognitive variables (Table S1), and on loneliness, perceived age and symptoms of depression (Table S2).
[DOCX File , 24 KB - aging_v5i3e36515_app1.docx]

Multimedia Appendix 2
Ordinal least squares regression model results for subjective age (Table S1), loneliness (Table S2), subjective age (Table S3) and depressive symptoms (Table S4).
[DOCX File , 29 KB - aging_v5i3e36515_app2.docx]

References

1. Groarke JM, Berry E, Graham-Wisener L, McKenna-Plumley PE, McGlinchey E, Armour C. Loneliness in the UK during the COVID-19 pandemic: cross-sectional results from the COVID-19 psychological wellbeing study. PLoS One 2020;15(9):e0239698 [FREE Full text] [doi: 10.1371/journal.pone.0239698] [Medline: 32970764]
2. Noone C, McSharry J, Smalle M, Burns A, Dwan K, Devane D, et al. Video calls for reducing social isolation and loneliness in older people: a rapid review. Cochrane Database Syst Rev 2020 May 21;5:CD013632 [FREE Full text] [doi: 10.1002/14651858.CD013632] [Medline: 32441330]
3. Savage RD, Wu W, Li J, Lawson A, Bronskill SE, Chamberlain SA, et al. Loneliness among older adults in the community during COVID-19: a cross-sectional survey in Canada. BMJ Open 2021 Apr 02;11(4):e044517 [FREE Full text] [doi: 10.1136/bmjopen-2020-044517] [Medline: 33811054]
4. Macy JT, Owens C, Mullis K, Middlestadt SE. The role of self-efficacy and injunctive norms in helping older adults decide to stay home during the COVID-19 pandemic. Front Public Health 2021;9:660813 [FREE Full text] [doi: 10.3389/fpubh.2021.660813] [Medline: 34150704]
5. Dickens AP, Richards SH, Greaves CJ, Campbell JL. Interventions targeting social isolation in older people: a systematic review. BMC Public Health 2011 Aug 15;11:647 [FREE Full text] [doi: 10.1186/1471-2458-11-647] [Medline: 21843337]
6. Pels F, Kleinert J. Loneliness and physical activity: a systematic review. Int Rev Sport Exerc Psychol 2016 Jun 07;9(1):231-260. [doi: 10.1080/1750984x.2016.1177849]
7. Alves GS, Casali ME, Veras AB, Carrilho CG, Bruno Costa E, Rodrigues VM, et al. A systematic review of home-setting psychoeducation interventions for behavioral changes in dementia: some lessons for the COVID-19 pandemic and post-pandemic assistance. Front Psychiatry 2020;11:577871 [FREE Full text] [doi: 10.3389/fpsyt.2020.577871] [Medline: 33132937]
8. Boekhout JM, Peels DA, Berendsen BA, Bolman C, Lechner L. A web-based and print-delivered computer-tailored physical activity intervention for older adults: pretest-posttest intervention study comparing delivery mode preference and attrition. J Med Internet Res 2019 Aug 28;21(8):e13416 [FREE Full text] [doi: 10.2196/13416] [Medline: 31464186]
9. Golsteijn RH, Peels DA, Evers SM, Bolman C, Mudde AN, de Vries H, et al. Cost-effectiveness and cost-utility of a web-based or print-delivered tailored intervention to promote physical activity among adults aged over fifty: an economic evaluation of the Active Plus intervention. Int J Behav Nutr Phys Act 2014 Sep 28;11:122 [FREE Full text] [doi: 10.1186/s12966-014-0122-z] [Medline: 25262435]
10. Fortuna KL, Torous J, Depp CA, Jimenez DE, Areán PA, Walker R, et al. A future research agenda for digital geriatric mental healthcare. Am J Geriatr Psychiatry 2019 Nov;27(11):1277-1285 [FREE Full text] [doi: 10.1016/j.jagp.2019.05.013] [Medline: 31196619]
11. Paggi ME, Jopp D, Hertzog C. The importance of leisure activities in the relationship between physical health and well-being in a life span sample. Gerontology 2016;62(4):450-458 [FREE Full text] [doi: 10.1159/000444415] [Medline: 26974682]

XSL•FO
RenderX

12. Sala G, Jopp D, Gobet F, Ogawa M, Ishioka Y, Masui Y, et al. The impact of leisure activities on older adults' cognitive function, physical function, and mental health. PLoS One 2019;14(11):e0225006 [FREE Full text] [doi: 10.1371/journal.pone.0225006] [Medline: 31703115]
13. Global recommendations on physical activity for health. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/44399/9789241599979_eng.pdf [accessed 2022-07-12]
14. Hupin D, Roche F, Gremeaux V, Chatard J, Oriol M, Gaspoz J, et al. Even a low-dose of moderate-to-vigorous physical activity reduces mortality by 22% in adults aged ≥60 years: a systematic review and meta-analysis. Br J Sports Med 2015 Oct;49(19):1262-1267. [doi: 10.1136/bjsports-2014-094306] [Medline: 26238869]
15. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. Can Med Assoc J 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
16. Södergren M, McNaughton SA, Salmon J, Ball K, Crawford DA. Associations between fruit and vegetable intake, leisure-time physical activity, sitting time and self-rated health among older adults: cross-sectional data from the WELL study. BMC Public Health 2012 Jul 25;12:551 [FREE Full text] [doi: 10.1186/1471-2458-12-551] [Medline: 22830932]
17. Abdelbasset WK, Alsubaie SF, Tantawy SA, Elyazed TI, Elshehawy AA. A cross-sectional study on the correlation between physical activity levels and health-related quality of life in community-dwelling middle-aged and older adults. Medicine (Baltimore) 2019 Mar;98(11):e14895 [FREE Full text] [doi: 10.1097/MD.0000000000014895] [Medline: 30882705]
18. Olson EA, McAuley E. Impact of a brief intervention on self-regulation, self-efficacy and physical activity in older adults with type 2 diabetes. J Behav Med 2015 Dec;38(6):886-898 [FREE Full text] [doi: 10.1007/s10865-015-9660-3] [Medline: 26162648]
19. Seguin RA, Economos CD, Palombo R, Hyatt R, Kuder J, Nelson ME. Strength training and older women: a cross-sectional study examining factors related to exercise adherence. J Aging Phys Act 2010 Apr;18(2):201-218 [FREE Full text] [doi: 10.1123/japa.18.2.201] [Medline: 20440031]
20. Dolansky MA, Stepanczuk B, Charvat JM, Moore SM. Women's and men's exercise adherence after a cardiac event. Res Gerontol Nurs 2010 Jan;3(1):30-38 [FREE Full text] [doi: 10.3928/19404921-20090706-03] [Medline: 20128541]
21. Elzen H, Slaets JP, Snijders TA, Steverink N. Do older patients who refuse to participate in a self-management intervention in the Netherlands differ from older patients who agree to participate? Aging Clin Exp Res 2008 Jun;20(3):266-271. [doi: 10.1007/BF03324777] [Medline: 18594195]
22. Jancey J, Lee A, Howat P, Clarke A, Wang K, Shilton T. Reducing attrition in physical activity programs for older adults. J Aging Phys Act 2007 Apr;15(2):152-165. [doi: 10.1123/japa.15.2.152] [Medline: 17556782]
23. Chinn DJ, White M, Howel D, Harland JO, Drinkwater CK. Factors associated with non-participation in a physical activity promotion trial. Public Health 2006 Apr;120(4):309-319. [doi: 10.1016/j.puhe.2005.11.003] [Medline: 16473376]
24. Picorelli AM, Pereira LS, 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 [FREE Full text] [doi: 10.1016/j.jphys.2014.06.012] [Medline: 25092418]
25. Lakerveld J, Ijzelenberg W, van Tulder MW, Hellemans IM, Rauwerda JA, van Rossum AC, et al. Motives for (not) participating in a lifestyle intervention trial. BMC Med Res Methodol 2008 Apr 10;8:17 [FREE Full text] [doi: 10.1186/1471-2288-8-17] [Medline: 18402683]
26. Justine M, Azizan A, Hassan V, Salleh Z, Manaf H. Barriers to participation in physical activity and exercise among middle-aged and elderly individuals. Singapore Med J 2013 Oct;54(10):581-586 [FREE Full text] [doi: 10.11622/smedj.2013203] [Medline: 24154584]
27. Wichmann F, Pischke CR, Jürgens D, Darmann-Finck I, Koppelin F, Lippke S, et al. Requirements for (web-based) physical activity interventions targeting adults above the age of 65 years - qualitative results regarding acceptance and needs of participants and non-participants. BMC Public Health 2020 Jun 11;20(1):907 [FREE Full text] [doi: 10.1186/s12889-020-08927-8] [Medline: 32527251]
28. Gourlan M, Bernard P, Bortolon C, Romain AJ, Lareyre O, Carayol M, et al. Efficacy of theory-based interventions to promote physical activity. A meta-analysis of randomised controlled trials. Health Psychol Rev 2016;10(1):50-66. [doi: 10.1080/17437199.2014.981777] [Medline: 25402606]
29. van Genugten L, Dusseldorp E, Webb TL, van Empelen P. Which combinations of techniques and modes of delivery in internet-based interventions effectively change health behavior? A meta-analysis. J Med Internet Res 2016 Jun 07;18(6):e155 [FREE Full text] [doi: 10.2196/jmir.4218] [Medline: 27268104]
30. Wilson TE, Hennessy EA, Falzon L, Boyd R, Kronish IM, Birk JL. Effectiveness of interventions targeting self-regulation to improve adherence to chronic disease medications: a meta-review of meta-analyses. Health Psychol Rev 2020 Mar;14(1):66-85 [FREE Full text] [doi: 10.1080/17437199.2019.1706615] [Medline: 31856664]
31. McEwan D, Beauchamp MR, Kouvousis C, Ray CM, Wyrough A, Rhodes RE. Examining the active ingredients of physical activity interventions underpinned by theory versus no stated theory: a meta-analysis. Health Psychol Rev 2019 Mar;13(1):1-17. [doi: 10.1080/17437199.2018.1547120] [Medline: 30412685]
32. Johnson BT, Scott-Sheldon LA, Carey MP. Meta-synthesis of health behavior change meta-analyses. Am J Public Health 2010 Nov;100(11):2193-2198 [FREE Full text] [doi: 10.2105/AJPH.2008.155200] [Medline: 20167901]

XSL•FO
RenderX

33. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF. Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis. Soc Sci Med 2017 Nov;192:125-133. [doi: 10.1016/j.socscimed.2017.09.021] [Medline: 28965003]
34. Rhodes RE, Janssen I, Bredin SS, Warburton DE, Bauman A. Physical activity: health impact, prevalence, correlates and interventions. Psychol Health 2017 Aug;32(8):942-975. [doi: 10.1080/08870446.2017.1325486] [Medline: 28554222]
35. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]
36. Powell SM, Larsen CA, Phillips SM, Pellegrini CA. Exploring beliefs and preferences for reducing sedentary behavior among adults with symptomatic knee osteoarthritis or knee replacement. ACR Open Rheumatol 2021 Jan;3(1):55-62 [FREE Full text] [doi: 10.1002/acr2.11216] [Medline: 33400397]
37. Nguyen MH, Smets EM, Bol N, Loos EF, Van Weert JC. How tailoring the mode of information presentation influences younger and older adults' satisfaction with health websites. J Health Commun 2018;23(2):170-180. [doi: 10.1080/10810730.2017.1421729] [Medline: 29345531]
38. Quittschalle J, Stein J, Luppa M, Pabst A, Löbner M, Koenig H, et al. Internet use in old age: results of a German population-representative survey. J Med Internet Res 2020 Nov 23;22(11):e15543 [FREE Full text] [doi: 10.2196/15543] [Medline: 33226351]
39. Neil-Sztramko SE, Smith-Turchyn J, Richardson J, Dobbins M. A mobility-focused knowledge translation randomized controlled trial to improve physical activity: process evaluation of the Move4Age study. J Med Internet Res 2019 Jun 20;21(6):e13965 [FREE Full text] [doi: 10.2196/13965] [Medline: 31223121]
40. Peels DA, Bolman C, Golsteijn RH, de Vries H, Mudde AN, van Stralen MM, et al. Long-term efficacy of a printed or a web-based tailored physical activity intervention among older adults. Int J Behav Nutr Phys Act 2013 Sep 02;10:104 [FREE Full text] [doi: 10.1186/1479-5868-10-104] [Medline: 24228862]
41. Tkatch R, Bazarko D, Musich S, Wu L, MacLeod S, Keown K, et al. A pilot online mindfulness intervention to decrease caregiver burden and improve psychological well-being. J Evid Based Complementary Altern Med 2017 Oct;22(4):736-743 [FREE Full text] [doi: 10.1177/2156587217737204] [Medline: 29228806]
42. Annear M, Keeling S, Wilkinson T, Cushman G, Gidlow B, Hopkins H. Environmental influences on healthy and active ageing: a systematic review. Ageing Soc 2012 Nov 06;34(4):590-622. [doi: 10.1017/s0144686x1200116x]
43. Beauchamp MR, Carron AV, McCutcheon S, Harper O. Older adults' preferences for exercising alone versus in groups: considering contextual congruence. Ann Behav Med 2007 Apr;33(2):200-206. [doi: 10.1007/BF02879901] [Medline: 17447872]
44. Mehra S, van den Helder J, Visser B, Engelbert RH, Weijs PJ, Kröse BJ. Evaluation of a blended physical activity intervention for older adults: mixed methods study. J Med Internet Res 2020 Jul 23;22(7):e16380 [FREE Full text] [doi: 10.2196/16380] [Medline: 32459652]
45. Volders E, de Groot RH, Coumans JM, Bolman CA, Lechner L. A randomized controlled trial into the cognitive effects of a computer-tailored physical activity intervention in older adults with chronic disease(s). Eur Rev Aging Phys Act 2021 Feb 25;18(1):3 [FREE Full text] [doi: 10.1186/s11556-021-00259-9] [Medline: 33632130]
46. Wentzel J, van der Vaart R, Bohlmeijer ET, van Gemert-Pijnen JE. Mixing online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016 Feb 09;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]
47. Li Y, Liu M, Miyawaki CE, Sun X, Hou T, Tang S, et al. Bidirectional relationship between subjective age and frailty: a prospective cohort study. BMC Geriatr 2021 Jun 29;21(1):395 [FREE Full text] [doi: 10.1186/s12877-021-02344-1] [Medline: 34187378]
48. Wiener J, Kuhlmann T, Fink S, Hambrecht R, Lippke S. Subjective age and planning strategies for physical activity: brief report about a domain-specific example from a cross-sectional observation study. Int J Sport Psychol 2017;48(4):448-458.
49. Muellmann S, Bragina I, Voelcker-Rehage C, Rost E, Lippke S, Meyer J, et al. Development and evaluation of two web-based interventions for the promotion of physical activity in older adults: study protocol for a community-based controlled intervention trial. BMC Public Health 2017 May 25;17(1):512 [FREE Full text] [doi: 10.1186/s12889-017-4446-x] [Medline: 28545506]
50. Muellmann S, Buck C, Voelcker-Rehage C, Bragina I, Lippke S, Meyer J, et al. Effects of two web-based interventions promoting physical activity among older adults compared to a delayed intervention control group in Northwestern Germany: results of the PROMOTE community-based intervention trial. Prev Med Rep 2019 Sep;15:100958 [FREE Full text] [doi: 10.1016/j.pmedr.2019.100958] [Medline: 31410347]
51. Pischke CR, Voelcker-Rehage C, Peters M, Ratz T, Pohlabeln H, Meyer J, et al. Implementation and effects of information technology-based and print-based interventions to promote physical activity among community-dwelling older adults: protocol for a randomized crossover trial. JMIR Res Protoc 2020 Apr 27;9(4):e15168 [FREE Full text] [doi: 10.2196/15168] [Medline: 32338622]

XSL•FO
RenderX

52. Pischke C, Voelcker-Rehage C, Ratz T, Peters M, Buck C, Meyer J, et al. Web-based versus print-based physical activity intervention for community-dwelling older adults: crossover randomized trial. JMIR Mhealth Uhealth 2022 Mar 23;10(3):e32212 [FREE Full text] [doi: 10.2196/32212] [Medline: 35319484]
53. Ratz T, Lippke S, Muellmann S, Peters M, Pischke CR, Meyer J, et al. Effects of two web-based interventions and mediating mechanisms on stage of change regarding physical activity in older adults. Appl Psychol Health Well Being 2020 Mar;12(1):77-100. [doi: 10.1111/aphw.12174] [Medline: 31332957]
54. Ratz T, Voelcker-Rehage C, Pischke CR, Muellmann S, Peters M, Lippke S. Health-related lifestyle and dropout from a web-based physical activity intervention trial in older adults: a latent profile analysis. Health Psychol 2021 Aug;40(8):481-490. [doi: 10.1037/hea0001091] [Medline: 34472906]
55. Forberger S, Bammann K, Bauer J, Boll S, Bolte G, Brand T, et al. How to tackle key challenges in the promotion of physical activity among older adults (65+): the Aequipa network approach. Int J Environ Res Public Health 2017 Apr 04;14(4):379 [FREE Full text] [doi: 10.3390/ijerph14040379] [Medline: 28375177]
56. Fleig L, Lippke S, Pomp S, Schwarzer R. Intervention effects of exercise self-regulation on physical exercise and eating fruits and vegetables: a longitudinal study in orthopedic and cardiac rehabilitation. Prev Med 2011 Sep;53(3):182-187. [doi: 10.1016/j.ypmed.2011.06.019] [Medline: 21784096]
57. Pomp S, Fleig L, Schwarzer R, Lippke S. Effects of a self-regulation intervention on exercise are moderated by depressive symptoms: a quasi-experimental study. Int J Clin Health Psychol 2013 Jan;13(1):1-8. [doi: 10.1016/S1697-2600(13)70001-2]
58. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
59. Nelson ME, Rejeski WJ, Blair SN, Duncan PW, Judge JO, King AC, et al. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007 Aug;39(8):1435-1445. [doi: 10.1249/mss.0b013e3180616aa2] [Medline: 17762378]
60. Lippke S, Ziegelmann JP, Schwarzer R, Velicer WF. Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. Health Psychol 2009 Mar;28(2):183-193 [FREE Full text] [doi: 10.1037/a0012983] [Medline: 19290710]
61. Nigg CR. There is more to stages of exercise than just exercise. Exerc Sport Sci Rev 2005 Jan;33(1):32-35. [Medline: 15640718]
62. Lippke S, Ziegelmann J, Schwarzer R. Behavioral intentions and action plans promote physical exercise: a longitudinal study with orthopedic rehabilitation patients. J Sport Exerc Psychol 2004;26(3):470-483. [doi: 10.1123/jsep.26.3.470]
63. Schwarzer R, Schuz B, Ziegelmann JP, Lippke S, Luszczynska A, Scholz U. Adoption and maintenance of four health behaviors: theory-guided longitudinal studies on dental flossing, seat belt use, dietary behavior, and physical activity. Ann Behav Med 2007 Apr;33(2):156-166. [doi: 10.1007/BF02879897] [Medline: 17447868]
64. Luszczynska A. An implementation intentions intervention, the use of a planning strategy, and physical activity after myocardial infarction. Soc Sci Med 2006 Feb;62(4):900-908. [doi: 10.1016/j.socscimed.2005.06.043] [Medline: 16095786]
65. Lippke S, Fleig L, Pomp S, Schwarzer R. Validity of a stage algorithm for physical activity in participants recruited from orthopedic and cardiac rehabilitation clinics. Rehabil Psychol 2010 Nov;55(4):398-408. [doi: 10.1037/a0021563] [Medline: 21171799]
66. Verplanken B, Orbell S. Reflections on past behavior: a self-report index of habit strength. J Appl Social Pyschol 2003 Jun;33(6):1313-1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
67. Radloff LS. The CES-D Scale: a self-report depression scale for research in the general population. Applied Psychol Measure 1977;1(3):385-401. [doi: 10.1177/014662167700100306]
68. Goeßwald A, Lange M, Kamtsiuris P, Kurth BM. DEGS: studie zur Gesundheit Erwachsener in Deutschland. Bundesweite Quer- und Laengsschnittstudie im Rahmen des Gesundheitsmonitorings des Robert Koch-Instituts. Bundesgesundheitsbl Gesundheitsforsch Gesundheitsschutz 2012;55:775-780. [doi: 10.5414/dbx00211]
69. Fragebögen nach Themenfeldern. Sozialplanung für Senioren. URL: http://www.sozialplanung-senioren.de/frageboegen-bausteine/frageboegen-nach-themenfeldern/index.html [accessed 2022-01-12]
70. UNESCO United Nations Educational, Scientific and Cultural Organization. International Standard Classification of Education, ISCED 1997. In: Hoffmeyer-Zlotnik JHP, Wolf C, editors. Advances in Cross-National Comparison. Boston, MA: Springer; 2003.
71. Body mass index - BMI. World Health Organization. URL: https://www.euro.who.int/en/health-topics/disease-prevention/nutrition/a-healthy-lifestyle/body-mass-index-bmi [accessed 2022-01-12]
72. 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]
73. Baek MJ, Kim K, Park YH, Kim S. The validity and reliability of the mini-mental state examination-2 for detecting mild cognitive impairment and Alzheimer's disease in a Korean population. PLoS One 2016;11(9):e0163792 [FREE Full text] [doi: 10.1371/journal.pone.0163792] [Medline: 27668883]
74. Rubin DB. Inference and Missing Data. Biometrika 1976 Dec;63(3):581-592. [doi: 10.2307/2335739]

XSL•FO
RenderX

75. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
76. Vandelanotte C, Duncan MJ. Individual characteristics associated with physical activity intervention delivery mode preferences among adults. Int J Behav Nutr Phys Act 2014 Feb 25;11(1):25 [FREE Full text] [doi: 10.1186/1479-5868-11-25] [Medline: 24568611]
77. Spittaels H, De Bourdeaudhuij I, Vandelanotte C. Evaluation of a website-delivered computer-tailored intervention for increasing physical activity in the general population. Prev Med 2007 Mar;44(3):209-217. [doi: 10.1016/j.ypmed.2006.11.010] [Medline: 17197015]
78. Cunningham JA, Kypri K, McCambridge J. Exploratory randomized controlled trial evaluating the impact of a waiting list control design. BMC Med Res Methodol 2013 Dec 06;13:150 [FREE Full text] [doi: 10.1186/1471-2288-13-150] [Medline: 24314204]
79. Oenema A, Brug J, Dijkstra A, de Weerdt I, de Vries H. Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial. Ann Behav Med 2008 Apr;35(2):125-135. [doi: 10.1007/s12160-008-9023-1] [Medline: 18363076]

Abbreviations

CES-D: Center for Epidemiologic Studies Depression
HAPA: Health Action Process Approach
IG: intervention group
MANOVA: multivariate analysis of variance
MMSE: Mini-Mental State Examination
MMSE-2: BV: Mini-Mental State Examination second edition: Brief Version
MVPA: moderate to vigorous physical activity
PA: physical activity
T0: time point 0
T1: time point 1
WHO: World Health Organization
WLCG: wait-list control group

Edited by T Leung, J Wang; submitted 18.01.22; peer-reviewed by A Joseph, M Leslie, K Wall; comments to author 24.02.22; revised version received 20.04.22; accepted 25.05.22; published 09.08.22.

Please cite as:
Lippke S, Ratz T, Keller FM, Juljugin D, Peters M, Pischke C, Voelcker-Rehage C
Mitigating Feelings of Loneliness and Depression by Means of Web-Based or Print-Based Physical Activity Interventions: Pooled Analysis of 2 Community-Based Intervention Trials
JMIR Aging 2022;5(3):e36515
URL: https://aging.jmir.org/2022/3/e36515
doi:10.2196/36515
PMID:35943790

©Sonia Lippke, Tiara Ratz, Franziska Maria Keller, Dennis Juljugin, Manuela Peters, Claudia Pischke, Claudia Voelcker-Rehage. Originally published in JMIR Aging (https://aging.jmir.org), 09.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Objectively Measured Physical Activity Levels and Associated Factors in Older US Women During the COVID-19 Pandemic: Cross-sectional Study

Renoa Choudhury[1], MSc; Joon-Hyuk Park[1,2], PhD; Ladda Thiamwong[3], PhD; Rui Xie[4], PhD; Jeffrey R Stout[5], PhD

[1]Department of Mechanical and Aerospace Engineering, University of Central Florida, Orlando, FL, United States

[2]Disability, Aging and Technology Cluster, University of Central Florida, Orlando, FL, United States

[3]College of Nursing, University of Central Florida, Orlando, FL, United States

[4]Department of Statistics and Data Science, University of Central Florida, Orlando, FL, United States

[5]School of Kinesiology and Physical Therapy, College of Health Professions and Sciences, University of Central Florida, Orlando, FL, United States

Corresponding Author:
Joon-Hyuk Park, PhD
Department of Mechanical and Aerospace Engineering
University of Central Florida
4000 Central Florida Blvd
Orlando, FL, 32816
United States
Phone: 1 407 823 1227
Email: Joonpark@ucf.edu

Abstract

Background: Physical activity (PA) is vital for attenuating the aging-related physiological and functional declines in women aged 60 years or above. However, little is known about the objectively assessed PA behavior in older women during the COVID-19 pandemic and its association with sociodemographics, health and physical function, and COVID-19 related factors.

Objective: This study aims to examine the objectively measured PA levels and associated factors among older US women who were living under the physical distancing guidelines during the second year of the pandemic.

Methods: In this cross-sectional study, we collected free-living PA data from 94 community-dwelling older women aged between 60 and 96 years (mean age 75.1 years, SD 7.3) using wrist-worn ActiGraph GT9X accelerometers between February and August 2021. We examined whether their daily duration spent in sedentary behavior (SB), light-intensity physical activity (LPA), and moderate-to-vigorous-intensity physical activity (MVPA) varied by sociodemographic characteristics, health and physical function, and COVID-19 related factors.

Results: On average, participants accumulated 12.4 (SD 1.9) hours/day in SB, 218.6 (SD 64.3) minutes/day in LPA, and 42.4 (SD 31.0) minutes/day in MVPA, exhibiting overall reduced PA levels than previously published pre–COVID-19 norms of older US women. Among participants aged ≥80 years, sedentary time was 7.5% (P=.003) higher and the time spent in LPA and MVPA was, respectively, 13.3% (P=.03) and 44.9% (P<.001) lower than those aged 60-79 years. More MVPA participation and a less sedentary lifestyle were observed in those who had a higher self-rated health score (MVPA: P=.001, SB: P=.04) and lower fear of falling (FOF; MVPA: P=.003, SB: P=.04). Poorer performance in the 30-second sit-to-stand (STS) test was independently associated with more SB (P=.01) and less LPA (P=.04) and MVPA (P=.001) time among participants. In addition, sedentary time was 5.0% higher (P=.03) in frail and prefrail participants than their healthy counterparts.

Conclusions: During the pandemic, older women spent the majority of their waking time being sedentary, while LPA accounted for a larger portion of their daily PA. Therefore, replacing SB with LPA (rather than MVPA) might provide a more feasible PA target for older women, particularly those aged ≥80 years or who have reduced physical function. In addition, targeted interventions might be beneficial in promoting an active lifestyle for those who live alone, are prefrail or frail, and have a high FOF in older age.

International Registered Report Identifier (IRRID): RR2-10.2196/27381

(JMIR Aging 2022;5(3):e38172) doi:10.2196/38172

XSL•FO
RenderX

KEYWORDS

physical activity; older women; COVID-19; sedentary behavior; wrist-worn accelerometers; ActiGraph; aging; elderly population; women's health; digital health; frail; healthy lifestyle

Introduction

Background

The worldwide outbreak of COVID-19, caused by a novel type of coronavirus (SARS-COV-2), was declared a global pandemic by the World Health Organization (WHO) in March 2020. As of March 22, 2022, there were 470,839,745 COVID-19 cases confirmed worldwide, resulting in approximately 5,944,342 deaths [1]. Age and immune-compromised states are directly linked to the severity and fatality of COVID-19, making older adults the most significantly afflicted, particularly those with pre-existing health conditions (eg, chronic respiratory diseases, diabetes, hypertension, cardiovascular diseases, chronic kidney diseases) [2]. Hence, the Centers for Disease Control and Prevention (CDC) has been recommending strong adherence to physical distancing guidelines (previously known as social distancing guidelines) to the older adult population [3]. Despite the easing of stay-at-home order restrictions and advancement in the rapid, safe production and distribution of authorized vaccines, the physical distancing recommendations were still effective in the United States during the second year of the pandemic (ie, 2021), accompanied by the ongoing vaccination process and the emergence of new variants of concern (eg, Delta, Omicron).

However, practicing physical distancing may cause many older adults to limit their social interactions and out-of-home activities in community settings, which, in turn, is likely to affect their habitual physical activity (PA) level. Concerns regarding reduced PA levels resulting from COVID-19 mitigation strategies are particularly relevant for women aged 60 years or above, as they are the least active segment of the US population when evaluated against the current PA guidelines (ie, ≥150 minutes/week of moderate-to-vigorous-intensity physical activity [MVPA]). Federal monitoring data show that less than 20% of older US women were engaged in insufficient MVPA, even prior to the pandemic [4].

Therefore, it is important to examine the factors that may influence PA behavior in older women during the pandemic so that age- and gender-appropriate interventions can be tailored. To date, many studies have reported the factors associated with PA participation in older adults across different regions of the world during different phases of the pandemic, but the majority of these studies relied on self-report questionnaires and surveys for PA assessment [3,5-10]. Despite large and statistically robust sample sizes, these prior studies should be interpreted with caution because they can often be subjected to measurement biases (eg, social desirability bias and recall bias) and may not accurately capture the lower end of the PA spectrum (ie, light-intensity physical activity [LPA]) [11]. Objective PA measures can overcome these limitations of self-report questionnaires and provide continuous evaluation of 1 or more dimensions of PA (eg, frequency, intensity, and duration) in free-living conditions [12]. So far, a small number of studies have investigated objectively measured PA levels among Japanese, Swedish, and Brazilian older adults during the pandemic [13-16]; however, no study has yet reported objectively measured PA levels among older US people in the context of the pandemic.

Identifying nonmodifiable factors (eg, age, race/ethnicity, educational attainment, living alone) [17] associated with PA participation can help us in recognizing and targeting subgroups of older women who have been at higher risk of negative health consequences resulting from physical distancing adherence and are in most need of tailored PA interventions. Conversely, understanding the role of potentially modifiable risk factors (eg, overweight status, prefrail symptoms, fear of falling [FOF], upper and lower extremity strength) for reduced PA level during the pandemic might aid in developing evidence-based programs to enhance PA behavior in older women. Studies have reported obesity [18,19], frailty incidence [20], and decline in physical function (eg, upper and lower body strength) [21,22] to be associated with lower PA levels in older adults in the prepandemic period. In addition, FOF has been previously linked with more sedentary time and less duration in all other PA domains (both light and moderate-to-vigorous intensity) in a large cohort of older British men [23]. During the pandemic, older adults have been more vulnerable to social isolation and disconnectedness compared to the prepandemic time, due to changes in their lifestyle under the physical distancing guidelines. Given that social isolation poses a higher risk of frailty progression [24,25], increased FOF [26], and reduced physical functioning [27], it is important to understand how these factors are associated with different PA intensities among older adults in the light of physical distancing recommendations. This, in turn, will allow us to learn from this COVID-19 pandemic regarding PA strategies for older adults during social distancing and pandemic-related regulations in order to better prepare us for any similar instances possible in the future.

Goal of This Study

This study aims to examine PA levels among a diverse sample of older US women who were living under the physical distancing guidelines during the second year of the COVID-19 pandemic using wrist-worn accelerometry-based analysis. More specifically, we investigated whether their daily time spent in sedentary behavior (SB), LPA, and MVPA differed by (1) sociodemographic status (ie, age, race/ethnicity, education level, and household composition), (2) health and physical function (ie, BMI, self-rated health, frailty, FOF, grip strength, and sit-to-stand [STS] performance), and (3) COVID-19–related factors (ie, history of being COVID-19 positive, fear of COVID-19, and perceived severity of COVID-19 in their community).

XSL•FO
RenderX

Methods

Study Design and Participants

In this cross-sectional study, an opportunistic sample of 94 community-dwelling older women aged 60-96 years was recruited from the region of Central Florida, USA, between February and August 2021. Participants were recruited via word of mouth and flyers distributed in their communities. The inclusion criteria were that participants must be aged ≥60 years, be able to walk (with or without assistive devices but not requiring assistance from another person), have no marked cognitive impairment, live in their own homes or apartments, and be fluent in English or Spanish. The exclusion criteria were (1) having a medical condition that may preclude engagement in PA (eg, shortness of breath, dizziness, tightness or pain in the chest, and unusual fatigue at rest or with light exertion) and (2) currently receiving treatment from a rehabilitation facility. This cross-sectional assessment required 1 visit to the study site during which participants completed the informed consent, a self-report questionnaire, the Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight (FRAIL) scale, and Short Falls Efficacy Scale-International (short FES-I), followed by the assessment of grip strength and STS performance. At the end of the visit, each participant was fitted with a wrist-worn accelerometer and given instructions on how to wear it during the PA-monitoring period.

Ethical Considerations

The study protocol was approved by the Institutional Review Board of the University of Central Florida (Protocol No. 2189; September 10, 2020). All procedures were approved by the University of Central Florida's Institutional Review Board (#00003029). All experimental procedures were performed in accordance with the University of Central Florida COVID-19 Human Subject Research Standard Safety Plan.

Objective Measurement of PA

ActiGraph GT9X Link (ActiGraph LLC., Pensacola, FL, USA) was used to measure PA levels among participants. It is a small (3.5 × 3.5 × 1 cm), lightweight (14 g) wrist-worn device and contains a triaxial accelerometer with a dynamic range of ±8 gravitational units (g). Participants were required to wear it on the nondominant wrist for 7 consecutive days in free-living conditions. They were instructed to remove it only during sleeping, showering or swimming, and medical imaging tests. The accelerometer was initialized to record data at a sampling rate of 30 Hz. After 7 days of PA data collection, the ActiGraph devices were collected from the participants. At least 6 valid days of data were required for a participant to be included in the analysis, and only days during which the accelerometer was worn for at least 14 hours were counted as valid days.

Raw acceleration data from the ActiGraph devices were downloaded and converted to ".csv" files using ActiLife 6 v6.13.4 (ActiGraph LLC.). Next, data processing was performed in R statistical software (R Foundation for Statistical Computing) using the GGIR package (version 2.4-0) [28]. Data processing steps in GGIR include (1) autocalibration of acceleration signals according to local gravity [29], (2) detection of nonwear time, and (3) calculation of the average magnitude of dynamic acceleration corrected for gravity (ie, Euclidean Norm Minus 1 gravitational unit [g; ENMO]) over 5-second epochs, with negative values rounded to 0 [30]. ENMO was expressed in milligravitational units (mg) and defined as [31]:

$$\text{ENMO (mg)} = r_i - 1000$$

where r_i is the i-th vector magnitude at each time point and 1 g=1000 mg.

Nonwear time and sustained abnormally high accelerations (ie, ≥5.5 g) were imputed using the default settings, described in detail elsewhere [30]. ENMO cut-off points were used to estimate the total time spent in SB, LPA, and MVPA among the participants. The following cut-off points for nondominant wrist-worn accelerometry for older adults were adopted from the literature [32,33]: (1) SB<30 mg, (2) 30 mg≤LPA<100 mg, and (3) MVPA≥100 mg. In addition, to understand the impact of the pandemic on PA levels among older women, our data were compared with 2 pre–COVID-19 observational studies that had large-scale older women population samples and used a similar protocol and data processing methods as ours [34,35].

Assessment of Factors

A self-report questionnaire was used to obtain sociodemographic characteristics of participants. Based on age, they were categorized into 2 groups, *60-79 years* and *≥80 years*. According to race/ethnicity, participants were grouped into *White* and *non-White* groups, where the non-White category included African American, Asian, and Hispanic older women. The level of education was divided into 2 categories, *high school or lower* and *college or higher*. Household composition was defined as *living alone* or *living with family*. Height was measured using a stadiometer, and weight was measured using a digital scale. The BMI was calculated as weight (kg) divided by the square of height (m^2). Based on BMI, the participants were categorized into *normal weight* (BMI<25 kg/m^2), overweight (BMI=25-29.9 kg/m^2), and obese (BMI≥30 kg/m^2) [36]. Self-rated health was obtained using a 5-point Likert scale, and participants were classified into health status categories of *excellent* (score=1), *very good* (score=2), and *good or fair* (score≤3). Frailty was assessed using FRAIL scale, a 5-item self-report tool measuring fatigue, resistance, ambulation, illnesses, and loss of weight [37]. Based on the FRAIL score, participants were screened as *healthy* (score=0), *prefrail* (score=1-2), and *frail* (score=3-5). FOF was assessed using short FES-I, a 7-item self-administered tool measuring the level of concern about falling while performing 7 activities on a 4-point Likert scale (1=not at all concerned to 4=very concerned) [38]. A short FES-I score between 7 and 10 indicated a low concern of falling, while a score between 11 and 28 indicated a high concern of falling.

Grip strength, an indicator of hand and forearm muscle strength, was measured using a hydraulic hand dynamometer (JAMAR 5030J1, Patterson Medical), following the procedures adopted by the American Society of Hand Therapists, described in detail elsewhere [39]. Participants were categorized into *low* and *regular* grip strength groups based on the revised sarcopenia cut-off point (ie, <16 kg) recommended by the European

XSL•FO
RenderX

Working Group on Sarcopenia in Older People (EWGSOP) [40]. The 30-second STS test (also known as the chair-stand test) was used to assess lower limb muscle strength, endurance, and balance among participants [39]. STS performance was divided into *below average*, *average*, and *above average* based on the age- and gender-specific normative scores provided by Rikli and Jones [41].

In addition, participants were asked whether they had ever tested positive for COVID-19. They also rated their perception of COVID-19 severity in their community over the past month on a 4-point Likert scale (1=extremely high, 2=moderately high, 3=severe, 4=not severe). Fear of COVID-19 among the participants was assessed using the Fear of COVID-19 Scale (FCV-19S), a 7-item, 4-point Likert scale adapted from Ahorsu et al [42]. An FCV-19S score between 7 and 21 was defined as normal fear of COVID, while a score between 22 and 35 indicated elevated fear of COVID.

Statistical Analysis

All statistical analyses were performed using R statistical software (version 4.1.2) with a significance level (α) of .05. According to the Shapiro-Wilk test, SB and LPA among the participants were normally distributed, but MVPA showed nonnormal distribution. Descriptive statistics of PA variables (expressed as % of total wake time) were presented as means and SDs for normally distributed data and as medians and IQRs for nonnormally distributed data. For normally distributed PA variables, differences across 2 categories and more than 2 categories were examined, respectively, using independent *t* tests and 1-way ANOVA, with Bonferroni adjustment for post hoc comparisons. For nonnormally distributed variable, we performed the Mann-Whitney *U* test and the Kruskal-Wallis test, which are nonparametric equivalences of independent *t* tests and ANOVA, respectively.

Multiple linear regression analysis was performed on each PA outcome variable (ie, SB, LPA, and MVPA, expressed in minutes/day) to examine their adjusted associations with different independent variables. The independent variables included age, household composition, self-rated health score, frailty status, FOF score, and STS performance. For each PA outcome, model 1 was adjusted for the BMI and total wear time and model 2 was adjusted for the BMI and other 2 PA intensities. A priori sample size calculation revealed that the minimum required sample size for 9 explanatory variables at a statistical power level of 0.8 and a medium effect size (Cohen f^2=0.2) would be 87; therefore, our sample size (ie, N=94) was sufficient for multiple regression. Before conducting the regression, multicollinearity was checked by examining the correlation matrix of independent variables for any correlation coefficient value >0.8. In addition, $\log_{10}$ transformation was performed on MVPA (minutes/day) in order to meet the linear regression assumption of normality of residuals. To aid interpretation, while presenting outcomes for MVPA models, regression coefficients were back-transformed using the formula $100\times(\exp^{\beta} - 1)$ to indicate the percentage change in MVPA (minutes/day) for 1-unit change in the corresponding independent variable [43].

Results

Participant Details

The mean age of participants was 75.1 (SD 7.3) years, and 23 (25%) participants were aged 80 years or above. The mean BMI was 26.85 (SD 5.42) kg/m^2, and 39 (42%) participants were screened as prefrail. The mean grip strength was 19.0 (SD 5.6) kg, and the mean 30-second STS score was 14 (SD 6) repetitions (reps). The median accelerometer wear period for participants was 16.5 (IQR 15.5-17.6) hours/day. In addition, 85 (90%) participants had valid data (ie, ≥14 hours/day) on all 7 days. For the remaining participants (n=9, 10%), valid data were available for 6 days. All participants were included in the analysis. Among participants, the mean time spent in SB, LPA, and MVPA was 12.4 (SD 1.9) hours/day, 218.6 (SD 64.3) minutes/day, and 42.4 (SD 31.0) minutes/day, respectively. When expressed as a percentage of total waking time, the mean time accumulated in SB, LPA, and MVPA was 74.0% (SD 7.9%), 21.8% (SD 6.0%), and 4.2% (SD 3.0%), respectively.

Table 1 summarizes the descriptive statistics and the results from univariate analyses (parametric: independent *t* tests and 1-way ANOVA; nonparametric: Mann-Whitney *U* test and Kruskal-Wallis test) between PA variables and all factors. The average sedentary time was significantly higher (*P*=.003) in participants aged 80 years or above compared to those aged 60-79 years (78.10%, SD 7.49%, vs 72.70%, SD 7.54%), as shown in Figure 1. In addition, participants in the ≥80 years age group accumulated significantly less time in LPA (19.50%, *P*=.03) and MVPA (2.12%, *P*=.001) than those in the 60-79 years age group (LPA: 22.50%; MVPA: 3.85%). We observed that time spent in MVPA was significantly higher (*P*=.001) in participants who lived with their family compared to those living alone. However, no significant group differences were observed across race, education level, and BMI categories for any of PA variables in the current sample.

XSL•FO
RenderX

Table 1. Time spent in SB[a], LPA[b], and MVPA[c] (expressed as % of total wake time), stratified by sociodemographic characteristics, health and physical function, and COVID-19–related factors.

Participant characteristics	Participants, n (%)	SB (%), mean (SD)	LPA (%), mean (SD)	MVPA (%), mean (SD)
Total	94 (100)	73.98 (7.85)	21.79 (6.04)	3.54 (3.53)
Age (years)				
60-79	71 (75)	72.70 (7.54)	22.50 (5.84)	3.85 (3.50)
≥80	23 (25)	78.10 (7.49)	19.50 (6.18)	2.12 (2.76)
P value	N/A[d]	.003	.03	<.001
Race				
Non-White	23 (25)	76.06 (7.73)	20.68 (6.03)	3.27 (1.61)
White	71 (75)	73.31 (7.82)	22.14 (6.04)	3.90 (3.74)
P value	N/A	.15	.32	.07
Education				
College or higher	67 (71)	73.73 (7.39)	21.71 (5.44)	3.54 (3.69)
High school or lower	27 (29)	74.62 (9.01)	21.97 (7.42)	3.36 (3.24)
P value	N/A	.63	.85	.08
Household composition				
Living alone	45 (48)	75.58 (8.58)	20.70 (6.5)	3.31 (3.07)
Living with family	49 (52)	72.52 (6.87)	22.78 (5.46)	3.97 (3.43)
P value	N/A	.06	.10	.001
BMI[e]				
Normal weight	38 (40)	72.23 (7.12)	22.67 (5.21)	3.94 (2.80)
Overweight	32 (34)	74.24 (8.84)	22.00 (7.11)	3.21 (3.83)
Obese	24 (26)	76.41 (7.11)	20.10 (5.59)	2.96 (3.15)
P value	N/A	.12	.26	.06
Self-rated health				
Excellent	14 (15)	70.07 (7.98)	23.64 (5.46)	5.02 (5.13)
Very good	36 (38)	73.18 (8.28)	22.04 (6.64)	3.98 (3.25)
Good or fair	44 (47)	75.88 (6.99)	20.98 (5.67)	3.26 (2.29)
P value	N/A	.04	.34	.001
Frailty status				
Prefrail or frail	39 (42)	76.10 (7.70)	20.23 (5.79)	3.15 (3.57)
Healthy	55 (58)	72.48 (7.66)	22.88 (6.02)	3.65 (3.22)
P value	N/A	.03	.34	.06
FOF[f]				
High	36 (38)	76.12 (8.76)	20.70 (7.20)	2.60 (3.80)
Low	58 (62)	72.66 (6.97)	22.45 (5.14)	3.94 (3.45)
P value	N/A	.04	.28	.003
Grip strength				
Low (<16 kg)	40 (43)	73.21 (8.02)	22.77 (5.77)	3.46 (2.77)
Regular (≥16 kg)	54 (57)	74.56 (7.74)	21.05 (6.18)	3.56 (3.55)
P value	N/A	.42	.17	.39
STS[g] performance				

XSL•FO
RenderX

Participant characteristics	Participants, n (%)	SB (%), mean (SD)	LPA (%), mean (SD)	MVPA (%), mean (SD)
Below average	13 (14)	80.42 (7.22)	17.72 (6.12)	1.13 (1.57)
Average	39 (41)	73.40 (8.15)	22.39 (6.60)	3.58 (3.46)
Above average	42 (45)	72.56 (6.87)	22.48 (5.03)	3.99 (3.30)
P value	N/A	.004	.03	<.001
History of being COVID-19 positive				
No	87 (93)	74.01 (8.03)	21.66 (6.13)	3.54 (3.83)
Yes	7 (7)	73.60 (5.37)	23.28 (4.84)	2.65 (1.15)
P value	N/A	.89	.50	.42
Fear of COVID-19				
Elevated fear	8 (9)	74.01 (5.48)	23.0 (4.73)	3.21 (1.26)
Normal fear	86 (91)	73.98 (8.06)	21.67 (6.16)	3.57 (3.84)
P value	N/A	.99	.56	.26
Perceived severity of COVID-19 in community				
Severe or moderately severe	42 (45)	74.69 (7.54)	21.29 (6.01)	3.38 (2.54)
Not severe	52 (55)	73.41 (8.11)	22.18 (6.09)	3.62 (4.13)
P value	N/A	.48	.60	.51

[a]SB: sedentary behavior.

[b]LPA: light-intensity physical activity.

[c]MVPA: moderate-to-vigorous-intensity physical activity.

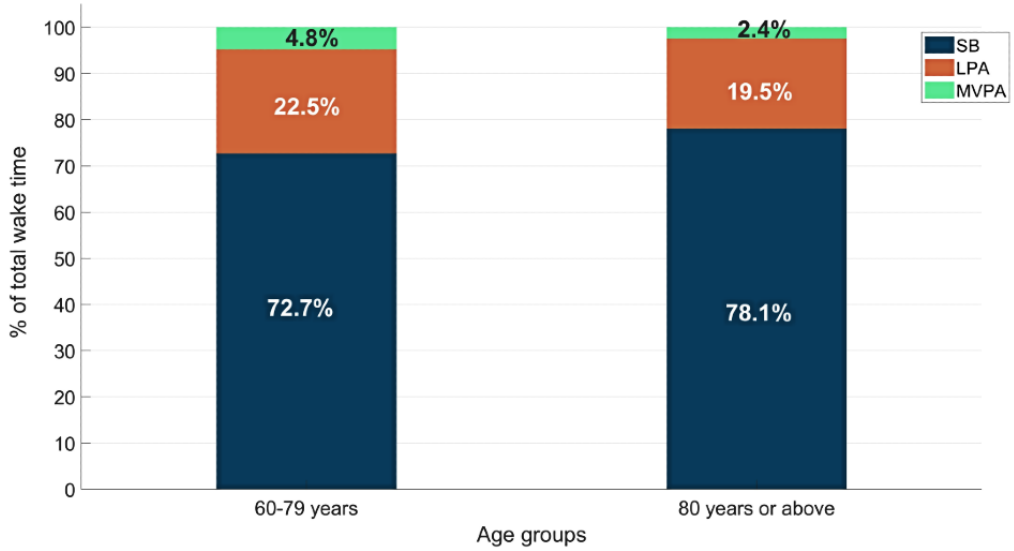
[d]N/A: not applicable.

[e]BMI: Body Mass Index.

[f]FOF: fear of falling.

[g]STS: sit-to-stand.

Figure 1. Distribution of mean PA levels by age group. LPA: light-intensity physical activity; MVPA: moderate-to-vigorous-intensity physical activity; PA: physical activity; SB: sedentary behavior.



The average sedentary time for participants with excellent health status was significantly lower than those rating their health as fair or poor (mean 70.07%, SD 7.98%, vs mean 75.88%, SD 6.99%, *P*=.01), as shown in Figure 2. Conversely, the mean time accumulated in MVPA was lower in participants with fair or poor health status (mean 3.26%, SD 2.29%) compared to those with excellent (mean 5.02%, SD 5.13%, *P*=.001) and very good (mean 3.98%, SD 3.25%, *P*=.04) health. We also found that sedentary time accumulated in prefrail and frail participants was significantly higher than that of participants with robust health (mean 76.10%, SD 7.70%, vs mean 72.48%, SD 7.66%, *P*=.03).

XSL•FO
RenderX

In the current sample, older women with low FOF participated more in MVPA (*P*=.003) and spent less time in being sedentary (*P*=.04) than their high-FOF counterparts (Figure 3). Regarding physical function, significant group differences were observed for STS performance but not for grip strength. Participants with below-average STS scores accumulated more sedentary time (mean 80.42%, SD 7.22%) and less MVPA time (mean 1.13%, SD 1.57%) than those with average (SB: mean 73.40%, SD 8.15%, *P*=.01; MVPA: mean 3.58%, SD 3.46%, *P*=.001) and above-average scores (SB: mean 72.56%, SD 6.87%, *P*=.001; MVPA: mean 3.99%, SD 3.30%, *P*<.001), as shown in Figure 4. In addition, time spent in LPA was significantly higher in participants with above-average STS scores compared to those with below-average scores (mean 22.48%, SD 5.03%, vs mean 17.72%, SD 6.12%, *P*=.01).

However, none of the COVID-19–related factors showed significant group differences for any of the PA variables. This might be attributed to the fact that only a small proportion of our study participants had a history of being COVID-19 positive and showed elevated fear of COVID-19 (ie, n=7, 7%, and n=8, 9%, respectively).

Figure 2. Time spent in (a) SB and (b) MVPA across categories of the self-rated health score. *P<.05 and **P<.01. MVPA: moderate-to-vigorous-intensity physical activity; SB: sedentary behavior.

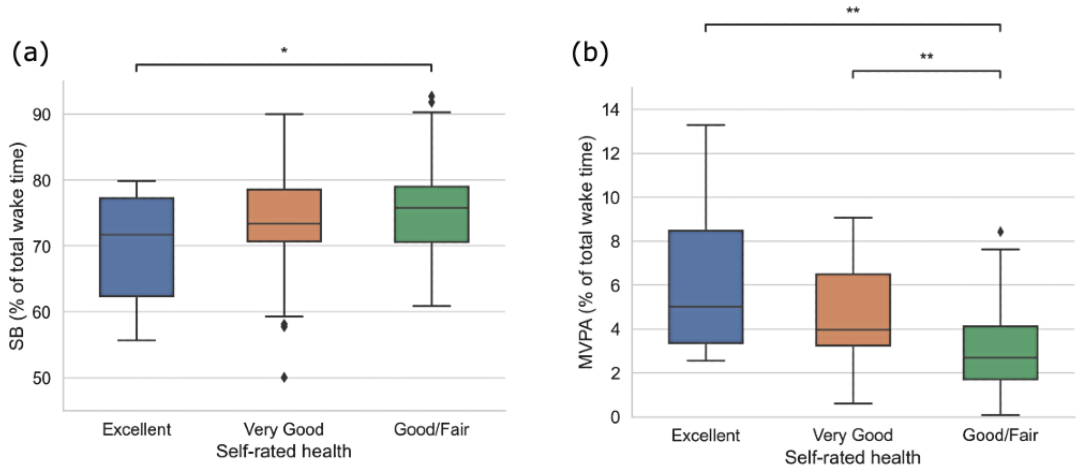


Figure 3. Time spent in (a) SB and (b) MVPA according to the FOF. *P<.05 and **P<.01. FOF: fear of falling; MVPA: moderate-to-vigorous-intensity physical activity; SB: sedentary behavior.

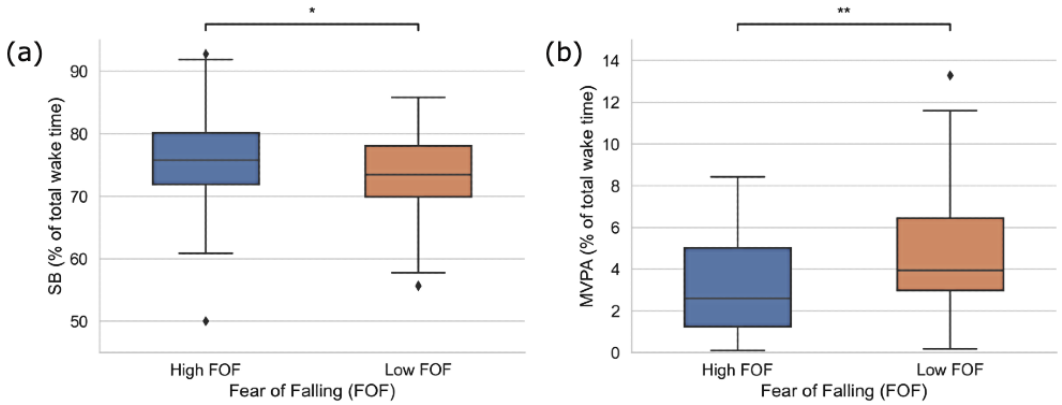
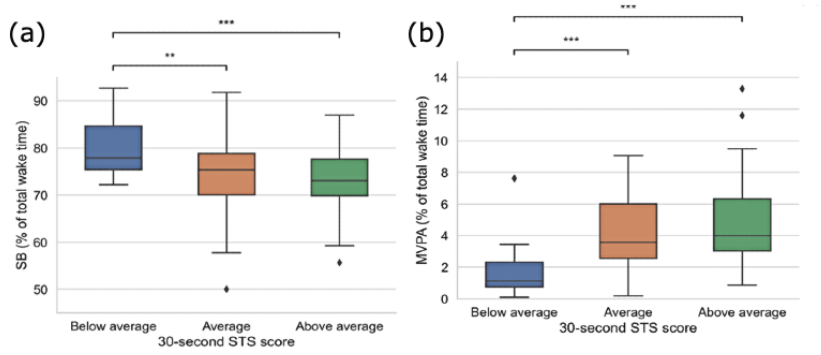


Figure 4. Time spent in (a) SB and (b) MVPA across categories of the 30-second STS score. **P<.01 and ***P<.001. MVPA: moderate-to-vigorous-intensity physical activity; SB: sedentary behavior; STS: sit-to-stand.



XSL•FO
RenderX

Table 2 presents the multiple regression models for sedentary time (minutes/day). In model 1, STS performance showed significant negative association (β=–3.92, *P*=.008) with the time spent in SB. This indicates that, in this current sample, an increase of 1-unit in STS score would result in a decrease of sedentary time by 3.92 minutes/day, after adjustment for total wear time and other variables. In model 2, after controlling for LPA and MVPA time, STS score did not show significant association with SB; however, FOF score was significantly positively associated with more time spent in SB (β=8.99, *P*=.01).

Results from multiple regression analysis for LPA (minutes/day) are reported in Table 3. In model 1 (adjusted for total wear time), STS performance had a significant positive association (*P*=.04) with LPA, indicating an increase of 2.39 minutes/day of LPA time for each 1-unit increase in the STS score. However, in model 2 (adjusted for SB and MVPA), no significant associations were observed between LPA and any independent variable.

Table 4 presents the back-transformed regression coefficients for MVPA models. In model 1, STS performance was significantly positively associated (*P*=.001) with MVPA, indicating a 5.13% change in MVPA time (minutes/day) for each 1-unit increase in the STS score. In model 2, after adjusting for SB and LPA time, the self-rated health score (β=20.92, *P*=.04) and STS performance (β=34.99, *P*=.02) showed a significant positive association with MVPA time.

Multimedia Appendix 1 reports the correlation matrix between all independent variables used in the linear regression analysis. All correlation coefficients were less than 0.8; therefore, no multicollinearity was detected.

Table 2. Association with SB[a] (minutes/day): results from multiple regression analysis.

Participant characteristics	Model 1[b]		Model 2[c]	
	β[d] (SE)	*P* value	β (SE)	*P* value
Age (years)	0.60 (1.14)	.60	–0.99 (1.84)	.59
Household composition (Ref.[e]: living alone)				
Living with family	–12.27 (16.42)	.46	18.82 (26.36)	.48
Self-rated health score	–0.37 (10.67)	.97	8.42 (17.35)	.63
Frailty status (Ref.: prefrail or frail)				
Robust health	–11.02 (16.78)	.51	–25.02 (26.62)	.35
FOF[f] score	2.52 (2.32)	.28	8.99 (3.57)	.01
STS[g] score (reps)	–3.92 (1.44)	.01	0.25 (2.42)	.92

[a]SB: sedentary behavior.

[b]Adjusted for the BMI and total wear time (minutes/day).

[c]Adjusted for the BMI, light-intensity physical activity (LPA; minutes/day), and moderate-to-vigorous-intensity physical activity (MVPA; minutes/day).

[d]β: standardized regression coefficient.

[e]Ref.: reference.

[f]FOF: fear of falling.

[g]STS: sit-to-stand.

XSL•FO
RenderX

Table 3. Association with LPA[a] (minutes/day): results from multiple regression analysis.

Participant characteristics	Model 1[b]		Model 2[c]	
	β[d] (SE)	*P* value	β (SE)	*P* value
Age (years)	–0.13 (0.92)	.89	0.53 (0.83)	.53
Household composition (Ref.[e]: living alone)				
Living with family	11.12 (13.20)	.40	2.74 (11.93)	.82
Self-rated health score	–8.03 (8.58)	.35	–12.60 (7.72)	.11
Frailty status (Ref.: prefrail or frail)				
Robust health	–12.72 (13.49)	.35	9.08 (12.04)	.45
FOF[f] score	–2.12 (1.87)	.26	1.28 (1.67)	.44
STS[g] score (reps)	2.39 (1.16)	.04	0.50 (1.09)	.65

[a]LPA: light-intensity physical activity.

[b]Adjusted for the BMI and total wear time (minutes/day).

[c]Adjusted for the BMI, sedentary behavior (SB; minutes/day), and moderate-to-vigorous-intensity physical activity (MVPA; minutes/day).

[d]β: standardized regression coefficient.

[e]Ref.: reference.

[f]FOF: fear of falling.

[g]STS: sit-to-stand.

Table 4. Association with MVPA[a] (minutes/day): results from multiple regression analysis.

Participant characteristics	Model 1[b]		Model 2[c]	
	β[d] (SE)	*P* value	β (SE)	*P* value
Age (years)	–1.98 (1.00)	.10	–1.78 (0.95)	.06
Household composition (Ref.[e]: living alone)				
Living with family	31.0 (17.35)	.11	22.14 (15.03)	.14
Self-rated health score	15.03 (11.63)	.20	20.92 (9.42)	.04
Frailty status (Ref.: prefrail or frail)				
Robust health	–1.00 (18.53)	.94	–10.42 (15.02)	.43
FOF[f] score	–3.92 (2.02)	.11	–1.39 (1.92)	.46
STS[g] score (reps)	5.13 (1.00)	.001	34.99 (1.21)	.02

[a]MVPA: moderate-to-vigorous-intensity physical activity.

[b]Adjusted for the BMI and total wear time (minutes/day).

[c]Adjusted for the BMI, sedentary behavior (SB; minutes/day), and light-intensity physical activity (LPA; minutes/day).

[d]β: standardized regression coefficient.

[e]Ref.: reference.

[f]FOF: fear of falling.

[g]STS: sit-to-stand.

Discussion

Principal Findings

This is the first study, to the best of our knowledge, to report the factors associated with objectively measured PA levels among a diverse sample of older US women during the COVID-19 pandemic. Our study findings indicate that participants spent the majority of their day being sedentary, confirming the high prevalence of a sedentary lifestyle among older adults reported in the literature [44]. Evidence suggests that a higher level of sedentary time remains associated with greater all-cause mortality risk among older adults, even among those who meet the national MVPA guidelines of 150 minutes/week [45]. We also observed that participants accumulated more time in LPA compared to MVPA, which confirms that LPA is the predominant form of PA behavior among women aged 60 years or above and accounts for a large portion of their daily activities [34]. Therefore, to combat a

XSL•FO
RenderX

sedentary lifestyle, replacing SB with LPA (rather than MVPA) could be a more achievable target for older women, particularly those with chronic conditions and low cardiorespiratory fitness. Studies have reported LPA to be associated with reduced mortality risk, more favorable cardiometabolic biomarkers, and reduced incident mobility disorders in older adults [46-48]. Therefore, future studies should focus on identifying the optimal amount of LPA that may elicit health benefits in older women, irrespective of their engagement in MVPA, and developing LPA recommendations for the heterogeneous older adult population.

Compared to the previously published literature reporting normative PA levels of older US women in the prepandemic period using accelerometry-based analysis [49], our study participants' PA levels were observed to be lower. Evenson et al [49] reported the average sedentary time in older US women, measured by hip-worn accelerometers, to be 510.6 (SD 98.8) minutes/day in the Women's Health Study (WHS) cohort (n=16,726, mean age 71.5 years, SD 5.7 years, age range 62-89 years) and 555.6 (SD 99.4) minutes/day in the Women's Health Initiative/Objective Physical Activity and Cardiovascular Health (WHI/OPACH) cohort (n=6126, mean age 78.7 years, SD 6.7 years, age range 63-97 years), which is less than the sedentary time among our participants (ie, mean 744, SD 114 minutes/day). Similarly, the average time accumulated in LPA (ie, mean 218.6, SD 64.3 minutes/day) and MVPA (ie, mean 42.4, SD 31.0 minutes/day) in our study sample was found to be less than the prepandemic PA norms reported in the WHS (LPA: mean 287.8, SD 54.63 minutes/day, MVPA: mean 91.9, SD 45.4 minutes/day) and WHI/OPACH cohorts (LPA: mean 286.9, SD 61.48 minutes/day, MVPA: mean 50.4, SD 34.4 minutes/day). These findings might qualitatively reflect the trend in PA and SB among older US women before and during the pandemic, despite the differences in accelerometer placement and PA intensity cut-off points. Previously, several studies have reported the negative impact of the pandemic on PA levels of older adults across different geographical regions during various phases of the pandemic [10]. Our results imply that despite the easing of stay-at-home orders in the United States, this trend of reduced PA participation among older women persisted during the second year of the pandemic, due to their adherence to the physical distancing guidelines. Previous findings suggest that physical distancing poses a risk of diminished social connectedness and can disproportionately impact older adults whose social interactions used to take place mostly outside the home (eg, community centers, volunteering services, places of worship), affecting their habitual PA level [50]. However, in this study, we were unable to quantify the longitudinal changes in PA behavior among our participants, since we did not have their free-living PA data from the prepandemic period.

This study also investigated the factors affecting the objectively measured PA and SB among older women during the pandemic. Our findings indicate that the time spent in SB increased with older age in our study sample. In general, older adults spend more time being sedentary than any other age groups [51], and our results support that even among older adults, older age is associated with a more sedentary lifestyle [52]. Prior studies have reported that age-related physiological and functional declines, as well as the prevalence of chronic diseases, may limit one's ability to participate in MVPA in older age [53,54]. Therefore, it is important to identify the barriers to PA participation (eg, poor health, lack of knowledge, lack of motivation) in older women, particularly those aged 80 years or above, and provide individual-specific MVPA recommendations based on their aerobic capacity. For older women who cannot achieve 150 minutes/week of MVPA due to poor health and limited functional capacity, PA interventions incorporating both LPA and MVPA might provide a more feasible and sustainable approach in maintaining an active lifestyle [52].

In our study sample, lower MVPA time was observed in participants living alone compared to those living with family. During the pandemic, older women who live alone have been more susceptible to social isolation than in the prepandemic time, and the associations between social isolation and lower self-reported MVPA among older adults have been noted previously [55]. In addition, a recent study in Japan reported that during the third wave of the pandemic (ie, January 2021), recovery in the total PA time since the first wave (ie, April 2020) was observed among most Japanese older adults, except those who were living alone and were socially inactive [56]. This indicates that living with family might contribute to better resilience against the negative impact of the pandemic on the PA behavior of older women, because they are more likely to obtain valuable knowledge, support, and motivation from family members for maintaining a healthy lifestyle. In addition, this highlights the need for increasing social support for older women living alone in order to effectively promote MVPA participation among them.

Our findings suggest that better self-rated health was positively associated with less SB and more MVPA engagement in our study sample. Previously, existing studies have identified the bidirectional associations of SB and MVPA with self-rated health [57-59]. A study on middle-aged US adults reported that poor self-rated health is linked with adverse longitudinal shifts toward a more sedentary lifestyle and less MVPA time [58]. Conversely, Beyer et al [59] found that older individuals with positive self-perceptions of aging are more likely to participate in PA, which, in turn, improves their self-rated health. Therefore, PA intervention programs for older women should foster positive self-perceptions of aging, in conjunction with healthy lifestyle behaviors (eg, proper nutrition and diet intake, adequate sleep, no/reduced smoking and alcohol intake) for enhancing their PA participation so that they can achieve favorable self-rated health in later life.

Regarding frailty, our finding is consistent with the previous evidence that irrespective of MVPA participation, a higher level of SB is associated with the increased odds of being frail or prefrail in older age [60-62]. This also emphasizes the need for developing targeted interventions to reduce sedentary time among prefrail and frail older women, with a particular focus on decelerating or possibly preventing further functional loss in prefrail individuals.

FOF causes older adults to limit their habitual PA level, which, in turn, may increase their risk of falling more [23], and our

XSL•FO
RenderX

result is in agreement with the existing literature [63,64]. Therefore, it is important to identify the barriers (both physical and psychological) to PA participation in older women with high FOF so that tailored interventions can be developed for those having irrational FOF despite having a low physiological risk of fall [65]. In addition, to reduce sedentary time in older women with high FOF and a high physiological risk of fall, focus should be given on integrated intervention approaches combining both cognitive behavioral therapy and balance exercises.

Our key findings indicate that STS performance was independently associated with all 3 PA variables (ie, SB, LPA, and MVPA time) in the regression analysis after adjusting for total wear time. The 30-second STS test is a widely used, well-validated functional performance measure in clinical research and practice, having good test-retest and interrater reliability [66]. STS performance is considered an indicator of lower limb strength among older adults and has been correlated with objective strength testing methods, such as leg-press resistance [67] and power rigs [68]. In our study, participants with below-average STS scores showed reduced PA levels compared to those with average and above-average scores. Previous findings have reported the bidirectional associations between SB and PA with lower body muscle strength [69], which is in agreement with our findings. In addition to lower body strength, STS scores have also been associated with dynamic balance and mobility [70] and are considered a proxy measure for physical performance in sarcopenia diagnosis [66]. These highlight the need for developing home-based multimodal intervention strategies during the pandemic to promote PA participation among older women, which will include (1) strength training for improving lower limb muscle mass and strength and (2) balance exercises for reducing the risk of falls.

Our focus was to identify the factors that were associated with higher sedentary time and less PA participation in our study sample during the pandemic. Based on our findings, considerations should be taken about an older individual's age and health status and whether the person lives alone, is frail, and has high FOF and poor STS performance while providing PA prescriptions. For instance, if an older person has high FOF but good dynamic balance (ie, STS performance), the study result informs that the PA intervention should integrate approaches to reduce this irrational FOF for promoting PA participation. Again, if an older woman is more than 80 years old and has multiple chronic diseases, then PA intervention focusing on increasing LPA might be more effective and feasible (rather than increasing MVPA) to combat the sedentary lifestyle. These examples indicate how this knowledge of different factors associated with PA participation can contribute to individually tailored PA prescriptions for older women, rather than a one-size-fits-all approach, even during their transition to the postpandemic lifestyle.

Strengths and Limitations

A strength of our study is the accelerometry-based measurement of the PA level during the pandemic, providing an objective and detailed description of SB, LPA, and MVPA patterns among women aged 60 years or above. In addition, identifying the factors associated with PA behavior provided evidence to develop informed strategies for maintaining or improving PA participation among older women in the context of the pandemic.

There are some recognized limitations of wrist-worn accelerometry-based studies, which apply to our study as well. For instance, wrist-worn accelerometers cannot accurately and reliably detect nonambulatory activities, such as resistance training or cycling. In addition, in some cases, a wrist-worn accelerometer can overestimate the PA level of the user while they perform activities that are primarily upper limb movements with low energy expenditure (eg, cleaning or sewing in a seated position). Furthermore, the cut-off points to classify PA intensity for wrist-worn accelerometers for older adults have not been firmly established yet. The nondominant wrist ENMO cut-off points for older adults, reported in the existing literature, range from 18 to 57 mg for the LPA threshold and 60 to 104 mg for the MVPA threshold [71-73], which limits the comparability of results among studies with different cut-off points. Moreover, due to the cross-sectional study design, we were not able quantify the change in PA behavior in our participants between the prepandemic period and the pandemic time, since we did not have their objectively measured pre–COVID-19 PA data. Another limitation of this study was the small, nonrepresentative nature of the sample. This sample was predominantly White, educated, relatively healthy, and active (75.6% meeting the national MVPA guidelines), which limits the generalizability of our findings.

Conclusion

This study investigated objectively measured SB and PA in a sample of older US women during the COVID-19 pandemic. When compared to pre–COVID-19 norms of older US women, it was observed that the total time spent in LPA and MVPA was lower during the pandemic, while the average sedentary time was higher. A more sedentary lifestyle was found in participants who were aged 80 years or above, had poorer self-rated health, were frail or prefrail, and had high FOF. The time spent in LPA was significantly lower among women aged 80 years or above. Participation in MVPA was higher for those who were aged 60-79 years, lived with family, had better self-rated health, and had low FOF. In addition, it was observed that STS performance was independently associated with increased PA levels among participants after adjustment for total accelerometer wear time. These findings can help design more sustainable and behavior-changing PA interventions for older women to promote healthy aging and mitigate long-term health consequences of the pandemic.

XSL•FO
RenderX

Acknowledgments

This work was supported by the National Institute on Aging (NIA), National Institutes of Health (NIH; grant no. 1R03AG06799). This work was also supported by the Learning Institute for Elders at the University of Central Florida. The funding body was not involved in the study design, data collection, data analysis, or writing of the manuscript.

Authors' Contributions

JP, LT, RX, and JRS contributed to the funding acquisition, conceptualization, and study design. RC, JP, and LT coordinated the study and collected data. RC, JP, and RX contributed to data analysis. RC wrote the original draft, and JP, LT, RX, and JRS contributed to the revision of the original draft. All authors have agreed to the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Correlation matrix of variables used in the regression analysis.
[DOCX File , 25 KB - aging_v5i3e38172_app1.docx]

References

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. URL: https://covid19.who.int [accessed 2022-03-20]
2. Shahid Z, Kalayanamitra R, McClafferty B, Kepko D, Ramgobin D, Patel R, et al. COVID-19 and older adults: what we know. J Am Geriatr Soc 2020 May 20;68(5):926-929 [FREE Full text] [doi: 10.1111/jgs.16472] [Medline: 32255507]
3. Callow DD, Arnold-Nedimala NA, Jordan LS, Pena GS, Won J, Woodard JL, et al. The mental health benefits of physical activity in older adults survive the COVID-19 pandemic. Am J Geriatr Psychiatry 2020 Oct 30;28(10):1046-1057 [FREE Full text] [doi: 10.1016/j.jagp.2020.06.024] [Medline: 32713754]
4. Wegner L, Mendoza-Vasconez AS, Mackey S, McGuire V, To C, White B, et al. Physical activity, well-being, and priorities of older women during the COVID-19 pandemic: a survey of Women's Health Initiative Strong and Healthy (WHISH) intervention participants. Transl Behav Med 2021 Dec 14;11(12):2155-2163 [FREE Full text] [doi: 10.1093/tbm/ibab122] [Medline: 34633465]
5. Carriedo A, Cecchini JA, Fernandez-Rio J, Méndez-Giménez A. COVID-19, psychological well-being and physical activity levels in older adults during the nationwide lockdown in Spain. Am J Geriatr Psychiatry 2020 Nov;28(11):1146-1155 [FREE Full text] [doi: 10.1016/j.jagp.2020.08.007] [Medline: 32919872]
6. Suzuki Y, Maeda N, Hirado D, Shirakawa T, Urabe Y. Physical activity changes and its risk factors among community-dwelling Japanese older adults during the COVID-19 epidemic: associations with subjective well-being and health-related quality of life. Int J Environ Res Public Health 2020 Sep 10;17(18):6591 [FREE Full text] [doi: 10.3390/ijerph17186591] [Medline: 32927829]
7. Bailey L, Ward M, DiCosimo A, Baunta S, Cunningham C, Romero-Ortuno R, et al. Physical and mental health of older people while cocooning during the COVID-19 pandemic. QJM 2021 Nov 13;114(9):648-653 [FREE Full text] [doi: 10.1093/qjmed/hcab015] [Medline: 33471128]
8. Wilke J, Mohr L, Tenforde AS, Edouard P, Fossati C, González-Gross M, et al. A pandemic within the pandemic? Physical activity levels substantially decreased in countries affected by COVID-19. Int J Environ Res Public Health 2021 Feb 24;18(5):2235 [FREE Full text] [doi: 10.3390/ijerph18052235] [Medline: 33668262]
9. Cheval B, Sivaramakrishnan H, Maltagliati S, Fessler L, Forestier C, Sarrazin P, et al. Relationships between changes in self-reported physical activity, sedentary behaviour and health during the coronavirus (COVID-19) pandemic in France and Switzerland. J Sports Sci 2021 Mar 29;39(6):699-704. [doi: 10.1080/02640414.2020.1841396] [Medline: 33118469]
10. Oliveira MR, Sudati IP, Konzen VDM, de Campos AC, Wibelinger LM, Correa C, et al. Covid-19 and the impact on the physical activity level of elderly people: a systematic review. Exp Gerontol 2022 Mar;159:111675 [FREE Full text] [doi: 10.1016/j.exger.2021.111675] [Medline: 34954282]
11. Marasso D, Lupo C, Collura S, Rainoldi A, Brustio PR. Subjective versus objective measure of physical activity: a systematic review and meta-analysis of the convergent validity of the Physical Activity Questionnaire for Children (PAQ-C). Int J Environ Res Public Health 2021 Mar 25;18(7):3413 [FREE Full text] [doi: 10.3390/ijerph18073413] [Medline: 33806106]
12. Strath SJ, Kaminsky LA, Ainsworth BE, Ekelund U, Freedson PS, Gary RA, Committee on Physical Activity of the American Heart Association Council on Lifestyle and Cardiometabolic Health. Guide to the assessment of physical activity: clinical and research applications: a scientific statement from the American Heart Association. Circulation 2013 Nov 12;128(20):2259-2279 [FREE Full text] [doi: 10.1161/01.cir.0000435708.67487.da] [Medline: 24126387]
13. Browne RA, Macêdo GAD, Cabral LL, Oliveira GT, Vivas A, Fontes EB, et al. Initial impact of the COVID-19 pandemic on physical activity and sedentary behavior in hypertensive older adults: an accelerometer-based analysis. Exp Gerontol 2020 Dec;142:111121 [FREE Full text] [doi: 10.1016/j.exger.2020.111121] [Medline: 33132145]

XSL•FO
RenderX

14. Leavy B, Hagströmer M, Conradsson DM, Franzén E. Physical activity and perceived health in people with Parkinson disease during the first wave of COVID-19 pandemic: a cross-sectional study from Sweden. J Neurol Phys Ther 2021 Oct 01;45(4):266-272 [FREE Full text] [doi: 10.1097/NPT.0000000000000372] [Medline: 34369451]
15. Miyahara S, Tanikawa Y, Hirai H, Togashi S. Impact of the state of emergency enacted due to the COVID-19 pandemic on the physical activity of the elderly in Japan. J Phys Ther Sci 2021 Apr;33(4):345-350 [FREE Full text] [doi: 10.1589/jpts.33.345] [Medline: 33935359]
16. Yamada Y, Yoshida T, Nakagata T, Nanri H, Miyachi M. Letter to the editor: age, sex, and regional differences in the effect of COVID-19 pandemic on objective physical activity in Japan: a 2-year nationwide longitudinal study. J Nutr Health Aging 2021 Jul 30;25(8):1032-1033 [FREE Full text] [doi: 10.1007/s12603-021-1662-y] [Medline: 34545926]
17. Prohaska T, Belansky E, Belza B, Buchner D, Marshall V, McTigue K, et al. Physical activity, public health, and aging: critical issues and research priorities. J Gerontol B Psychol Sci Soc Sci 2006 Sep;61(5):S267-S273. [doi: 10.1093/geronb/61.5.s267] [Medline: 16960240]
18. Zbrońska I, Mędrela-Kuder E. The level of physical activity in elderly persons with overweight and obesity. Rocz Panstw Zakl Hig 2018;69(4):369-373 [FREE Full text] [Medline: 30525327]
19. Barreira T, Harrington D, Katzmarzyk P. Cardiovascular health metrics and accelerometer-measured physical activity levels: National Health and Nutrition Examination Survey, 2003-2006. Mayo Clin Proc 2014 Jan;89(1):81-86. [doi: 10.1016/j.mayocp.2013.10.001] [Medline: 24388025]
20. Tolley AP, Ramsey KA, Rojer AG, Reijnierse EM, Maier AB. Objectively measured physical activity is associated with frailty in community-dwelling older adults: a systematic review. J Clin Epidemiol 2021 Sep;137:218-230 [FREE Full text] [doi: 10.1016/j.jclinepi.2021.04.009] [Medline: 33915264]
21. Mattioli R, Cavalli AS, Ribeiro JAB, Silva MCD. Association between handgrip strength and physical activity in hypertensive elderly individuals. Rev Bras Geriatr Gerontol 2015 Dec;18(4):881-891. [doi: 10.1590/1809-9823.2015.14178]
22. Rosenberg DE, Bellettiere J, Gardiner PA, Villarreal VN, Crist K, Kerr J. Independent associations between sedentary behaviors and mental, cognitive, physical, and functional health among older adults in retirement communities. J Gerontol A Biol Sci Med Sci 2016 Jan 13;71(1):78-83 [FREE Full text] [doi: 10.1093/gerona/glv103] [Medline: 26273024]
23. Jefferis BJ, Iliffe S, Kendrick D, Kerse N, Trost S, Lennon LT, et al. How are falls and fear of falling associated with objectively measured physical activity in a cohort of community-dwelling older men? BMC Geriatr 2014 Oct 27;14(1):114 [FREE Full text] [doi: 10.1186/1471-2318-14-114] [Medline: 25348492]
24. Jarach C, Tettamanti M, Nobili A, D'avanzo B. Social isolation and loneliness as related to progression and reversion of frailty in the Survey of Health Aging Retirement in Europe (SHARE). Age Ageing 2021 Jan 08;50(1):258-262 [FREE Full text] [doi: 10.1093/ageing/afaa168] [Medline: 32915990]
25. Pizano-Escalante MG, Anaya-Esparza LM, Nuño K, Rodríguez-Romero J, Gonzalez-Torres S, López-de la Mora DA, et al. Direct and indirect effects of COVID-19 in frail elderly: interventions and recommendations. J Pers Med 2021 Oct 02;11(10):999 [FREE Full text] [doi: 10.3390/jpm11100999] [Medline: 34683141]
26. Nakamura M, Imaoka M, Nakao H, Hida M, Imai R, Tazaki F, et al. Increased anxiety about falls and walking ability among community-dwelling Japanese older adults during the COVID-19 pandemic. Psychogeriatrics 2021 Sep 07;21(5):826-831 [FREE Full text] [doi: 10.1111/psyg.12750] [Medline: 34363431]
27. Del Pozo Cruz B, Perales F, Alfonso-Rosa RM, Del Pozo-Cruz J. Impact of social isolation on physical functioning among older adults: a 9-year longitudinal study of a U.S.-representative sample. Am J Prev Med 2021 Aug;61(2):158-164. [doi: 10.1016/j.amepre.2021.02.003] [Medline: 33849775]
28. Migueles J, Rowlands A, Huber F, Sabia S, van HV. GGIR: a research community–driven open source R package for generating physical activity and sleep outcomes from multi-day raw accelerometer data. JMPB 2019;2(3):188-196. [doi: 10.1123/jmpb.2018-0063]
29. van Hees VT, Fang Z, Langford J, Assah F, Mohammad A, da Silva ICM, et al. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. J Appl Physiol (1985) 2014 Oct 01;117(7):738-744 [FREE Full text] [doi: 10.1152/japplphysiol.00421.2014] [Medline: 25103964]
30. van Hees VT, Gorzelniak L, Dean León EC, Eder M, Pias M, Taherian S, et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. PLoS One 2013 Apr 23;8(4):e61691 [FREE Full text] [doi: 10.1371/journal.pone.0061691] [Medline: 23626718]
31. Bakrania K, Yates T, Rowlands AV, Esliger DW, Bunnewell S, Sanders J, et al. Intensity thresholds on raw acceleration data: Euclidean Norm Minus One (ENMO) and mean amplitude deviation (MAD) approaches. PLoS One 2016 Oct 5;11(10):e0164045 [FREE Full text] [doi: 10.1371/journal.pone.0164045] [Medline: 27706241]
32. Suorsa K, Pulakka A, Leskinen T, Pentti J, Holtermann A, Heinonen O. Comparison of sedentary time between thigh-worn and wrist-worn accelerometers. JMPB 2020;3(3):234-243. [doi: 10.1123/jmpb.2019-0052]
33. Menai M, van Hees VT, Elbaz A, Kivimaki M, Singh-Manoux A, Sabia S. Accelerometer assessed moderate-to-vigorous physical activity and successful ageing: results from the Whitehall II study. Sci Rep 2017 Apr 03;8(1):45772 [FREE Full text] [doi: 10.1038/srep45772] [Medline: 28367987]
34. LaCroix AZ, Bellettiere J, Rillamas-Sun E, Di C, Evenson KR, Lewis CE, Women's Health Initiative (WHI). Association of Light Physical Activity Measured by Accelerometry and Incidence of Coronary Heart Disease and Cardiovascular

XSL•FO
RenderX

Disease in Older Women. JAMA Netw Open 2019 Mar 01;2(3):e190419 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.0419] [Medline: 30874775]
35. Lee IM, Shiroma EJ, Evenson KR, Kamada M, LaCroix AZ, Buring JE. Using Devices to Assess Physical Activity and Sedentary Behavior in a Large Cohort Study, the Women's Health Study. J Meas Phys Behav 2018 Jun 01;1(2):60-69 [FREE Full text] [doi: 10.1123/jmpb.2018-0005] [Medline: 30666321]
36. Nuttall F. Body mass index: obesity, BMI, and health: a critical review. Nutr Today 2015;50(3):117. [doi: 10.1097/nt.0000000000000092]
37. Morley JE, Malmstrom TK, Miller DK. A simple frailty questionnaire (FRAIL) predicts outcomes in middle aged African Americans. J Nutr Health Aging 2012 Jul 11;16(7):601-608 [FREE Full text] [doi: 10.1007/s12603-012-0084-2] [Medline: 22836700]
38. Kempen GIJM, Yardley L, van Haastregt JCM, Zijlstra GAR, Beyer N, Hauer K, et al. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. Age Ageing 2008 Jan 11;37(1):45-50. [doi: 10.1093/ageing/afm157] [Medline: 18032400]
39. Thiamwong L, Stout JR, Park J, Yan X. Technology-based fall risk assessments for older adults in low-income settings: protocol for a cross-sectional study. JMIR Res Protoc 2021 Apr 07;10(4):e27381 [FREE Full text] [doi: 10.2196/27381] [Medline: 33825688]
40. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, Writing Group for the European Working Group on Sarcopenia in Older People 2 (EWGSOP2), the Extended Group for EWGSOP2. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019 Jan 01;48(1):16-31 [FREE Full text] [doi: 10.1093/ageing/afy169] [Medline: 30312372]
41. Rikli R, Jones C. Functional fitness normative scores for community-residing older adults, ages 60-94. J Aging Phys Act 1999;7(2):162-181. [doi: 10.1123/japa.7.2.162]
42. Ahorsu DK, Lin C, Imani V, Saffari M, Griffiths MD, Pakpour AH. The Fear of COVID-19 Scale: development and initial validation. Int J Ment Health Addict 2020 Mar 27:1-9 [FREE Full text] [doi: 10.1007/s11469-020-00270-8] [Medline: 32226353]
43. Singer J, Willett J, Willett J. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford, UK: Oxford University Press; 2003.
44. Gennuso KP, Gangnon RE, Matthews CE, Thraen-Borowski KM, Colbert LH. Sedentary behavior, physical activity, and markers of health in older adults. Med Sci Sports Exerc 2013 Aug;45(8):1493-1500 [FREE Full text] [doi: 10.1249/MSS.0b013e318288a1e5] [Medline: 23475142]
45. Chastin S, Gardiner P, Harvey JA, Leask CF, Jerez-Roig J, Rosenberg D, et al. Interventions for reducing sedentary behaviour in community-dwelling older adults. Cochrane Database Syst Rev 2021 Jun 25;6(6):CD012784 [FREE Full text] [doi: 10.1002/14651858.CD012784.pub2] [Medline: 34169503]
46. LaMonte MJ, Buchner DM, Rillamas-Sun E, Di C, Evenson KR, Bellettiere J, et al. Accelerometer-measured physical activity and mortality in women aged 63 to 99. J Am Geriatr Soc 2018 May 16;66(5):886-894 [FREE Full text] [doi: 10.1111/jgs.15201] [Medline: 29143320]
47. Healy G, Dunstan D, Salmon J, Cerin E, Shaw J, Zimmet P, et al. Objectively measured light-intensity physical activity is independently associated with 2-h plasma glucose. Diabetes Care 2007 Jun;30(6):1384-1389. [doi: 10.2337/dc07-0114] [Medline: 17473059]
48. Glass NL, Bellettiere J, Jain P, LaMonte MJ, LaCroix AZ, Women's Health Initiative. Evaluation of light physical activity measured by accelerometry and mobility disability during a 6-year follow-up in older women. JAMA Netw Open 2021 Feb 01;4(2):e210005 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.0005] [Medline: 33620446]
49. Evenson KR, Bellettiere J, Cuthbertson CC, Di C, Dushkes R, Howard AG, et al. Cohort profile: the Women's Health Accelerometry Collaboration. BMJ Open 2021 Nov 29;11(11):e052038 [FREE Full text] [doi: 10.1136/bmjopen-2021-052038] [Medline: 34845070]
50. Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly. Lancet Public Health 2020 May;5(5):e256. [doi: 10.1016/s2468-2667(20)30061-x]
51. Dogra S, Ashe MC, Biddle SJH, Brown WJ, Buman MP, Chastin S, et al. Sedentary time in older men and women: an international consensus statement and research priorities. Br J Sports Med 2017 Nov 19;51(21):1526-1532 [FREE Full text] [doi: 10.1136/bjsports-2016-097209] [Medline: 28724710]
52. Diaz K, Howard V, Hutto B, Colabianchi N, Vena J, Blair S. Patterns of sedentary behavior in US middle-age and older adults: the REGARDS study. Med Sci Sports Exerc 2016;48(3):430. [doi: 10.1249/mss.0000000000000792]
53. Lee P, Cigolle C, Ha J, Min L, Murphy S, Blaum C, et al. Physical function limitations among middle-aged and older adults with prediabetes: one exercise prescription may not fit all. Diabetes Care 2013 Oct;36(10):3076-3083 [FREE Full text] [doi: 10.2337/dc13-0412] [Medline: 23757432]
54. Volders E, de Groot RHM, Bolman CAW, Lechner L. The longitudinal associations between change in physical activity and cognitive functioning in older adults with chronic illness (es). BMC Geriatr 2021 Sep 04;21(1):478 [FREE Full text] [doi: 10.1186/s12877-021-02429-x] [Medline: 34481465]

XSL•FO
RenderX

55. Kobayashi LC, Steptoe A. Social isolation, loneliness, and health behaviors at older ages: longitudinal cohort study. Ann Behav Med 2018 May 31;52(7):582-593 [FREE Full text] [doi: 10.1093/abm/kax033] [Medline: 29860361]
56. Yamada M, Kimura Y, Ishiyama D, Otobe Y, Suzuki M, Koyama S, et al. The influence of the COVID-19 pandemic on physical activity and new incidence of frailty among initially non-frail older adults in Japan: a follow-up online survey. J Nutr Health Aging 2021 Apr 26;25(6):751-756 [FREE Full text] [doi: 10.1007/s12603-021-1634-2] [Medline: 34179929]
57. Gregg E, Kriska A, Fox K, Cauley J. Self-rated health and the spectrum of physical activity and physical function in older women. J Aging Phys Act 1996;4(4):349-361. [doi: 10.1123/japa.4.4.349]
58. Gibbs BB, Sternfeld B, Whitaker KM, Brach JS, Hergenroeder AL, Jacobs DR, et al. Bidirectional associations of accelerometer-derived physical activity and stationary behavior with self-reported mental and physical health during midlife. Int J Behav Nutr Phys Act 2021 Jun 06;18(1):74 [FREE Full text] [doi: 10.1186/s12966-021-01145-4] [Medline: 34090471]
59. Beyer A, Wolff JK, Warner LM, Schüz B, Wurm S. The role of physical activity in the relationship between self-perceptions of ageing and self-rated health in older adults. Psychol Health 2015 Feb 27;30(6):671-685. [doi: 10.1080/08870446.2015.1014370] [Medline: 25720739]
60. Blodgett J, Theou O, Kirkland S, Andreou P, Rockwood K. The association between sedentary behaviour, moderate-vigorous physical activity and frailty in NHANES cohorts. Maturitas 2015 Feb;80(2):187-191. [doi: 10.1016/j.maturitas.2014.11.010] [Medline: 25542406]
61. da Silva Coqueiro R, de Queiroz BM, Oliveira DS, das Merces MC, Oliveira Carneiro JA, Pereira R, et al. Cross-sectional relationships between sedentary behavior and frailty in older adults. J Sports Med Phys Fitness 2017 Jun;57(6):825-830. [doi: 10.23736/s0022-4707.16.06289-7]
62. Santos ISD, Silva CDFR, Ohara DG, Matos AP, Pinto ACPN, Pegorari MS. Association between frailty syndrome and sedentary behavior among community-dwelling older adults in the Amazon region: a cross-sectional study. Sao Paulo Med J 2021 Jun;139(3):226-233. [doi: 10.1590/1516-3180.2020.0546.r1.14122020]
63. Mendes da Costa E, Pepersack T, Godin I, Bantuelle M, Petit B, Levêque A. Fear of falling and associated activity restriction in older people. Results of a cross-sectional study conducted in a Belgian town. Arch Public Health 2012 Jan 03;70(1):1-8 [FREE Full text] [doi: 10.1186/0778-7367-70-1] [Medline: 22958732]
64. Zijlstra GAR, van Haastregt JCM, van Eijk JTM, van Rossum E, Stalenhoef PA, Kempen GIJM. Prevalence and correlates of fear of falling, and associated avoidance of activity in the general population of community-living older people. Age Ageing 2007 May 22;36(3):304-309. [doi: 10.1093/ageing/afm021] [Medline: 17379605]
65. Thiamwong L, Ng BP, Kwan RYC, Suwanno J. Maladaptive fall risk appraisal and falling in community-dwelling adults aged 60 and older: implications for screening. Clin Gerontol 2021 Jul 08;44(5):552-561. [doi: 10.1080/07317115.2021.1950254] [Medline: 34233599]
66. Yee XS, Ng YS, Allen JC, Latib A, Tay EL, Abu Bakar HM, et al. Performance on sit-to-stand tests in relation to measures of functional fitness and sarcopenia diagnosis in community-dwelling older adults. Eur Rev Aging Phys Act 2021 Jan 08;18(1):1 [FREE Full text] [doi: 10.1186/s11556-020-00255-5] [Medline: 33419399]
67. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999 Jun;70(2):113-119. [doi: 10.1080/02701367.1999.10608028] [Medline: 10380242]
68. Hardy R, Cooper R, Shah I, Harridge S, Guralnik J, Kuh D. Is chair rise performance a useful measure of leg power? Aging Clin Exp Res 2013 Jul 26;22(5-6):412-418. [doi: 10.1007/bf03324942]
69. Buckinx F, Peyrusqué É, Granet J, Aubertin-Leheudre M. Impact of current or past physical activity level on functional capacities and body composition among elderly people: a cross-sectional analysis from the YMCA study. Arch Public Health 2021 Apr 15;79(1):50 [FREE Full text] [doi: 10.1186/s13690-021-00573-9] [Medline: 33858506]
70. Lord SR, Murray SM, Chapman K, Munro B, Tiedemann A. Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. J Gerontol A Biol Sci Med Sci 2002 Aug 01;57(8):M539-M543. [doi: 10.1093/gerona/57.8.m539] [Medline: 12145369]
71. Sanders GJ, Boddy LM, Sparks SA, Curry WB, Roe B, Kaehne A, et al. Evaluation of wrist and hip sedentary behaviour and moderate-to-vigorous physical activity raw acceleration cutpoints in older adults. J Sports Sci 2019 Jun 17;37(11):1270-1279. [doi: 10.1080/02640414.2018.1555904] [Medline: 30558487]
72. Migueles JH, Cadenas-Sanchez C, Alcantara JMA, Leal-Martín J, Mañas A, Ara I, et al. Calibration and cross-validation of accelerometer cut-points to classify sedentary time and physical activity from hip and non-dominant and dominant wrists in older adults. Sensors (Basel) 2021 May 11;21(10):3326 [FREE Full text] [doi: 10.3390/s21103326] [Medline: 34064790]
73. Fraysse F, Post D, Eston R, Kasai D, Rowlands AV, Parfitt G. Physical activity intensity cut-points for wrist-worn GENEActiv in older adults. Front Sports Act Living 2020 Jan 15;2:579278 [FREE Full text] [doi: 10.3389/fspor.2020.579278] [Medline: 33521631]

Abbreviations

CDC: Centers for Disease Control and Prevention
ENMO: Euclidean Norm Minus 1 g
FCV-19S: Fear of COVID-19 Scale

XSL•FO
RenderX

FES-I: Falls Efficacy Scale-International
FOF: fear of falling
FRAIL: Fatigue, Resistance, Ambulation, Illnesses, and Loss of weight
g: gravitational unit
LPA: light-intensity physical activity
mg: milligravitational unit
MVPA: moderate-to-vigorous-intensity physical activity
PA: physical activity
SB: sedentary behavior
STS: sit-to-stand
WHI/OPACH: Women's Health Initiative/Objective Physical Activity and Cardiovascular Health
WHO: World Health Organization
WHS: Women's Health Study

Edited by T Leung, J Wang; submitted 23.03.22; peer-reviewed by E Rillamas-Sun, A McGuire; comments to author 23.05.22; revised version received 27.06.22; accepted 14.07.22; published 22.08.22.

Please cite as:
Choudhury R, Park JH, Thiamwong L, Xie R, Stout JR
Objectively Measured Physical Activity Levels and Associated Factors in Older US Women During the COVID-19 Pandemic: Cross-sectional Study
JMIR Aging 2022;5(3):e38172
URL: https://aging.jmir.org/2022/3/e38172
doi:10.2196/38172
PMID:35994346

©Renoa Choudhury, Joon-Hyuk Park, Ladda Thiamwong, Rui Xie, Jeffrey R Stout. Originally published in JMIR Aging (https://aging.jmir.org), 22.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementing Affordable Socially Assistive Pet Robots in Care Homes Before and During the COVID-19 Pandemic: Stratified Cluster Randomized Controlled Trial and Mixed Methods Study

Hannah Bradwell[1], PhD; Katie J Edwards[1], MSc; Rhona Winnington[2], PhD; Serge Thill[3], PhD; Victoria Allgar[1], PhD; Ray B Jones[1], PhD

[1]Centre for Health Technology, University of Plymouth, Plymouth, United Kingdom

[2]School of Clinical Sciences, Auckland University of Technology, Auckland, New Zealand

[3]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, Netherlands

Corresponding Author:
Hannah Bradwell, PhD
Centre for Health Technology
University of Plymouth
Drake Circus
Plymouth, PL4 8AA
United Kingdom
Phone: 44 7975927341
Email: hannah.bradwell@plymouth.ac.uk

Abstract

Background: Robot pets may assist in the challenges of supporting an aging population with growing dementia prevalence. Prior work has focused on the impacts of the robot seal Paro on older adult well-being, but recent studies have suggested the good acceptability and implementation feasibility of more affordable devices (Joy for All [JfA] cats and dogs).

Objective: We aimed to address the limited effectiveness research on JfA devices.

Methods: We conducted an 8-month, stratified, cluster randomized controlled trial in 8 care homes in Cornwall, United Kingdom. Over 4 months, 4 care homes each received 2 JfA devices (1 cat and 1 dog; intervention group), and 4 homes received care as usual (control group). Psychometrics were collected before and after the intervention to compare the change from baseline to follow-up between the groups. In the final 4 months, all 8 care homes received devices, but only qualitative data were collected owing to COVID-19 and reduced capacity. The primary outcome was neuropsychiatric symptoms (Neuropsychiatric Inventory [NPI] Nursing Home version). Care provider burden was a secondary outcome (occupational disruptiveness NPI subscale), alongside the Challenging Behavior scale, the Holden communication scale, the Campaign to End Loneliness questionnaire, and medication use. Qualitative data were collected through care staff observation calendars and end-of-study interviews to understand use, experience, and impact. We also collected demographic data and assessed dementia severity. In total, 253 residents had robot interaction opportunities, and 83 were consented for direct data collection.

Results: There was a significant difference in the total change from baseline to follow-up between the intervention and control groups for NPI (P<.001) and occupational disruptiveness (P=.03). Neuropsychiatric symptoms increased in the control group and decreased in the intervention group. No significant difference was seen for communication issues or challenging behavior. For NPI subdomains, there were significant differences from baseline to follow-up in delusions (P=.03), depression (P=.01), anxiety (P=.001), elation (P=.02), and apathy (P=.009), all of which decreased in the intervention group and increased slightly in the control group. The summative impact results suggested that most residents (46/54, 85%) who interacted with robots experienced a positive impact. Those who interacted had significantly higher dementia severity scores (P=.001). The qualitative results suggested good adoption, acceptability, and suitability for subjectively lonely individuals and lack of a novelty effect through sustained use, and demonstrated that the reasons for use were entertainment, anxiety, and agitation.

Conclusions: Affordable robot pets hold potential for improving the well-being of care home residents and people with dementia, including reducing neuropsychiatric symptoms and occupational disruptiveness. This work suggests no novelty effect and contributes toward understanding robot pet suitability. Moreover, interactions were more common among residents with more moderate/severe dementia and those subjectively lonely.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT04168463; https://www.clinicaltrials.gov/ct2/show/NCT04168463

(JMIR Aging 2022;5(3):e38864) doi:10.2196/38864

KEYWORDS

social robots; companion robots; well-being; older adults; dementia; robot pets; COVID-19

Introduction

Background

Robot pets may offer a psychosocial method of improving well-being in older adults and people with dementia. The most well researched robot pet is Paro, the robot seal [1,2]. The use of Paro for individuals in care homes or those with dementia suggests benefits of reduced agitation and depression [3], more adaptive stress responses [4], reduced loneliness [5], reduced care provider burden [4,6], and reduced psychoactive and analgesic medication use [7]. However, Paro is expensive at approximately GBP 5000 (approximately US $6000) per robot, and this limits the number of people able to benefit from interactions [8]. The impact of this cost is evident in the limited number of real-world implementations of Paro. Additionally, robot pet alternatives to Paro have received much less research interest, creating further requirement for work such as this study.

The Joy for All (JfA) cat and dog seem to be preferred over Paro by older people in the United Kingdom, are more affordable [9-12], and are now widely used [13], although there is relatively little formal research on their benefits. A longitudinal 6-month staff diary study suggested that JfA devices had potential benefits of reduced agitation, increased communication, positive experiences, and de-escalated situations [14]. Other studies of JfA devices suggested possible positive impacts [15], including for communication, with conversations being facilitated [16,17], as well as providing companionship [18,19] and improving loneliness, mental well-being, and purpose [20]. However, there were some incidences of negative responses, such as jealousy and over-attachment [14], or dislike and rejection [16]. Much of the prior work with JfA devices has been conducted with community-dwelling older adults [16-18,20] and is mainly qualitative, with small samples [17,18].

Wexler et al [21] conducted a randomized controlled trial (RCT) with a JfA cat and dog in older adults who were hospitalized. A total of 160 older adults took part, with 80 receiving animals for the duration of their hospitalization and 80 receiving 15-minute visits from a nursing student (control group). Participants who received the JfA robot pet experienced less delirium and loneliness, and fewer falls. There was no significant effect found for cognition or depression. However, the study was conducted within a hospital rather than a care home, and each participant received a robot, which would be costly for care homes, even at the more affordable price. The study also did not measure impacts on symptoms, such as agitation and anxiety, which are commonly reported outcomes for people receiving robot pets [2,14]. It is unclear at present why the participants had been hospitalized, if any participants had dementia, or if the participants usually resided in the community or in care facilities. The duration of hospitalization is also unclear.

Marsilio et al [19] conducted the most relevant study. In their study, they provided a JfA cat to 11 care home residents for 6 weeks. They measured agitation, oxygen saturation, heart rate, and medication use at baseline and following the intervention. Qualitative weekly reflections were also maintained. They observed a decrease in agitation and an increase in oxygen saturation. However, the study had a small sample, was over a short timeframe, and had no control group. The authors provided limited details on device implementation, such as quantity, intervention dose, intervention schedule, or method of use (eg, facilitated/unfacilitated interactions or individual/group sessions).

Aims

We aimed to address the lack of longer-term real-world research by performing a study with a large sample of care home residents and exploring the effectiveness of affordable robot pets. In particular we (1) explored if affordable robot pets led to improved well-being for an intervention group in comparison with a control group; (2) aimed to provide an indication of whether robots are robust and engaging over 8 months; and (3) identified under what circumstances and for which care home residents the robot pets were used.

Methods

Ethics Approval and Trial Registration

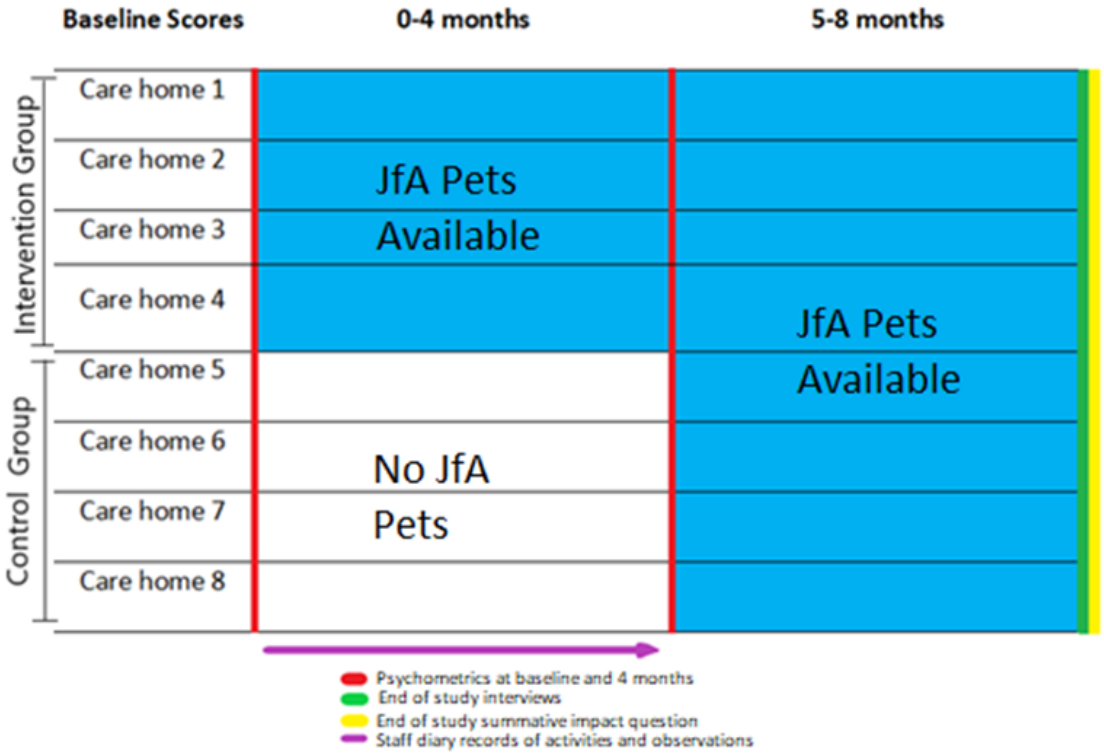
This study received ethical approval from the Health Research Authority (November 13, 2019; North East – Newcastle & North Tyneside 2 REC; Integrated Research Application System number: 268571). This study was registered on ClinicalTrials.gov (November 19, 2019; reference NCT04168463), and is reported following the CONSORT 2010 statement: extension to cluster randomized trials [22].

Research Design

This study was planned as a stepped-wedge, stratified, cluster RCT [23]. The clusters were 8 care homes. However, the trial commenced in January 2020, and the COVID-19 pandemic, resultant care home lock downs, staff workloads, and resident deaths meant that we were unable to carry out the RCT as originally planned. This variation in the planned RCT is described in Multimedia Appendix 1. The study, as conducted (Figure 1), comprised a 4-month, parallel, stratified, cluster RCT with 4 care homes in each arm. This was followed by a qualitative study over an additional 4 months, where all 8 care homes received robots, which ended with staff telephone interviews and a summative impact question. The summative impact question was a simple tool designed by the authors, where staff were asked what impact the robots had for each resident (no impact, positive impact, negative impact, or no interaction).

XSL•FO
RenderX

Figure 1. Research design and data collection. JfA: Joy for All.



As demonstrated in Figure 1, the quantitative scales represented a parallel control trial, where metrics were collected for residents in the control and intervention groups at baseline and in the following 4 months. As care staff capacity was limited by the pandemic, scales were not repeated at 8 months. Diary records were maintained in both the control homes and intervention homes for the first 4 months. Due to limited staff capacity during pressures of the pandemic, diary entries were not recorded from 4 to 8 months. The qualitative impact of robots for all residents in all 8 homes was collected at 8 months through telephone interviews and a summative impact question.

Collaborating Sites

Eight residential care facilities present in rural towns in Cornwall, comprising 4 care homes with nursing care and 4 residential-only care homes, with a total resident population of 253, agreed to collaborate before the start of the project. Sixteen care staff became collaborators for the purpose of completing scales and recording observations of residents. Homes were eligible for participation if they provided residential care or nursing to older adults, were situated in Cornwall, United Kingdom, and allowed regular researcher visits.

Recruitment of Residents for the Collection of Individual Data

In November 2019, researchers and care staff talked to residents or residents' relatives to gauge interest in participation. Prior to randomization, written informed consent was obtained directly from 30 individuals with the capacity to consent and from 53 authorized third parties for individuals without capacity. Where consent involved advice from a consultee of a participant, care home collaborators were encouraged to use measures of assent throughout the trial to ensure participant comfort. Care staff were asked to be mindful not to cause residents distress if they did not like the robots. The 83 care home residents recruited for directly collected data comprised 61 females and 22 males, and represented 32.8% (83/253) of all residents who had access to the robot pets. To allow stratified randomization, staff assessed consenting residents using the Dementia Severity Rating Scale [24]. This provides a score ranging from 0 to 54, with 0-18 indicating mild dementia, 19-36 indicating moderate dementia, and 37-54 indicating severe dementia.

Randomization

The 8 care homes were stratified into 4 pairs based on the number of consented residents, average age, and average dementia severity (as key factors likely to influence behavior) using randomly permuted blocks of size 2 by HB. Each member of the pair was then randomly allocated to either group A or group B, and finally group A and group B were randomly allocated by a separate researcher (KJE) using a random number generator to the intervention or control arm in a 1:1 ratio (homes 1-4 and homes 5-8).

Data Collection

Individual Participant Data

We aimed to collect pre/post data on 5 scales for 83 residents who had consented. The primary outcome was neuropsychiatric symptoms, measured with the Neuropsychiatric Inventory (NPI) by staff [25], with higher scores indicating higher symptom prevalence. Secondary outcomes were measured with the Challenging Behavior Scale [26], Holden Communication Scale [27], and NPI occupational disruptiveness subdomain scale by staff. Residents were assisted in directly completing the 3-item Campaign to End Loneliness [28] questionnaire. The 5 scales were completed at baseline (December 2019) and at 4 months

XSL•FO
RenderX

(May 2020). Finally, staff indicated, through a summative impact question, whether (1) each resident had no interaction with robots, (2) robots had a negative impact, (3) robots had no impact, or (4) robots had a positive impact for all participants at 8 months, as part of an *end of study reflection*, when the intervention group had been using robots for 8 months and the control group had been using robots for 4 months.

Data Collection Tools for Individual Outcomes at Baseline and 4 Months

The primary outcome was assessed with the NPI Nursing Home version [25] (total score 0-120), with higher scores indicating higher symptom prevalence.

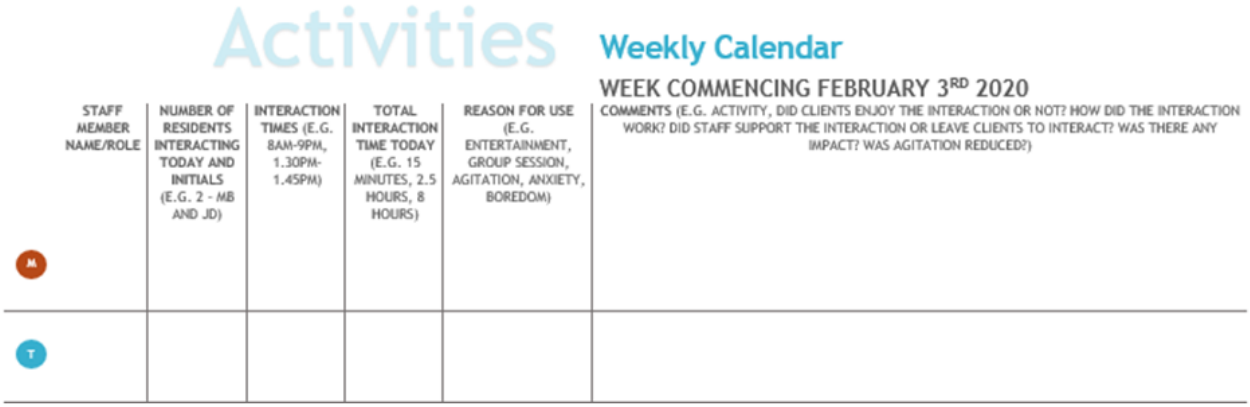
The secondary outcomes were assessed with the following: (1) Challenging Behavior Scale [26] (scored 0-400), with higher scores indicating more challenging behavior; (2) Holden Communication Scale [27] (scored 0-48), with higher scores indicating greater communication challenges; (3) Campaign to End Loneliness Measurement Tool (3-item) [28] (scored 0-12), with higher scores indicating greater loneliness; and (4) NPI subdomain scales [25] (scored 0-12) and NPI occupational disruptiveness scale (scored 0-50), with higher scores indicating more disruptiveness.

Cluster (Care Home) Level Data at 8 Months

Moyle et al [29] noted that behavioral and psychological improvements were not always shown through the chosen scales, and that an evaluation should look beyond these for a picture of overall effectiveness, including comments and observations of care staff and family members. Collaborating care staff in all homes were encouraged to record observations on their calendars using an experience sampling method [30]. Based on our previous use of diaries [14], we created wall-hung calendars for data entry (Figure 2).

Staff were asked to record notes on the calendar each time they observed resident-robot interactions, where possible. We also conducted qualitative semistructured interviews with care staff at 8 months, with open questions aiming to understand the robots' use, engagement, and impact, and the experiences of the staff and residents (Textbox 1).

Figure 2. Example calendar for recording activities. Monday and Tuesday rows are shown (full page includes all days of the week).



Textbox 1. Semistructured interview guide.

Questions (text in brackets was not spoken but provided as notes for the researchers; additionally, questions on benefits were only asked if benefits were mentioned)

- Tell me about your experience with the robot pets here at (name of home)?
- How were the robot pets used?
- Was there any impact? (positive or negative impact – follow up questions included: please explain how many residents benefitted [if benefits were mentioned] and how?)
- (If benefits were mentioned above) Which residents benefited? Would you say there were residents the pets were more or less suited to based on your first-hand experience?
- Were there any particular features of the pets you perceived positively based on their use here with residents?
- Were there any particular features of the pets you perceived negatively based on their use here with residents?
- How did the residents engage with the robot pets?
- Has there been any change in their use over time?
- Has there been any change in reactions over time?
- Any additional comments or observations?
- Were there any practical considerations? (eg, robustness, cleaning, and batteries)
- How did the COVID-19 pandemic and lockdown affect use?

XSL•FO
RenderX

Intervention

In mid-January 2020, homes in the intervention group were gifted a JfA cat and dog to keep indefinitely and to use or not use as they felt appropriate. The researcher provided infection control information [31], providing care homes with the cleaning protocol and informing them of products to use. This study aimed to respond to limitations of trials with highly controlled intervention doses, and explore robot pet effectiveness rather than efficacy [32]. The researcher discussed past research with care home staff, providing examples and ideas, including prior work that implemented robots with structured daily group or individual sessions [3,29], or used robots *when required* for reducing loneliness, anxiety, depression, or agitation, as in previous research with Paro [33]. Decisions on robot use were then left to the professional judgement of care staff. In the fourth month of the trial, the pandemic resulted in changes to the use of robots, with homes tending to reserve robots for specific individuals during specific times, rather than conducting group activities with robots passed between residents.

Sample Size

The sample size was primarily informed by feasibility and the number of residents in each home providing consent, but we calculated the minimum number required for the total sample. Based on previous work reporting on the minimal clinically important difference for the original NPI [34], we calculated using the lower value of 2.77, with an estimated SD of 3.31. To detect a difference of 2.77 between groups, based on 80% power and 5% significance, a sample size of 25 per condition would be required, inflated by 20% to account for any loss to follow-up, resulting in a total sample size of at least 70 individuals.

Data Analysis

Descriptive statistics are presented as mean (SD), median (IQR), and n (%). The changes from baseline to 4 months on the primary and 3 secondary ordinal scales were compared between the control and intervention groups using the Mann Whitney *U* test. SPSS 25 (IBM Corp) was used for statistical analysis. A *P* value <0.05 was considered to indicate statistical significance.

Qualitative diaries and interviews were individually subject to content analysis, and have been reported together due to great similarity of themes. Content analysis follows similar processes to thematic analysis, involving coding and categorizing of textual information; however, the frequency of occurrence is of additional importance [35].

Quantitative Scales

For the quantitative measures, we first report the primary (NPI) and secondary psychometric outcomes (communication, challenging behavior, and occupational disruptiveness) and report intention-to-treat (ITT) results for all residents (as randomized) who survived to the 4-month follow-up (n=63). We then report NPI subdomain results, followed by the summative impact question completed by a staff member at 8 months, to indicate the overall robot impact for each consented resident (n=83). We subsequently report a comparison of the characteristics of residents who did and did not interact with robots during the study, to comment on the suitability of the devices, based on residents who survived till follow-up (n=63). This was because of the possibility that residents who died never had the opportunity to interact with the robots, rather than, for example, rejected the robots owing to a lack of suitability.

Results

Participants

The average age of consented participants was 87.21 years (SD 7.42 years), and the average dementia score was 32.11 (SD 10.52) (Tables 1 and 2). Twenty of the 83 residents recruited died during the study, leaving 63 participants for analysis (49 females and 14 males) (Table 1; Figure 3). There was no difference in the dementia severity (U=513; n=63; *P*=.65) or age (U=549; n=63; *P*=.34) of residents included in the analysis between the intervention and control groups.

Figure 3 shows that a greater number of deaths occurred in the intervention group than in the control group. Considering our concerns on infection control and the timing of the trial in the early stages of the COVID-19 pandemic, we carried out more detailed analysis of deaths and enquired with care home staff. Further details are given in Multimedia Appendix 2.

XSL•FO
RenderX

Table 1. Demographic characteristics of the participating homes and consented participants.

Home	Site type	Staff collaborators (N=16)	Total residents (N=253)	Consented residents (N=83)	Gender (22 M, 61 F)[a]	Residents included in the analysis (N=63)
1[b]	Nursing	2	33	9	3 M, 6 F	3
2[b]	Residential	2	16	11	1 M, 10 F	10
3[b]	Nursing	2	36	9	4 M, 5 F	4
4[b]	Residential	2	36	12	4 M, 8 F	9
5	Nursing	2	36	7	4 M, 3 F	4
6	Residential	2	27	13	4 M, 9 F	12
7	Nursing	2	31	13	1 M, 12 F	12
8	Residential	2	38	9	1 M, 8 F	9

[a]M: male; F: female.

[b]Homes included in the intervention group (see Figure 1).

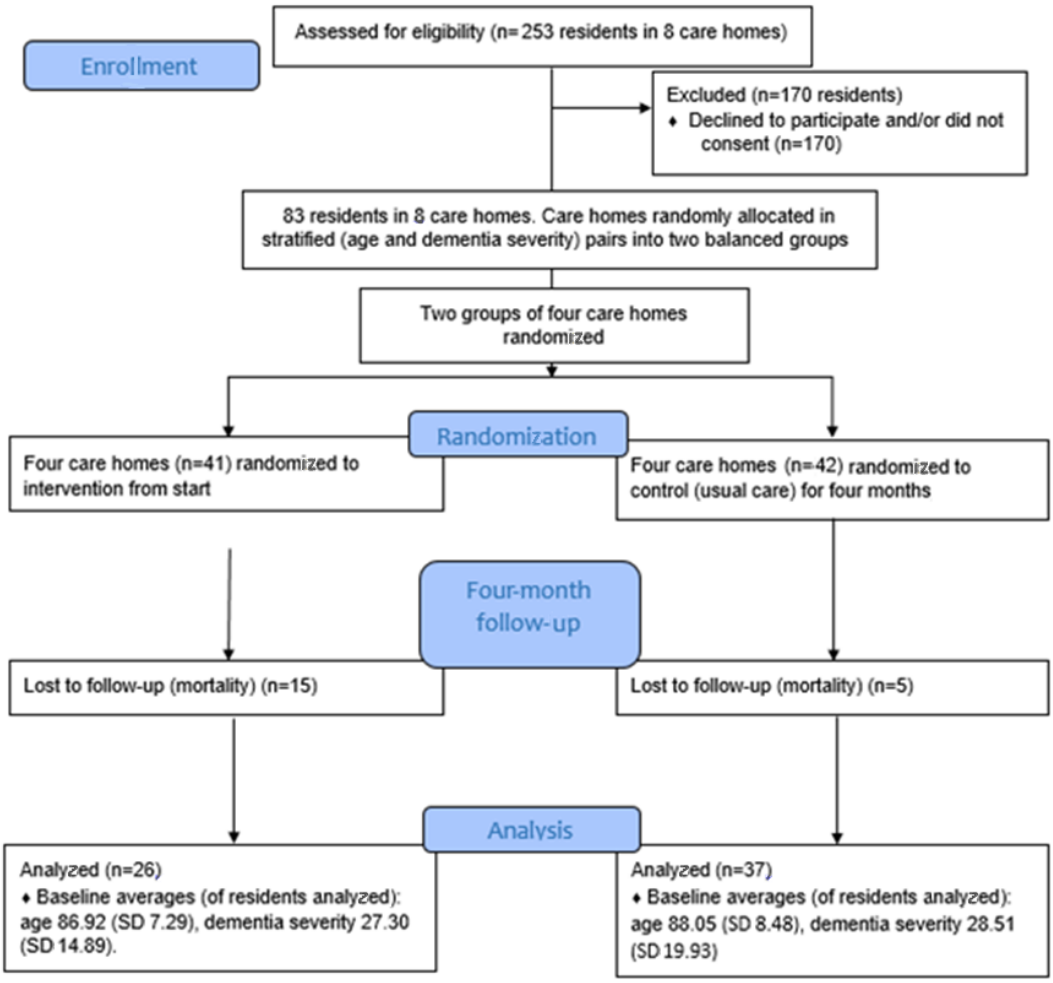
Table 2. Demographic characteristics of the consented participants.

Home	Age of the consented residents (years), mean (SD)	Age of the residents analyzed (years), mean (SD)	Dementia severity score for the consented residents (scored 0-54), mean (SD)	Dementia severity score for the analyzed residents (scored 0-54), mean (SD)
1[a]	87.67 (6.73)	86.33 (7.37)	40.56 (9.38)	43.33 (9.71)
2[a]	90.73 (7.85)	90.10 (7.97)	19.63 (12.82)	17.30 (10.76)
3[a]	82.89 (2.51)	83.00 (7.39)	44.11 (8.25)	37.50 (7.59)
4[a]	85.08 (6.33)	85.33 (6.1)	32.58 (15.77)	28.56 (15.58)
5	86.29 (10.05)	87.75 (9.60)	36.14 (10.07)	35.75 (7.58)
6	90.46 (9.53)	89.42 (9.14)	5.23 (5.93)	4.75 (5.93)
7	85.15 (8.34)	85.75 (8.41)	46.77 (6.13)	47.33 (6.03)
8	89.44 (8.00)	89.44 (8.00)	31.89 (15.84)	31.89 (15.84)
Overall	87.21 (7.42)	87.14 (8.00)	32.11 (10.52)	30.80 (9.88)

[a]Homes included in the intervention group (see Figure 1).

XSL•FO
RenderX

Figure 3. CONSORT diagram of trial recruitment, allocation, and analysis of data.



Psychometric Analysis

The data for communication issues, challenging behavior, neuropsychiatric symptoms, and occupational disruptiveness are presented in Tables 3 and 4.

Table 4 demonstrates that based on ITT analysis, there was a significant difference in the total change in NPI and occupational disruptiveness scores between the intervention and control groups. Neuropsychiatric symptoms increased in the control group and decreased in the intervention group. No significant difference was present for communication issues or challenging behavior from baseline to follow-up between the control and intervention groups.

NPI data for the intervention and control groups are presented in Tables 5 and 6.

When looking at the individual domains, there was a significant difference between the control and intervention groups in the total change from baseline to follow-up for delusions, depression, anxiety, elation, and apathy, all of which decreased in the intervention group and increased slightly in the control group. There was no significant difference from baseline to follow-up between the 2 groups for other subdomains. Multimedia Appendix 3 presents issues in the normality of the data, justifying the choice of nonparametric analysis.

XSL•FO
RenderX

Table 3. Baseline and 4-month scores for communication issues, challenging behavior, neuropsychiatric symptoms, and occupational disruptiveness in the control and intervention groups.

Scale (scoring)	Intention-to-treat analysis (as randomized) (N=63)[a]							
	Baseline				Follow-up			
	Control (n=37)		Intervention (n=26)		Control (n=37)		Intervention (n=26)	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Communication (0-48)	20.57 (15.13)	21.00 (29.5)	16.58 (11.85)	15.00 (20.75)	21.97 (15.12)	22.00 (30.00)	17.23 (15.33)	14.00 (29.75)
Challenging behavior (0-400)	54.86 (56.95)	32.00 (82.00)	43.38 (43.02)	26.00 (53.00)	48.22 (53.98)	29.00 (73.00)	31.85 (38.39)	19.50 (36.00)
Neuropsychiatric Inventory (0-120)	16.64 (16.41)	16.00 (13.50)	19.19 (17.08)	15.00 (22.50)	19.41 (18.72)	11.00 (26.00)	9.62 (7.83)	9.00 (10.75)
Neuropsychiatric Inventory occupational disruptiveness (0-50)	5.51 (6.37)	4.00 (8.00)	4.42 (4.86)	3.00 (7.00)	5.46 (6.26)	3.00 (8.50)	3.19 (4.54)	1.00 (3.25)

[a]The intention-to-treat analysis excludes the 20 residents who died but includes the 63 who potentially had access to the robots.

Table 4. Difference in communication issues, challenging behavior, neuropsychiatric symptoms, and occupational disruptiveness from baseline to follow-up between the control and intervention groups.

Scale (scoring)[a]	Intention-to-treat analysis (as randomized) (N=63)[b]		
	Difference (baseline to follow-up), mean (SD)		Test of difference (control vs intervention)
	Control (n=26)	Intervention (n=20)	Mann-Whitney *U* test *P* value
Communication (0-48)	1.41 (6.00)	0.65 (7.54)	.18
Challenging behavior (0-400)	−6.65 (25.65)	−11.54 (23.92)	.35
Neuropsychiatric Inventory (0-120)	2.76 (9.43)	−9.58 (14.06)	<.001
Neuropsychiatric Inventory occupational disruptiveness (0-50)	−0.05 (2.47)	−1.23 (2.53)	.03

[a]For all scales, higher scores indicate greater prevalence of challenges.

[b]The intention-to-treat analysis excludes the 20 residents who died but includes the 63 who potentially had access to the robots.

XSL•FO
RenderX

Table 5. Baseline and 4-month Neuropsychiatric Inventory domain data for the intervention and control groups.

Scale (scored 0-12)	Baseline				Follow-up			
	Control (n=37)		Intervention (n=26)		Control (n=37)		Intervention (n=26)	
	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)	Mean (SD)	Median (IQR)
Delusions	0.76 (2.46)	0.00 (0.00)	1.57 (3.34)	0.00 (0.25)	1.43 (3.18)	0.00 (.50)	0.19 (0.80)	0.00 (0.00)
Hallucinations	0.49 (2.04)	0.00 (0.00)	0.73 (1.95)	0.00 (0.00)	1.03 (2.69)	0.00 (0.00)	0.27 (0.87)	0.00 (0.00)
Agitation	4.68 (3.86)	4.00 (7.50)	3.42 (4.20)	2.50 (6.00)	3.70 (4.27)	2.00 (7.00)	1.00 (2.40)	0.00 (0.25)
Depression	2.43 (3.21)	2.00 (3.00)	2.08 (2.53)	0.50 (4.50)	3.03 (2.94)	2.00 (5.00)	1.62 (3.03)	0.00 (2.50)
Anxiety	2.30 (3.19)	1.00 (3.50)	3.31 (4.25)	0.00 (8.00)	2.92 (3.55)	2.00 (6.00)	0.84 (2.12)	0.00 (0.00)
Elation	2.30 (3.19)	0.00 (0.00)	1.31 (2.65)	0.00 (2.00)	0.84 (2.28)	0.00 (0.00)	0.92 (2.61)	0.00 (0.00)
Apathy	2.24 (2.56)	2.00 (4.00)	3.58 (3.30)	4.00 (6.00)	2.76 (3.55)	2.00 (4.00)	2.38 (3.45)	0.00 (4.00)
Disinhibition	0.78 (2.76)	0.00 (0.00)	0.37 (1.30)	0.00 (0.00)	0.78 (2.76)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Irritability	2.62 (3.36)	1.00 (4.00)	1.54 (3.05)	0.00 (2.00)	2.59 (3.48)	0.00 (6.00)	1.19 (2.83)	0.00 (1.25)
Motor behaviors	0.14 (0.67)	0.00 (0.00)	1.31 (2.69)	0.00 (0.75)	0.32 (1.11)	0.00 (0.00)	1.19 (2.68)	0.00 (0.00)
Sleep behaviors	1.22 (2.85)	0.00 (0.50)	1.38 (2.74)	0.00 (2.25)	0.24 (1.04)	0.00 (0.00)	1.27 (2.91)	0.00 (0.50)
Eating behaviors	0.46 (1.10)	0.00 (0.00)	1.81 (4.10)	0.00 (0.00)	0.35 (0.92)	0.00 (0.00)	0.88 (2.80)	0.00 (0.00)

Table 6. Difference in the Neuropsychiatric Inventory domains from baseline to follow-up between the intervention and control groups.

Scale (scored 0-12)	Intention-to-treat analysis (as randomized) (N=63)		
	Difference (baseline to follow-up), mean (SD)		Test of difference (control vs intervention)
	Control (n=37)	Intervention (n=26)	Mann-Whitney *U* test *P* value
Delusions	0.68 (2.85)	−1.38 (3.46)	.03
Hallucinations	0.54 (1.48)	−0.46 (2.21)	.06
Agitation	−0.97 (2.93)	−2.42 (3.76)	.22
Depression	0.56 (2.30)	−0.46 (3.19)	.01
Anxiety	0.62 (1.93)	−2.46 (4.37)	.001
Elation	0.62 (2.00)	−0.38 (2.47)	.02
Apathy	0.51 (2.43)	−1.19 (3.14)	.009
Disinhibition	0.00 (0.00)	−0.35 (1.29)	.08
Irritability	−0.03 (3.47)	−0.35 (3.39)	.55
Motor behaviors	0.19 (0.81)	−0.12 (0.59)	.10
Sleep behaviors	−0.97 (2.98)	−0.12 (0.99)	.19
Eating behaviors	−0.11 (0.66)	−0.92 (3.26)	.34

Summative Impact Question

The summative question asked about care staff members' perceptions of robot use and the impact for all residents at the 8-month time point after all homes had received robots and had been implementing them for either 4 or 8 months. Among the residents reported to interact with the robots (54/81), 85% (46/54) were reported to have positive experiences. Table 7 demonstrates that most residents who survived the 8 months and were included in the analysis (61/81) interacted with the pets (46/61, 75%), and that most (40/61, 66%) experienced a positive impact, with only 1 resident (male) experiencing a negative impact. This summative question provided the perception of 1 member of the staff in each home, and thus, there may be inaccuracies based on different staff members observing robot use with different residents, although the collaborating staff member was always the staff member in each home with the most insight and experience. Additionally, this observation may suffer from memory strain, with staff asked to reflect over the prior 8 months. However, as nearly a quarter (15/61) of the residents included in the analysis did not interact with the robots (Table 7), we performed a comparison of the characteristics of residents who did and did not interact with the robots to comment on suitability.

XSL•FO
RenderX

Table 7. Care staff summative estimation of the impact of robot pets for each resident at 8 months (N=83).

Care home	Total number of residents	Consented participants	Died by the 4-month follow-up	No interaction	Negative impact	No impact	Positive impact
Intervention care home, n							
1	33	9	6	4	0	1	4
2	16	11	1	2	0	0	9
3	36	9	5	5	0	1	3
4	36	12	3	2	0	1	9
Overall	121	41	15	13	0	3	25
Control care home, n							
5	36	7	3	0	0	1	3
6[a]	27	13[a]	1	7	0	0	3
7	31	13	1	0	0	3	9
8	38	9	0	2	1	0	6
Overall	132	42	5	9	1	4	21
All participants, n (%)	253 (100)	83 (32.8)	20 (24.1)	22 (27.2)[a]	1 (1.2)[a]	7 (8.6)[a]	46 (56.8)[a]
Residents included in the RCT[b] analysis at 4 months (n=61[a]), n (%)	N/A[c]	N/A	N/A	15 (24.6)[a]	1 (1.6)[a]	5 (8.2)[a]	40 (65.6)[a]

[a]Data on interaction are missing for 2 people in home 6.

[b]RCT: randomized controlled trial.

[c]N/A: not applicable.

Difference Between Interacting and Noninteracting Residents

Residents who subsequently went on to interact with robots had significantly higher dementia severity scores than residents who did not interact (Table 8). On average, residents who did interact were considered to be at the higher end of moderate dementia (19-36), while residents who did not interact were considered to have mild dementia (0-18). The interacting residents also had significantly poorer communication scores and scored significantly higher for challenging behaviors and NPI occupational disruptiveness. There was no difference in the overall NPI score, age, or gender.

Table 8. Baseline characteristics of residents who did and those who did not interact with robots.

Scale	Residents who did interact (n=46), mean (SD)	Residents who did not interact (n=15), mean (SD)	Mann-Whitney *U* test *P* value
Communication	22.22 (13.29)	11.20 (11.98)	.005
Challenging behavior	61.02 (54.73)	22.20 (26.27)	.003
Neuropsychiatric Inventory	20.28 (18.09)	11.40 (9.06)	.06
Neuropsychiatric Inventory occupational disruptiveness	6.15 (6.23)	2.27 (2.84)	.01
Dementia severity	33.46 (15.60)	14.73 (16.03)	.001
Age (years)	87.02 (7.68)	88.47 (9.08)	.32

The above findings suggest that robots are perhaps more suited to residents scoring higher for dementia severity, who also experience more communication issues and challenging behavior as associated symptoms.

The fact that many care homes restricted shared robot use from 4 months onwards would have influenced some residents not interacting, particularly in control homes where robots were only provided from month 5. However, homes reported aiming to allow interested participants opportunities to interact (individually after robot cleaning rather than group sessions), and robots tended to become *adopted* by residents who found particular benefit. Staff reported not pursuing interactions with residents who were disinterested, feeling they were best placed with *adoptees* in any case.

XSL•FO
RenderX

Qualitative Calendar Entries

During the first 4 months, staff in the 4 control homes provided 139 days of calendar entries describing usual resident activities and moods. Staff in the 4 intervention homes provided 109 days of calendar entries. In total, about 25% (248/960 [8×120]) of care-home-days were captured. The diaries reported a total of 516.3 hours of interaction with the robots over the 4 months, with an average interaction length of 3.9 hours. The range of interaction lengths varied from 0.25 hours to 24 hours, where residents kept the robot with them all day and night. On average, 4 residents interacted with the robots on each reported day (range 1-8). The main reasons recorded in the *reason for use* of robots were entertainment, anxiety, and agitation (Table 9). In control homes, typical activities included singing, manicures, reminiscence, television, garden games, hairdresser visits, and quizzes.

Table 10 demonstrates the themes resulting from analysis of comments made in the calendars and interviews. The full table of themes with quotes as example evidence is available in Multimedia Appendix 4, further to a full narrative on the themes.

Table 9. Reported reasons for using robots (N=109).

Reason	Value, n
Entertainment	40
Anxiety	33
Agitation	31
Boredom	30
Group session	10
Company	7
Love	6
Cuddles	4
Nurturing	3
Loneliness	3
Affection	2
Stress	1
Distress	1
Distraction	1
Observation	1
Sadness	1
Reassurance	1

XSL•FO
RenderX

Table 10. Content analysis of qualitative interviews and calendar entries.

Theme (*explanation*)	Codes (interviews: n; diaries: n)
Adoption (*Evidence strongly supported good robot adoption into services, and usually by particular "adoptee" residents*)	Love (interviews: 11; diaries: 13)
	Ownership (interviews: 18; diaries: 6)
	Individual use (interviews: 9; diaries: 14)
	High level of usage (diaries: 12)
	Jealousies or possessiveness (interviews: 6; diaries: 6)
	No novelty (interviews: 9)
	Naming (interviews: 7)
	Group sessions (diaries: 5)
	Personalizing (interviews: 1)
Well-being effects, particularly mood (*Evidence strongly supported well-being benefits*)	Calming (interviews: 10; diaries: 20)
	Enjoyment (interviews: 1; diaries: 19)
	Anxiety reduced (interviews: 3; diaries: 13)
	Companionship (interviews: 7; diaries: 6)
	Smiles, happiness (interviews: 1; diaries: 9)
	Engaging resident (interviews: 10)
	Relaxing or settling (diaries: 7)
	Mood improved (interviews: 7)
	Provides a focus (interviews: 5)
	Distraction (interviews: 3; diaries: 2)
	Agitation reduced (diaries: 5)
	Entertainment and laughter (interviews: 1; diaries: 3)
	Therapeutic (interviews: 3)
	Reassurance (interviews: 3)
	Sundowner (interviews: 2)
	Reduced boredom (interviews: 1; diaries: 1)
	Enabled eating (diaries: 1)
Effects on communication (*Evidence supported robot impact on residents' communication, with the pets and people, further to improving speech capabilities*)	Communication-pet (diaries: 25)
	Communication with others, and speech (interviews: 19; diaries: 2)
	Reminiscence (interviews: 5; diaries: 1)
	Interaction (interviews: 4)
Isolation and COVID (*Evidence showed particular benefits of robot pets as a supporting strategy against loneliness and isolation in response to the COVID-19 pandemic*)	COVID use (interviews: 15)
	Cleanliness and infection control (interviews: 9)
	Isolation (interviews: 5)
Design (*Staff suggested a few possible design improvements based on their experience, and commented on positive and negative design factors*)	Improvements (interviews: 11)
	Realistic (interviews: 9)
	Sound off (interviews: 8)
	Expectations (interviews: 8)
	Weight and size (interviews: 7)
	Breakage (interviews: 7)
	Battery life (interviews: 4)
	Importance of movement (interviews: 4)
	Purring as relaxing (interviews: 2; diaries: 2)
	Heartbeat enjoyable (interviews: 1; diaries: 2)

XSL•FO
RenderX

Theme (*explanation*)	Codes (interviews: n; diaries: n)
Suitability (*The data gave some insight into the most suitable use context for use with residents, including those with dementia and those isolated, perhaps due to mobility impairments*)	Dementia severity (interviews: 31) Limited interest (diaries: 17) Think it is real (interviews: 14) Dislike (interviews: 2; diaries: 9) Wide appeal (interviews: 7) Reduced mobility (interviews: 5; diaries: 1) Previous pets (interviews: 3; diaries: 1) Infantilizing (interviews: 4) Staff dislike (interviews: 1)
Nurture (*Evidence suggested residents tended to care for robots and treat them as living animals*)	Cuddled and fussed (diaries: 29) Feeding (interviews: 8; diaries: 5) Care for and nurture the pet (interviews: 8; diaries: 5)

Discussion

Overview

Our results suggest that affordable robot pets are able to produce important well-being impacts for older adult care home residents, with further potential positive impacts for staff through reduced occupational disruptiveness.

Principal Findings and Comparison With Prior Work

This study strongly supports the usefulness and benefits of implementing affordable robot pets in care homes for older adults. It contributes toward limited literature in this area, with most prior companion robot research focusing on Paro [1,29], a device with limited acceptability among older people [9,10] and too expensive for widespread implementation [12,36]. Previous work considering alternative more affordable robots had been mainly conducted within the community [16,20] or in hospital settings [37,38], with limited generalizability to care home residents [39], and with smaller samples and short time frames [19,40]. Additionally, much previous work has involved highly controlled intervention doses [7,29], thus assessing efficacy rather than potential real-world effectiveness [32]. This study therefore provides an important and novel contribution to companion robot literature.

JfA robots demonstrated significant improvements from baseline to follow-up for the primary outcome of neuropsychiatric symptoms and secondary outcome of occupational disruptiveness based on ITT analysis between the control and intervention groups. The reduction in neuropsychiatric symptoms in the intervention group was an encouraging result, suggesting important effects of affordable robot use, as the NPI measures key behavioral and psychological symptoms associated with dementia [25]. There were no significant differences for the secondary outcomes of communication impairments and challenging behavior. The NPI subscale of occupational disruptiveness was used as an indicator of care provider burden, and the reduction seen here is congruent with results from the study by Saito et al [6] who suggested Paro could decrease care provider burden. We did not use a specific care provider burden scale, with the stigmatizing wording found to discourage carer responses in a pilot study. However, the significant difference in occupational disruptiveness could suggest that the implementation of pets aided in easing the challenges of the carer's role.

When analyzing the individual NPI subdomains, the results suggested significant differences in the mean change from baseline to follow-up between the intervention and control groups for delusions, depression, elation, anxiety, and apathy. This finding suggests that JfA devices can achieve similar well-being outcomes to those reported for Paro, particularly around reducing depression [3,5-7]. The support for impact on delusions is also congruent with the work of Schulman-Marcus et al [38], who reported on stakeholders feeling that JfA devices were useful for hospital patients with delirium. The potential for these more affordable devices to produce promising therapeutic benefits is an important result, with implications for research and practice. Interestingly, we did not find a significant impact for agitation, as previous work did for Paro [3]. Similarly, in the cluster RCT conducted by Moyle et al [29], there was no significant effect on agitation in the Paro intervention group. Moyle et al [29] suggested that chosen psychometrics can sometimes miss behavioral improvements, and suggested complementing scales with qualitative feedback.

Our evidence from qualitative calendars shows the robot's effect on anxiety and agitation as the second and third most common *reasons for robot use*, respectively, strengthening the suggestion that affordable robot pets can produce well-being outcomes. Furthermore, interviews and calendar free-text observations demonstrated that the robots were calming, reduced anxiety, improved mood, relaxed residents, reduced agitation, and provided reassurance. The calendars also demonstrated that the primary *reason for use* of the pets was entertainment, thus providing a meaningful activity. This is congruent with the significantly greater reduction in apathy from baseline to follow-up in the intervention group compared to the control group. The importance of meaningful activities for older adults in care homes cannot be overstated, impacting physical and mental well-being [41]. Reduced apathy and greater engagement in an activity creates an improvement in the quality of life. The calendar and interview data suggest that older adults cared for and nurtured robots, which perhaps provided a sense of responsibility and purpose. Although most nurturing seemed to involve cuddling and fussing the animals, there were also counts

XSL•FO
RenderX

of residents feeding, dressing, and grooming the pets, thus providing care.

In contrast to prior work suggesting that robots could improve communication and interactions [42], our Holden communication scale results demonstrated no significant difference in communication as a result of robot implementation. However, our qualitative results suggested that robots encouraged communication, mediating social connection as shown in previous work with Paro [42]. The communication scale we selected provides a measure of resident speech and conversational ability [27], a possible limitation of our work. Future research may seek to employ measures of social cohesion and quality of interactions. Interestingly, our qualitative results did demonstrate evidence of speech and conversational ability improving in some instances, such as residents with severe aphasia showing no signs of the disease upon communicating with the pet dog. This is a profound result, although not replicated in the chosen scale, thus requiring further exploration in future research.

Our experience of sampling observations through calendars [30] also provided insights into the types of uses of the robots. As we did not provide an intervention dose, this aids in understanding the likely real-world use of the devices. The calendars demonstrated a range of uses, from short 15-minute sessions to 24/7 use by some residents who *adopted* the pets, keeping them day and night, until care staff retrieved them to be cleaned and shared. This result highlights a limitation of prior robot pet trials with highly controlled and prescribed intervention doses [7,29], as real-world use is likely more flexible and variable. Our results demonstrate that robots had high levels of use, and were clearly well adopted into daily practice. Observing staff reported evidence of residents loving pets and displaying ownership tendencies. Importantly, the study demonstrated no novelty effect for devices over 8 months, providing evidence against novelty as a concern for robot pet research and implementation [43]. Regarding use type, there were only 10 counts of group sessions recorded as the *reason for use*; however, these were all recorded prior to COVID-19 restrictions. Evidence in interviews after the 8-month study suggested that most robot use was on an individual basis. Previous work has varied in either group [2,3,5] or individual robot interventions [16,17,37]. While our work suggests that individual intervention is most common, we are unable to comment on the generalizability of this result to nonpandemic contexts. However, availability of multiple devices appears desirable owing to some issues in sharing and jealousies evidenced in our qualitative results.

The qualitative evidence also gave some further insights into robot design, based on longitudinal experience with robot pets. As in our previous work [10], stakeholders commented on hygiene as a design limitation of current devices, requesting removable shells for easier cleaning. Participants again supported the importance of realistic design, life-simulation features, and interactivity. Stakeholders felt the JfA cat had more appropriate vocalizations than the dog, although the importance of mute options (which the JfA devices have) was highlighted. Ultimately, design preferences seen here in the longitudinal work are consistent with the results of our previous cross-sectional design studies, supporting the validity of our earlier results [9-12]. In contrast to our previous work [14] suggesting that the devices were suitably robust, this study reported cases of breakage. We know of 5 broken pets throughout this trial, from a total of 18 pets (16 original and 2 replacements). One JfA cat had broken limbs (cause unknown), which did not hinder its use; 1 cat was dropped in urine, which rendered it unusable; and 3 dogs had technical malfunctions. The variance in reported robustness between prior work and this study could have resulted from the different settings (supported living vs care homes) and the thorough exploration with more devices, creating greater opportunity for issues to become evident. Despite the issues, only 2 devices required replacing as the other 3 remained mainly usable.

Owing to the timing of this trial, we were able to gather some understanding of the use and impact of robot pets during the COVID-19 pandemic, and resultant lockdown and isolation, which is entirely novel. The evidence suggests, in line with a previous report [44] and our suggestions, that homes took extra precautions for shared robot use. Despite this, the pets were highly valuable tools during the pandemic and lockdowns, with care staff reporting strongly on the value during the unprecedented times. The pets aided in reducing loneliness and providing company and comfort for residents experiencing long periods without visitors or usual excursions. The pets were also used for residents shielding in self-isolation, and were beneficial for those alone in their bedroom. This is a positive result and has implications for care homes and other aged care services, suggesting that the provision of robot pets for individual use during pandemic situations may ease the challenges of isolation. Isolation is particularly pertinent for care home residents [39], highlighting the value of this finding. Despite these benefits, the risk of use during pandemic situations must be thoroughly assessed, in light of the risks detailed in a previous report [44]. Here, our results demonstrated high numbers of mortalities in collaborating homes. While our enquiries suggested that the deaths were unrelated to robot presence, the risk needs to be considered appropriately, as with all shared surfaces, social contact, and cleaning procedures in the homes.

Regarding general acceptability, the summative impact question demonstrated that, encouragingly, 85% of residents who interacted with the robots experienced a positive impact based on carer observations and 74% of residents included in the analysis interacted with the robots. However, the finding that almost a quarter of the residents included in the analysis did not interact with the robots indicates that the devices lack universal appeal. This result, combined with 11 qualitative counts of robot dislike, is congruent with previous research reporting variation in the response to Paro [8,29], who was described as a "therapeutic tool that's not for everybody" [45]. In contrast to the prior work with Paro, where acceptability was reported to be 50% [45], the JfA devices seem more generally acceptable.

Regarding device suitability, the results demonstrated that residents who interacted with the robots had, on average, more severe dementia, communication issues, and challenging behavior. Previous work has also suggested that companion robots were more suitable for individuals with dementia [44]. This could suggest cognitive impairment and dementia severity

XSL•FO
RenderX

as predictive of likely robot acceptance and benefit; however, this contradicts our earlier work, which demonstrated robot pet acceptability among independent older adults [9] and care home residents without dementia [10]. It is possible that the impact of COVID-19 and the restriction on sharing robot pets in groups led to the prioritization of interactions for more impaired residents. In the qualitative data, evidence suggested that robots were most enjoyed by and beneficial to older adults who had dementia, and also those who were bed bound (due to mobility or illness), less socially engaged (due to dementia), or in isolation (due to COVID shielding). Additionally, residents who were disinterested in the robots were more socially engaged, preferring to play games and socialize with other people. While social engagement appears to be negatively correlated with dementia severity, the results may indicate that both dementia severity and social isolation predict the likelihood of accepting and benefiting from robot pet interventions. This could explain the acceptability of robot pets by more independent older people in prior work [9], as despite not having dementia, the older people lived in individual flats and reported requirement for social company. In previous work with independent older adults living in the community, 4 of 12 robots were rejected [16], with community-dwelling older people less vulnerable to isolation and loneliness [38]. Additionally, Pino et al [46] reported on healthy older adults feeling able to benefit from socially assistive robot (SAR) support, while Tkatch et al [20] reported positive benefits of JfA devices for *self-reported lonely individuals* despite them living in the community. Loneliness and dementia severity are thus likely to be predictive factors in the acceptance and benefit of robot pets in future implementations.

Strengths and Limitations

A strength of this work is the pragmatic mixed-method approach. The use of calendars to support interviews and psychometrics allowed for ecologically valid appraisal of subjective experiences, yielding comprehensive views of activities that may be difficult to assess using cross-sectional questionnaires or interviews, which can cause memory strain and aggregation [30]. A second strength is the somewhat novel approach to this trial that did not specify an intervention dose. This allowed for ecological validity and assessment of the effects on resident well-being based on the likely real-world use of robot pets, with the intervention dose reflecting real-world circumstances. In this regard, our results demonstrate the effectiveness of robot pets and the impact robot pets may genuinely have with real-world implementation, rather than their efficacy, as well as the impact these robots have under highly controlled research contexts with specified intervention doses [32]. Furthermore, not defining an intervention dose avoided the ethical concerns of encouraging robot interaction when residents were resistant and removing robots when they were being enjoyed, as encountered previously [8]. One limitation of this work is the lack of participant responses to the loneliness measure, resulting in the inability to assess the impact on loneliness quantitatively. We had also originally intended to collect medication records, but due to the impact of the COVID-19 pandemic, this was not possible. Prior work with Paro has suggested resultant decreases in the use of psychoactive and analgesic mediation [7], thus this remains a topic for future research. A second limitation is that our analysis reports on the NPI subdomain scores, further to the NPI total, with previous work cautioning that while use of NPI subscales has been popular, validity and reliability are mainly established for the total measure, with the validity of individual scales requiring further testing [47]. A third methodological limitation results from the inability to blind collaborators to conditions. It is possible that the significantly improved outcome measures in the intervention group were a consequence of the inability to blind collaborators. This challenge has been reported in prior Paro RCTs, where the influence of participating in the research itself raised staff awareness to improvements and contributed toward positive findings [3]. It is not possible to distinguish this effect from the intervention. Thus, there is some possibility of positive reporting bias from our collaborators. Additionally, the inability of 2 care home staff to co-jointly complete the 4-month outcome measures may have reduced the validity of the 4-month scores.

The use of a cluster RCT may also be perceived as a limitation over standard RCTs [3]. However, research with older adults and in care home environments presents specific challenges, differing greatly from clinical environments or laboratories. Residents often have dementia, and the ability to randomize residents individually within homes to receive/not receive a robot intervention would be challenging and unethical. Creating clusters from care homes as units, rather than randomizing residents individually, allows for research such as this [3,29]. A final consideration is that the psychometric scales we selected were all designed and validated for older adults and those with dementia. Not all of our participants had dementia; however, the scales were deemed appropriate by our collaborators owing to the high prevalence of dementia in long-term care facilities, such as care homes [29]. Additionally, the content of the chosen scales appears appropriate for older adults with and without dementia, and even those without diagnosed dementia who sometimes experience onset symptoms. Indeed, very few of our participants received a very low score on the dementia severity scale.

Conclusion

Our results suggest that affordable robot pets may have important well-being effects in older adults, including reduced neuropsychiatric symptoms (depression, delusions, elation, anxiety, and apathy), with qualitative accounts also supporting reductions in agitation. This work also suggests that robot use impacted occupational disruptiveness, as an indicator of care provider burden. The findings also support a no novelty effect for affordable robot pets and suggest that the best practice is the permanent availability of multiple devices. One key finding is the contribution to the discussion on the suitability of robot pets. Previous work has suggested that robots are best suited to residents with more severe dementia. This was supported in our work; however, we also suggest that subjective loneliness may be a predictive factor in the acceptance and benefit of robot pets. This work has also demonstrated the important value of the individual use of robot pets during the COVID-19 pandemic, with easing of the challenges of isolation through the provision of social companionship.

XSL•FO
RenderX

Acknowledgments

The authors would like to sincerely thank the care home managers, staff, residents, and family members who made this study possible, particularly during such a challenging period of time for the care sector during the COVID-19 pandemic. Preparatory work for the study was funded by a University of Plymouth PhD studentship. The work was subsequently supported by the EPIC (eHealth Productivity and Innovation in Cornwall and the Isle of Scilly) project, which was part funded by the European Regional Development Fund. Additional funding for the EPIC project was received from South West Academic Health Science Network and University of Plymouth. No specific funding was allocated for this study, but the general funds for the EPIC project allowed for researchers to support data collection and analysis.

Data Availability

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

Conflicts of Interest

None declared.

Multimedia Appendix 1
Variation from the planned stepped-wedge trial due to the COVID-19 pandemic.
[DOCX File , 87 KB - aging_v5i3e38864_app1.docx]

Multimedia Appendix 2
Further analysis of deaths during the trial and the impact of COVID-19.
[DOCX File , 17 KB - aging_v5i3e38864_app2.docx]

Multimedia Appendix 3
Histograms demonstrating normality issues for the primary outcome of neuropsychiatric symptoms at baseline and 4 months in both the intervention and control groups.
[DOCX File , 92 KB - aging_v5i3e38864_app3.docx]

Multimedia Appendix 4
Full table of themes and evidence.
[DOCX File , 52 KB - aging_v5i3e38864_app4.docx]

Multimedia Appendix 5
CONSORT 2010 checklist for Cluster Trials.
[PDF File (Adobe PDF File), 199 KB - aging_v5i3e38864_app5.pdf]

References

1. Pu L, Moyle W, Jones C, Todorovic M. The effectiveness of social robots for older adults: A systematic review and meta-analysis of randomized controlled studies. Gerontologist 2019 Jan 09;59(1):e37-e51. [doi: 10.1093/geront/gny046] [Medline: 29897445]
2. Abbott R, Orr N, McGill P, Whear R, Bethel A, Garside R, et al. How do "robopets" impact the health and well-being of residents in care homes? A systematic review of qualitative and quantitative evidence. Int J Older People Nurs 2019 Sep;14(3):e12239 [FREE Full text] [doi: 10.1111/opn.12239] [Medline: 31070870]
3. Jøranson N, Pedersen I, Rokstad AMM, Ihlebæk C. Effects on symptoms of agitation and depression in persons with dementia participating in robot-assisted activity: A cluster-randomized controlled trial. J Am Med Dir Assoc 2015 Oct 01;16(10):867-873. [doi: 10.1016/j.jamda.2015.05.002] [Medline: 26096582]
4. Wada K, Shibata T, Saito T, Sakamoto K, Tanie K. Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In: Proceedings of the 2005 IEEE International Conference on Robotics and Automation. 2005 Presented at: 2005 IEEE International Conference on Robotics and Automation; April 18-22, 2005; Barcelona, Spain. [doi: 10.1109/robot.2005.1570535]
5. Robinson H, Macdonald B, Kerse N, Broadbent E. The psychosocial effects of a companion robot: a randomized controlled trial. J Am Med Dir Assoc 2013 Sep;14(9):661-667. [doi: 10.1016/j.jamda.2013.02.007] [Medline: 23545466]
6. Saito T, Shibata T, Wada K, Tanie K. Relationship between interaction with the mental commit robot and change of stress reaction of the elderly. In: Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics

XSL•FO
RenderX

and Automation. Computational Intelligence in Robotics and Automation for the New Millennium (Cat. No.03EX694). 2003 Presented at: 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation; July 16-20, 2003; Kobe, Japan p. 119-124. [doi: 10.1109/cira.2003.1222074]
7. Petersen S, Houston S, Qin H, Tague C, Studley J. The utilization of robotic pets in dementia care. J Alzheimers Dis 2017;55(2):569-574 [FREE Full text] [doi: 10.3233/JAD-160703] [Medline: 27716673]
8. Moyle W, Jones C, Murfield J, Thalib L, Beattie E, Shum D, et al. Using a therapeutic companion robot for dementia symptoms in long-term care: reflections from a cluster-RCT. Aging Ment Health 2019 Mar 28;23(3):329-336. [doi: 10.1080/13607863.2017.1421617] [Medline: 29282989]
9. Bradwell HL, Edwards KJ, Winnington R, Thill S, Jones RB. Companion robots for older people: importance of user-centred design demonstrated through observations and focus groups comparing preferences of older people and roboticists in South West England. BMJ Open 2019 Sep 26;9(9):e032468 [FREE Full text] [doi: 10.1136/bmjopen-2019-032468] [Medline: 31558461]
10. Bradwell H, Edwards K, Shenton D, Winnington R, Thill S, Jones RB. User-centered design of companion robot pets involving care home resident-robot interactions and focus groups with residents, staff, and family: qualitative study. JMIR Rehabil Assist Technol 2021 Nov 01;8(4):e30337 [FREE Full text] [doi: 10.2196/30337] [Medline: 34723825]
11. Bradwell HL, Winnington R, Thill S, Jones RB. Morphology of socially assistive robots for health and social care: A reflection on 24 months of research with anthropomorphic, zoomorphic and mechanomorphic devices. 2021 Presented at: 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN); August 08-12, 2021; Vancouver, BC, Canada. [doi: 10.1109/RO-MAN50785.2021.9515446]
12. Bradwell H, Winnington R, Thill S, Jones RB. Prioritising Design Features for Companion Robots Aimed at Older Adults: Stakeholder Survey Ranking Results. In: Social Robotics. ICSR 2021. Lecture Notes in Computer Science, vol 13086. Cham: Springer; 2021:774-779.
13. Lancashire care homes get robotic pets to help residents with dementia. Lancashire Telegraph. 2019. URL: https://www.lancashiretelegraph.co.uk/news/17957962.lancashire-care-homes-get-robotic-pets-help-residents-dementia/ [accessed 2020-05-21]
14. Bradwell HL, Winnington R, Thill S, Jones RB. Longitudinal diary data: Six months real-world implementation of affordable companion robots for older people in supported living. In: HRI '20: Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. 2020 Presented at: ACM/IEEE International Conference on Human-Robot Interaction; March 23-26, 2020; Cambridge, United Kingdom. [doi: 10.1145/3371382.3378256]
15. Koh WQ, Ang FXH, Casey D. Impacts of low-cost robotic pets for older adults and people with dementia: Scoping review. JMIR Rehabil Assist Technol 2021 Feb 12;8(1):e25340 [FREE Full text] [doi: 10.2196/25340] [Medline: 33497349]
16. Pike J, Picking R, Cunningham S. Robot companion cats for people at home with dementia: A qualitative case study on companotics. Dementia (London) 2021 May 16;20(4):1300-1318. [doi: 10.1177/1471301220932780] [Medline: 32668978]
17. Hudson J, Ungar R, Albright L, Tkatch R, Schaeffer J, Wicker ER. Robotic pet use among community-dwelling older adults. J Gerontol B Psychol Sci Soc Sci 2020 Oct 16;75(9):2018-2028 [FREE Full text] [doi: 10.1093/geronb/gbaa119] [Medline: 32789476]
18. Picking R, Pike J. Exploring the effects of interaction with a robot cat for dementia sufferers and their carers. 2017 Presented at: 2017 Internet Technologies and Applications (ITA); September 12-15, 2017; Wrexham, United Kingdom. [doi: 10.1109/itecha.2017.8101940]
19. Marsilio J, McKittrick S, Garner M, Umbell L, Maiewski S, Wenos J. Effects of a robotic cat on agitation and quality of life in individuals with dementia in a long- term care facility. JMU Scholarly Commons Physician Assistant Capstone. 2018. URL: https://commons.lib.jmu.edu/pacapstones/35/ [accessed 2022-07-29]
20. Tkatch R, Wu L, MacLeod S, Ungar R, Albright L, Russell D, et al. Reducing loneliness and improving well-being among older adults with animatronic pets. Aging Ment Health 2021 Jul;25(7):1239-1245. [doi: 10.1080/13607863.2020.1758906] [Medline: 32363903]
21. Wexler S, Drury L, Pollak C, Disomma B, Narducci S, Iovino S, et al. Robotic pet therapy improves outcomes for hospitalized older adults. Innovation in Aging 2018;2(Suppl 1):731. [doi: 10.1093/geroni/igy023.2699]
22. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012 Sep 04;345:e5661. [doi: 10.1136/bmj.e5661] [Medline: 22951546]
23. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 2015 Mar 06;350(feb06 1):h391 [FREE Full text] [doi: 10.1136/bmj.h391] [Medline: 25662947]
24. Clark CM, Ewbank DC. Performance of the Dementia Severity Rating Scale. Alzheimer Disease & Associated Disorders 1996;10(1):31-39. [doi: 10.1097/00002093-199603000-00006]
25. Wood S, Cummings JL, Hsu MA, Barclay T, Wheatley MV, Yarema KT, et al. The use of the neuropsychiatric inventory in nursing home residents. Characterization and measurement. Am J Geriatr Psychiatry 2000;8(1):75-83. [doi: 10.1097/00019442-200002000-00010] [Medline: 10648298]
26. Moniz-Cook E, Woods R, Gardiner E, Silver M, Agar S. The Challenging Behaviour Scale (CBS): development of a scale for staff caring for older people in residential and nursing homes. Br J Clin Psychol 2001 Sep;40(3):309-322. [doi: 10.1348/014466501163715] [Medline: 11593958]

XSL•FO
RenderX

27. Strøm BS, Engedal K, Šaltytė Benth J, Grov E. Psychometric evaluation of the Holden Communication Scale (HCS) for persons with dementia. BMJ Open 2016 Dec 13;6(12):e013447 [FREE Full text] [doi: 10.1136/bmjopen-2016-013447] [Medline: 27965255]
28. Measuring Your Impact on Loneliness in Later Life. Campaign to End Loneliness. URL: https://www.campaigntoendloneliness.org/wp-content/uploads/Loneliness-Measurement-Guidance1-1.pdf [accessed 2021-03-01]
29. Moyle W, Jones CJ, Murfield JE, Thalib L, Beattie ERA, Shum DKH, et al. Use of a robotic seal as a therapeutic tool to improve dementia symptoms: A cluster-randomized controlled trial. J Am Med Dir Assoc 2017 Sep 01;18(9):766-773 [FREE Full text] [doi: 10.1016/j.jamda.2017.03.018] [Medline: 28780395]
30. Verhagen SJW, Hasmi L, Drukker M, van Os J, Delespaul PAEG. Use of the experience sampling method in the context of clinical trials. Evid Based Ment Health 2016 Aug;19(3):86-89 [FREE Full text] [doi: 10.1136/ebmental-2016-102418] [Medline: 27443678]
31. Bradwell HL, Johnson CW, Lee J, Winnington R, Thill S, Jones RB. Microbial contamination and efficacy of disinfection procedures of companion robots in care homes. PLoS One 2020 Aug 26;15(8):e0237069 [FREE Full text] [doi: 10.1371/journal.pone.0237069] [Medline: 32845891]
32. Kim SY. Efficacy versus effectiveness. Korean J Fam Med 2013 Jul;34(4):227 [FREE Full text] [doi: 10.4082/kjfm.2013.34.4.227] [Medline: 23904951]
33. Bemelmans R, Gelderblom GJ, Jonker P, de Witte L. Effectiveness of robot Paro in intramural psychogeriatric care: A multicenter quasi-experimental study. J Am Med Dir Assoc 2015 Nov 01;16(11):946-950. [doi: 10.1016/j.jamda.2015.05.007] [Medline: 26115817]
34. Mao H, Kuo C, Huang W, Cummings JL, Hwang T. Values of the minimal clinically important difference for the Neuropsychiatric Inventory Questionnaire in individuals with dementia. J Am Geriatr Soc 2015 Jul;63(7):1448-1452. [doi: 10.1111/jgs.13473] [Medline: 26046666]
35. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
36. Bradwell HL, Winnington R, Thill S, Jones RB. Ethical perceptions towards real-world use of companion robots with older people and people with dementia: survey opinions among younger adults. BMC Geriatr 2020 Jul 14;20(1):244 [FREE Full text] [doi: 10.1186/s12877-020-01641-5] [Medline: 32664904]
37. Brecher DB. Use of a robotic cat to treat terminal restlessness: A case study. J Palliat Med 2020 Mar 01;23(3):432-434. [doi: 10.1089/jpm.2019.0157] [Medline: 31328997]
38. Schulman-Marcus J, Mookherjee S, Rice L, Lyubarova R. New approaches for the treatment of delirium: A case for robotic pets. Am J Med 2019 Jul;132(7):781-782. [doi: 10.1016/j.amjmed.2018.12.039] [Medline: 30738728]
39. Siniscarco MT, Love-Williams C, Burnett-Wolle S. Video conferencing: An intervention for emotional loneliness in long-term care. Activities, Adaptation & Aging 2017 Jun 15;41(4):316-329. [doi: 10.1080/01924788.2017.1326763]
40. McBride V, Adorno A, Monaco A, Ferrini R. Robocats/Robopups: Awakening the Isolated with Robotic Animals. CALTCM. URL: https://tinyurl.com/2s4c92tc [accessed 2021-03-01]
41. Smith N, Towers A, Plamer S, Beecham J, Welch E. Being occupied: supporting 'meaningful activity' in care homes for older people in England. Ageing and Society 2017 Jul 11;38(11):2218-2240. [doi: 10.1017/s0144686x17000678]
42. Wood N, Sharkey A, Mountain G, Millings A. The Paro robot seal as a social mediator for healthy users. 2015 Presented at: 4th International Symposium on New Frontiers in Human-Robot Interaction; April 21-22, 2015; Canterbury, United Kingdom URL: https://www.cs.kent.ac.uk/events/2015/AISB2015/proceedings/hri/4-Wood-theparorobot.pdf
43. Kachouie R, Sedighadeli S, Khosla R, Chu M. Socially assistive robots in elderly care: A mixed-method systematic literature review. International Journal of Human-Computer Interaction 2014 Apr;30(5):369-393. [doi: 10.1080/10447318.2013.873278]
44. Bradwell H, Johnson CW, Lee J, Soler-Lopez M, Jones RB. Potential transmission of SARS-CoV-2 via robot pets in care homes. ResearchGate. 2020. URL: https://tinyurl.com/ncn6ddt3 [accessed 2022-07-29]
45. Birks M, Bodak M, Barlas J, Harwood J, Pether M. Robotic seals as therapeutic tools in an aged care facility: A qualitative study. J Aging Res 2016;2016:8569602 [FREE Full text] [doi: 10.1155/2016/8569602] [Medline: 27990301]
46. Pino M, Boulay M, Jouen F, Rigaud A. "Are we ready for robots that care for us?" Attitudes and opinions of older adults toward socially assistive robots. Front Aging Neurosci 2015;7:141 [FREE Full text] [doi: 10.3389/fnagi.2015.00141] [Medline: 26257646]
47. Lai CKY. The merits and problems of Neuropsychiatric Inventory as an assessment tool in people with dementia and other neurological disorders. Clin Interv Aging 2014;9:1051-1061 [FREE Full text] [doi: 10.2147/CIA.S63504] [Medline: 25031530]

Abbreviations

ITT: intention to treat
JfA: Joy for All
NPI: Neuropsychiatric Inventory
RCT: randomized controlled trial

XSL•FO
RenderX

Edited by T Leung, J Wang; submitted 19.04.22; peer-reviewed by WQ Koh, S Schüssler; comments to author 14.05.22; revised version received 10.06.22; accepted 08.07.22; published 24.08.22.

Please cite as:
Bradwell H, Edwards KJ, Winnington R, Thill S, Allgar V, Jones RB
Implementing Affordable Socially Assistive Pet Robots in Care Homes Before and During the COVID-19 Pandemic: Stratified Cluster Randomized Controlled Trial and Mixed Methods Study
JMIR Aging 2022;5(3):e38864
URL: https://aging.jmir.org/2022/3/e38864
doi:10.2196/38864
PMID:35830959

©Hannah Bradwell, Katie J Edwards, Rhona Winnington, Serge Thill, Victoria Allgar, Ray B Jones. Originally published in JMIR Aging (https://aging.jmir.org), 24.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mobile Videoconferencing for Occupational Therapists' Assessments of Patients' Home Environments Prior to Hospital Discharge: Mixed Methods Feasibility and Comparative Study

Karine Latulippe[1,2], PhD; Dominique Giroux[3,4,5], PhD; Manon Guay[6,7], PhD; Dahlia Kairy[1,8,9], PhD; Claude Vincent[3,10], PhD; Katia Boivin[3,11], MSc; Ernesto Morales[3,10], PhD; Natasa Obradovic[6,7], MA; Véronique Provencher[6,7], PhD

[1]Center for Interdisciplinary Research in Rehabilitation of Metropolitan Montreal, Montréal, QC, Canada
[2]School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada
[3]Department of Rehabilitation, Université Laval, Quebec, QC, Canada
[4]Centre de Recherche du Centre Hospitalier Universitaire de Québec, Quebec, QC, Canada
[5]Centre d'Excellence sur le Vieillissement de Québec, Québec, QC, Canada
[6]School of Rehabilitation, Université de Sherbrooke, Sherbrooke, QC, Canada
[7]Research Center on Aging, Sherbrooke, QC, Canada
[8]School of Rehabilitation, Université de Montréal, Montréal, QC, Canada
[9]Institut Universitaire sur la Réadaptation en Déficience Physique de Montréal, Montréal, QC, Canada
[10]Center for Interdisciplinary Research in Rehabilitation and Social Integration, Quebec, QC, Canada
[11]Centre Hospitalier Universitaire de Québec-Université Laval, Quebec, QC, Canada

Corresponding Author:
Karine Latulippe, PhD
Center for Interdisciplinary Research in Rehabilitation of Metropolitan Montreal
6363, chemin Hudson
bureau 061 Pavillon Lindsay de l'IURDPM
Montréal, QC, H3S 1M9
Canada
Phone: 1 418 435 8541
Email: karine.latulippe@mcgill.ca

Abstract

Background: Occupational therapists who work in hospitals need to assess patients' home environment in preparation for hospital discharge in order to provide recommendations (eg, technical aids) to support their independence and safety. Home visits increase performance in everyday activities and decrease the risk of falls; however, in some countries, home visits are rarely made prior to hospital discharge due to the cost and time involved. In most cases, occupational therapists rely on an interview with the patient or a caregiver to assess the home. The use of videoconferencing to assess patients' home environments could be an innovative solution to allow better and more appropriate recommendations.

Objective: The aim of this study was (1) to explore the added value of using mobile videoconferencing compared with standard procedure only and (2) to document the clinical feasibility of using mobile videoconferencing to assess patients' home environments.

Methods: Occupational therapists assessed home environments using, first, the standard procedure (interview), and then, videoconferencing (with the help of a family caregiver located in patients' homes, using an electronic tablet). We used a concurrent mixed methods design. The occupational therapist's responsiveness to telehealth, time spent on assessment, patient's occupational performance and satisfaction, and major events influencing the variables were collected as quantitative data. The perceptions of occupational therapists and family caregivers regarding the added value of using this method and the nature of changes made to recommendations as a result of the videoconference (if any) were collected as qualitative data, using questionnaires and semistructured interviews.

Results: Eight triads (6 occupational therapists, 8 patients, and 8 caregivers) participated. The use of mobile videoconferencing generally led occupational therapists to modify the initial intervention plan (produced after the standard interview). Occupational therapists and caregivers perceived benefits in using mobile videoconferencing (eg, the ability to provide real-time comments or

XSL•FO
RenderX

feedback), and they also perceived disadvantages (eg, videoconferencing requires additional time and greater availability of caregivers). Some occupational therapists believed that mobile videoconferencing added value to assessments, while others did not.

Conclusions: The use of mobile videoconferencing in the context of hospital discharge planning has raised questions of clinical feasibility. Although mobile videoconferencing provides multiple benefits to hospital discharge, including more appropriate occupational therapist recommendations, time constraints made it more difficult to perceive the added value. However, with smartphone use, interdisciplinary team involvement, and patient participation in the videoconference visit, mobile videoconferencing can become an asset to hospital discharge planning.

International Registered Report Identifier (IRRID): RR2-10.2196/11674

(JMIR Aging 2022;5(3):e24376) doi:10.2196/24376

KEYWORDS

caregivers; feasibility; mixed methods; mobile videoconferencing; mobile phone; occupational therapy; discharge planning; home assessment

Introduction

When planning hospital discharge, it is important to ensure that patients have optimal conditions for a safe return home and that patients' care and services needs are met [1]. The occupational therapist can play an important role in achieving this objective [1] by providing recommendations (eg, technical aids, site planning, care services) to promote the patient's autonomy and safety upon returning home. The home visit is a way for occupational therapists to obtain reliable information about the environment [2], which is essential for making recommendations that support the best fit between the person, their activities, and their environment. A randomized controlled trial [2] found that home predischarge assessment decreased the risk of falls, reduced the number of rehospitalizations, and increased the level of functional independence in patients with hip fractures. However, clinical (eg, patient fatigue or anxiety), organizational (eg, time available), and financial (eg, travel time and costs) constraints limit the implementation of home visits, despite their relevance [1-5].

Alternative means are currently used to assess the home environment when planning hospital discharge, such as interviews [6], consultation of home photos taken by caregivers [7], video [8], and virtual reality [9]. Interviews quickly give an idea of the environmental constraints perceived by the patient and caregivers. The interpretation of the occupational therapist is then based on this self-reported information, without having the opportunity to confirm it through direct observation [3]. Photos provided by the caregivers allow the occupational therapist to observe the patient's environment [10]; however, this observation is dependent on the choice of photos and the angle used by the caregiver. Video also makes it possible to observe the environment [8]; however, similar to photos, it is an asynchronous means, and the occupational therapist's observation is contingent upon what the relatives choose to show. Other methods such as virtual reality [9] and 3D photography [11,12] are currently being explored and are in the experimental stage [13].

Based on a growing body of literature, the use of mobile videoconferencing for remote rehabilitation interventions has potential clinical benefits [10,13,14]. By providing a detailed and real-time view of the home environment, mobile conferencing may help occupational therapists to improve the reliability of the data collected, which in turn guarantees appropriate recommendations. The occupational therapist, therefore, gives instructions to the caregiver who, using the electronic tablet, shows the facilities in the home for which more information is needed. However, empirical evidence is lacking to clinically support its use [15,16]. The aim of this study is (1) to explore the added value of using mobile videoconferencing compared with the standard procedure and (2) to document the clinical feasibility of using an electronic tablet to assess the patient's home environment through videoconferencing.

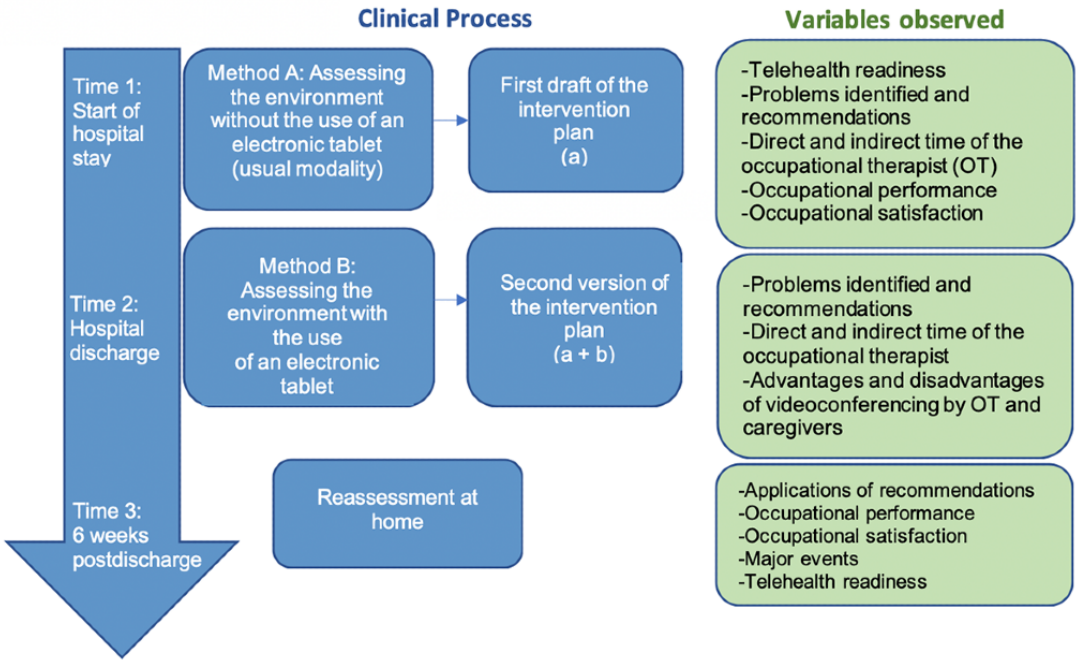
Methods

Design

The methods used for this study are detailed in a published protocol [17]. We conducted a concurrent mixed methods feasibility study to compare 2 home assessment methods (Figure 1). In method A, information about the home environment was collected during an interview with the patient. In method B, evaluation of the home was carried out through mobile videoconferencing using an electronic tablet. For the videoconferencing evaluation, some occupational therapists used their work computer (when the facility allowed the installation of Skype for Business and a webcam was available), others, as well as the patients, used an HP elite pad and iPad tablet with a 3G mobile connection. An electronic tablet was provided to each caregiver with the exception of one caregiver who chose to use her own smartphone. Skype for Business was used for videoconferencing. The home assessment was conducted from the hospital center. The 2 assessment methods were compared to highlight the contribution of mobile videoconferencing to the standard evaluation (*A* versus *A and B*).

XSL•FO
RenderX

Figure 1. Study flowchart. OT: occupational therapist.



Sampling and Recruitment

A purposeful sampling strategy was used in 2 regional hospitals in the province of Quebec, Canada. A triad consisted of (1) an adult patient with a loss of functional autonomy mainly due to physical disability, (2) their caregiver, and (3) the occupational therapist who conducted the assessment. The eligibility criteria for patients were (1) being hospitalized at the time of recruitment and (2) having a return home (including retirement homes) planned. Caregivers had to be able to (1) clearly express themselves orally (in French or English) and (2) walk without technical assistance. The occupational therapist had to have at least one year of clinical experience. Patients were excluded if they (1) had regular home monitoring by an occupational therapist in the community prior to hospitalization and (2) were unable to express themselves in a functional manner. The initial target sample was 18 triads (8 occupational therapists and 18 patient-related dyads).

Data Collection Methods

Qualitative and quantitative data were collected in parallel.

Quantitative Data

We collected data on (1) occupational therapists' receptivity to the use of mobile videoconferencing, using the French-Canadian version of the practitioner and organizational telehealth readiness assessment tools, for which a score >80 indicates that practitioners are well positioned to use telehealth, a score from 60 to 80 indicates that there are factors or elements that may negatively affect telehealth use, and a score <60 indicates that there are barriers to successful telehealth use by practitioners [18]; (2) the time spent evaluating the environment at the time of discharge (discussions, making an appointment with the caregiver, explanations prior to the assessment) using each method (with and without mobile videoconferencing), which was estimated by the occupational therapist; (3) major events after discharge, which were documented with the Social Readjustment Rating Questionnaire [19]; and (4) patient occupational performance and satisfaction was measured using the Canadian Professional Performance Measurement [20].

Qualitative Data

We also collected data on (1) the advantages and disadvantages of using mobile videoconferencing (the individual and semistructured interview guides—occupational therapists and caregivers' versions—addressed previous and current experience with mobile videoconferencing use, the barriers and problems encountered with the use of mobile videoconferencing in the study, and the perceived benefits of adding this method); (2) occupational therapist's professional recommendations, intervention plan, and the follow-up, which were charted using a pretested grid; and (3) the relevance and application of the recommendations, which were documented using a monitoring grid during an interview with the patient, with questions regarding the level of appreciation and barriers to implementation of the recommendations (approximately 20 minutes).

Study Process

A participant's guide was developed for the research assistant and occupational therapists. First, a research assistant invited the occupational therapists (in person or by telephone) to sign the consent form and to complete the French-Canadian version of the practitioner and organizational telehealth readiness assessment tools [18] and sociodemographic data form.

Second, the health care teams and occupational therapists identified patients who would potentially benefit from a home assessment prior to hospital discharge. These patients were offered the opportunity to participate in the study, and

XSL•FO
RenderX

occupational therapists made sure to specify that a refusal would affect neither access to nor the quality of their assessment. If patients decided to participate in the study, the research assistant made an appointment with the patient and their caregiver to discuss the study in order to obtain informed consent and verify that the inclusion and exclusion criteria were met. Subsequently, the research assistant (who also has background in occupational therapy) conducted the Canadian Professional Performance Measurement [20] and collected sociodemographic data.

The occupational therapist conducted the home assessment by interviewing the patient and caregiver (method A). The occupational therapist recorded the time (direct, when the patient and the caregiver were physically or remotely present, and indirect, when the patient and the caregiver were not present) that it took to complete method A and documented any problems identified and recommendations (the first draft of intervention plan).

Next, the occupational therapist made an appointment with the caregiver for the home assessment via mobile videoconferencing (method B), which included the time necessary to pick up the electronic tablet and to teach the caregiver how to use the device. The occupational therapist recorded the time (direct and indirect) that it took to complete method B and modified the intervention plan where relevant.

Finally, the research assistant conducted a semistructured interview with the occupational therapist to identify prior and current experience with mobile videoconferencing, the barriers and problems encountered with the use of mobile videoconferencing in the study, and the perceived benefit of incorporating this method. The occupational therapist also completed the French-Canadian version of the practitioner and organizational telehealth readiness assessment tools [18] a second time. Six weeks after hospital discharge, the research assistant went to the patient's home. In the presence of the caregiver, the research assistant completed the Canadian Professional Performance Measurement [20] again, as well as the Social Readjustment Rating Scale [19]. She also conducted a semistructured interview related to the patient's satisfaction and the applicability of the recommendations that had been given at the time of discharge.

Training

Occupational therapists were not formally trained in the study procedure; however, a step-by-step guide was provided on how to use the electronic tablet and videoconferencing app. A research assistant was available to answer questions and provide further guidance as needed.

Analysis

Descriptive analyses were conducted for the receptivity scores collected from occupational therapists at the beginning (before the first patient) and at the end of the study (after the last patient), patient satisfaction and performance, as well as the type and number of unplanned events (confounding variables). We compared the recommendations from method A with those from the combination of methods A and B by identifying the differences and the nature of the differences (addition, modification). Finally, the application of the occupational therapist's recommendations (methods A and B) by the patient was also evaluated 6 weeks after hospital discharge. Semistructured interviews were conducted, in which the perceived benefits and barriers of mobile videoconferencing were discussed, recorded, and transcribed verbatim. Using analytical questioning [21], the transcriptions were categorized by theme and analyzed by group (interview with occupational therapists, interview with caregivers, and occupational therapist's professional recommendations). We used MAXQDA software (version 2018.1; Verbi GmbH) for analyses. Quantitative and qualitative data were integrated based on 2 analytical questions related to the study objectives: Which results inform us about the added value (or absence thereof) of mobile videoconferencing? Which results inform us about the clinical feasibility of using mobile videoconferencing for the purpose of home environment assessment before hospital discharge?

Ethics Approval

The project was approved by the Research Ethics Committee of the Centre intégré universitaire de santé et de services sociaux de l'Estrie – Centre hospitalier universitaire de Sherbrooke (MP-31-2017-1485) and the Research Ethics Committee of the Quebec University Hospital – Université Laval (2017-3047). Mobile videoconferences were not recorded. Aside from the occupational therapists who performed the home assessment, no one could observe the patient's home.

Results

Participants

Eight triads (6 occupational therapists, 8 patients, and 8 caregivers) were enrolled between April 2017 and April 2019 (Table 1).

The number of triads originally targeted was not reached. To better understand the issues surrounding patient recruitment and the feasibility of using mobile videoconferencing for home assessment, the research team decided to add open questions to the receptivity questionnaire (1) for occupational therapists who used mobile videoconferencing with at least one patient (n=6) and (2) for occupational therapists who participated in the project but who did not recruit patients (n=7) (Multimedia Appendix 1). In 1 instance, the mobile videoconferencing could not be used due to the absence of internet coverage in the municipality; the occupational therapist used video as an alternative. In addition, 1 patient could not be reached 6 weeks after discharge. (The occupational therapist who followed this patient thinks that he may have relocated to a different province.)

XSL•FO
RenderX

Table 1. Participant characteristics.

Group and characteristic	Value (n=8)
Participants (n=8)	
Age (years)	
Mean	79.5
Range	68-90
Sex, n (%)	
Female	4 (50)
Male	4 (50)
In-hospital stay (days)	
Mean	39
Range	10-96
Principal diagnosis leading to hospitalization, n	
Infectious	1
Orthopedics	6
Neurology	1
Medical complications, n	
Delirium	1
Postop shock	1
Comorbidities, n	8
Caregivers (n=8)	
Age (years)	
Mean	58
Range	36-80[a]
Sex, n (%)	
Female	7 (88)
Male	1 (12)
Relationship to the patient, n	
Spouse	3
Child	3
Sibling	1
Familiarity level with technology, n	
Poor	5
Average	1
Good	2
Occupational therapists (n=6)	
Sex, n (%)	
Female	6 (100)
Male	0 (0)
Overall work experience (years)	
Mean	8
Range	1-13
Program work experience (years)	
Mean	6

XSL•FO
RenderX

Group and characteristic	Value (n=8)
Range	0.5-12

[a]n=7, one value is missing.

Difference Between Recommendations Before and After Mobile Videoconferencing

The majority of recommendations made before mobile videoconferencing (n=28) remained applicable afterward (n=25), except for 3 recommendations. Observation of the home environment through mobile videoconferencing made it possible to identify the lack of space required in the room to implement the planned recommendation and some incorrect information gathered during the interview (taps on the right side of the bathtub rather than on the left). These observations led to changes in the intervention plan.

Of the 8 patients, 7 patients' initial intervention plans were modified after videoconferencing (Table 2). The changes were aimed at improving the person's autonomy and safety, reducing the risk of falling, offering a cheaper or simpler solution than the one initially planned, and revising the initial recommendation in light of new information about the home environment that was not discussed during the interview. Occupational therapists also revised their recommendations for a better fit between the patient, their occupation, and the environment. Finally, 4 recommendations were completely modified as a result of videoconferencing—in 3 instances because the initial recommendation was not applicable and, in a fourth instance, because observing the environment made it possible to consider a return home following rehabilitation at an intensive functional rehabilitation unit instead of relocating the patient to a seniors' residence, as recommended initially. In one instance, the intervention plan was not modified after mobile videoconferencing.

Table 2. Modifications of initial intervention plans after mobile videoconferencing.

Type and reason	Changes (n=18), n	Example
Adding recommendations		
Optimizing the person's autonomy	1	The installation of a support bar allowed the patient to transfer on her own rather than accompanied as in the initial recommendation
Optimizing safety and reducing the risk of falling	2	Adding a grab bar and stool in the shower to maximize safety
Offering a cheaper and simpler solution	1	Adding a grab bar to the wall instead of a toilet support frame as in the original recommendation.
Adjusting to new information that was not discussed at the interview	6	Observation of the environment identified 2 potential places for taking meals (a high table with stools and a standard table and chairs in the dining room); due to physical difficulties, using a stool was not recommended
Precision of recommendations		
Ensuring a better match between the patient, the patient's occupation, and the environment	4	Precision about the orientation of the shower bench and support bars initially recommended
Change of recommendations		
The original recommendation was not applicable	3	The safety support on the righthand side is irrelevant given the countertop at an adequate height to the right of the toilet and the lack of space to install the grab bar on that side
Viewing the environment led to a return home	1	Once the environment is seen, there appeared to be no major architectural barrier to a return home if the patient manages to regain autonomy in her transfers and travel with the help of accessories

Perceived Benefits of Using Mobile Videoconferencing to Conduct a Remote Home Assessment

Overall, participants appreciated the use of the tablet and felt that "it adds something" (occupational therapist 6) to the standard evaluation. Specifically, caregivers perceived that the use of mobile videoconferencing allowed occupational therapists to obtain more precise information (Table 3).

According to participants, data collected from interviews can be wrong or incomplete because the caregiver neglects to take into consideration certain aspects. In fact, mobile videoconferencing induced modification of recommendations, such as correcting the provision for assistive devices to match the patient's environment.

> *Yes, in fact the lady had given me inaccurate information about where the bath faucet was located, it was on the opposite side. I recommended a transfer*

XSL•FO
RenderX

bench with the handle on the wrong side. So, I adjusted that. [Occupational therapist 3]

Another advantage of videoconferencing perceived by participants and occupational therapists was the opportunity for therapists to ask questions and provide feedback to the caregiver in real time. Caregivers felt guided in the assessment and identification of measures required, and the occupational therapists were able to document patients' lifestyles and which elements of the environment they wanted to see.

I watched her take some measurements, some of which I may not have thought of not knowing [what the environment was] but since I was watching her, I was able to ask her to measure this and that. It's great, I made a diagram. Seeing what she was doing was of great help to me. [Occupational therapist 4]

Mobile videoconferencing was useful for estimating distances between various elements in the home environment. In addition, caregivers said that the videoconference visit reassured them. All caregivers mentioned that the mobile videoconferencing experience had been positive despite the fact that some encountered a few technical glitches.

Table 3. Main advantages and disadvantages associated with the use of mobile videoconferencing by occupational therapists and caregivers.

Characteristic	Description
Advantages	
Common to occupational therapists and caregivers	• Ability to make comments or provide feedback in real time • Confirming the information obtained by the patient and caregivers • Providing additional information on the patient's lifestyle • Ensuring the best choice of equipment • Making sure that the caregiver is taking the right measurements and reassuring them about how they are doing • Seeing the general condition of the environment (eg, cleanliness) • Avoiding travel expenses and time
For occupational therapists	• Discovering unanticipated barriers • Dissipating remaining doubt and avoiding mistakes • The involvement of the caregiver helps the patient to remember the recommendations • Improving communication between the occupational therapist and the caregiver • Promoting discussion between the occupational therapist and the patient if the latter participates in mobile videoconferencing • For patients transferred to the intensive functional rehabilitation unit, mobile videoconferencing makes it possible to specify the rehabilitation objectives • Seeing details and offering more specific recommendations • Determining the pertinence of a home visit by the CLSC[a] occupational therapist
For caregivers	• Allowing the occupational therapist to identify problems that the caregiver had not thought of • Feeling guided in the return home process • Allowing patients to reconnect with their home and reflect on their return home • Avoiding the need to explain everything • Providing recommendations that don't need tweaking • Reassuring the caregiver
Disadvantages	
Common to occupational therapists and caregivers	• Videoconferencing requires being comfortable with technology • Videoconferencing requires additional time and availability of caregivers • Videoconferencing constitutes additional stress for caregivers who are uncomfortable with taking measurements or using the tablet
For occupational therapists	• Inability to observe the interaction between person, occupation, and environment • No overview such as during a home visit in person • Inaccessible if there is no Internet coverage in the municipality • More time consuming than an interview
For caregivers	—[b]

[a]CLSC: centres locaux de services communautaires (local community service centers).

[b]There were no other perceived disadvantages.

XSL•FO
RenderX

Disadvantages of Using Mobile Videoconferencing to Conduct a Remote Home Assessment

One occupational therapist reported that mobile videoconferencing does not show the interaction between the person, the environment, and the person's activities. In addition, mobile videoconferencing requires more time than interview assessment. Five caregivers mentioned that there were no disadvantages to mobile videoconferencing. The concerns brought up by caregivers were the same as those identified by occupational therapists (ie, stress of having to take measurements or use a tablet, especially if they are not comfortable with the device). Caregivers also highlighted the extra time that it took to pick up the equipment, undergo training, and conduct the videoconference.

Perception of the Added Value of Mobile Videoconferencing

Mobile videoconferencing provided added value according to 3 of 5 occupational therapists. These 3 occupational therapists expressed readiness to promote the use of mobile videoconferencing to their peers.

> *Mobile videoconferencing takes longer to complete, but the recommendations are more specific. The ratio of time to what mobile videoconferencing requires in terms of logistics versus what it provides in terms of intervention offers added value.* [Occupational therapist 3]

> *For (another) patient, it helped define specific goals for her rehabilitation. It ensures that the recommendations are correct and feasible. Occupational therapists are often told that recommendations don't work. Mobile videoconferencing is not pertinent in all cases but when it applies, it really offers added value. It applies when the occupational therapist has doubts about what the patient said.* [Occupational therapist 4]

Another occupational therapist, who was part of 2 triads, also failed to perceive any added value associated with the use of mobile videoconferencing.

> *The changes that the mobile videoconferencing made to the intervention plan were not essential to leave. They were aimed more at optimizing safety and could have been done by the occupational therapist at home. Although the mobile videoconferencing is more concrete than the interview and interesting, the time invested, and the minor changes made to the intervention plan mean that there is no added value.* [Occupational therapist 5]

Changes in Satisfaction and Occupational Performance

There was a clinical difference between hospital stay and postdischarge patient performance scores (hospital: mean 4.0, SD 2.7; postdischarge: mean 6.2, SD 2.8) and satisfaction scores (hospital: mean 4.1, SD 3.1; postdischarge: mean 7.1, SD 2.1). A change of 2 points is considered a clinically relevant improvement or deterioration [22].

Time Required for Mobile Videoconferencing

The mean direct time that occupational therapists reported having spent evaluating the environment through videoconferencing at the time of discharge (discussions, making an appointment with the caregiver, providing explanations prior to the assessment) was 104 minutes (SD 74). The mean indirect time (environment evaluation) was 64 minutes (SD 87).

Occupational Therapists' Receptivity to Mobile Videoconferencing

Assessment of the receptivity of occupational therapists who had recruited at least one patient indicated that there were barriers to successful telehealth use by practitioners (Table 4).

Table 4. Receptivity of occupational therapists who recruited at least one patient.

French-Canadian version of the Practitioner and Organizational Telehealth Readiness Assessment section	Score, mean (SD)	
	Before intervention (n=6)	Postdischarge (n=5)
Total score (out of 85)	51 (10)	56 (9)
In order to meet the requirements for core readiness (out of 15)	7 (1)	8 (2)
In order to meet the requirements for engagement readiness (out of 35)	25 (3)	26 (3)
In order to meet the requirements for structural readiness (out of 35)	20 (7)	22 (5)

Factors Influencing the Choice to Use Mobile Videoconferencing

Several factors appeared to influence whether or not mobile videoconferencing was used by occupational therapists who recruited at least one participant.

One occupational therapist mentioned that, due to her workload, she could not always prioritize mobile videoconferencing over other tasks and did not always have time to do it. The necessary caregiver training on mobile videoconferencing also added to the time constraints associated with this method. Consequently, the occupational therapist's perception of the time that training would take, dependent on whether or not the caregiver was comfortable with the technology, influenced their choice. According to the occupational therapists, meetings with caregivers to introduce mobile videoconferencing, scheduling virtual visits, and material recovery added to their workload, as well as, that of caregivers. Some occupational therapists doubted their ability to teach caregivers how to use mobile videoconferencing, insofar as this required availability, motivation, and collaboration. As reported by occupational

XSL•FO
RenderX

therapists, a number of families refused to engage in mobile videoconferencing because the process seemed too cumbersome. However, for 6 of the caregivers, logistics were not a problem. In addition, some of the occupational therapists felt that meetings with caregivers involved discussions that went beyond mobile videoconferencing and therefore, in a context of limited time, this aspect is a challenge in terms of feasibility.

> *Sometimes it's harder to get someone to go film or have a caregiver who is in step with current technologies.* [Occupational therapist 3]

> *You see the person, you do not just fix it and then move on to something else … she told me a lot of things and then they also have difficult things to do with them personally.* [Occupational therapist 1]

Occupational therapists sometimes anticipated the fact that the patient would be discharged from hospital before they had time to do the mobile videoconferencing or that it is not pertinent in view of the patient's transfer to a rehabilitation unit. Finally, occupational therapists' daytime work schedule did not match the availability of some caregivers, in which case, mobile videoconferencing was not considered.

Caregiver Level of Comfort With Technology and Mobile Videoconferencing Training

Based on their own perceptions, caregivers' comfort level with technology was poor (n=2), moderate (n=4), and good (n=2). Most felt that the training offered by the occupational therapist and the 2-page instruction booklet they were given helped them to learn how to use the tablet.

> *I was afraid I might not to be able to do it, but with that short training, it seemed simple enough and I enjoyed trying to help.* [Caregiver 5]

For one caregiver, however, the training was not sufficient. This caregiver used help from a third party (siblings) during the videoconference. Two other caregivers received help from a third, although their levels of comfort with the technology were moderate and high.

> *It didn't take long; the hardest part was to learn how to operate the tablet and finally it was my sister who turned it on because I had already forgotten how it worked...I'm not used to that myself.* [Caregiver 2]

Perceptions of the Relevance of Home Assessment, Mobile Videoconferencing, and Recruitment Difficulties

Relevance of Home Assessment in Hospital Discharge

Of the occupational therapists who participated in the project but did not recruit patients, 5 occupational therapists considered home assessment prior to hospital discharge to be pertinent.

> *It's important for the safety of the patient and in the prevention of falls, in the maintenance of autonomy also.* [Occupational therapist G]

> *It's an integral part of my job.* [Occupational therapist B]

However, 2 occupational therapists believed that it was the community occupational therapist and not the hospital occupational therapist's role to do the home assessment.

> *I think it's the role of the occupational therapist at the CLSC [centres locaux de services communautaires (local community service center)] to do the home assessment because she has that expertise.* [Occupational therapist D]

Trust in the Interview as a Home Assessment Method

According to most occupational therapists, the amount of trust that can be placed in an interview method depends on the patient. If the patient has no cognitive impairments and the family confirms the information, then it can be relied upon. Conversely, the method cannot be used with patients who have impaired memory or difficulty expressing themselves. The method is even less reliable if a caregiver is not present, which was mentioned by one occupational therapist, who also stressed the possible discrepancy between patient, patient family, and professional perceptions.

> *We are confused by the patient's speech. For example, the patient considers that his home allows to circulate with a walker while a professional would judge the opposite following assessment.* [Occupational therapist C]

In the opinion of some occupational therapists, when doubt exists, photos can be requested from the family or a referral sent to the CLSC occupational therapist. However, there may be a significant delay if the home assessment is done by the CLSC occupational therapist due to their own workload.

Prerequisites for the Use of Mobile Videoconferencing by Occupational Therapists

Many occupational therapists (n=3) commented that they did not have the necessary prerequisites to use mobile videoconferencing (ie, good knowledge of how to use the tablet, ability to solve technical contingencies, and ability to teach the family how to use it). One occupational therapist believed that with good training she could manage. The others felt comfortable using mobile videoconferencing (n=3).

Profile of Patients Who Could Benefit From Mobile Videoconferencing

According to occupational therapists, the patients who would benefit from mobile videoconferencing are patients who have permanent motor disorders, who are already known to the therapist, who are young adults, who are alone, who need a walker in the home, who have cognitive impairments and need to be tested in conditions that are similar to what they are used to, or whose entourage is comfortable with technology and available. Some occupational therapists said that this patient profile is quite common in practical settings, while others disagreed.

Reasons for Nonrecruitment

Finally, in order to explain the reasons why they were unable to recruit patients in the context of the project, occupational therapists mentioned the movement of staff, the impression that

XSL•FO
RenderX

it would be asking too much of the caregiver, the lack of time, the difficulty of coordinating the availability of caregivers with their own, having caregivers at ease with technology, the perception of duplicating work with the CLSC, thinking of recruiting patients, having patients who correspond to the inclusion criteria, and work overload.

Problems Encountered With the Use of Mobile Videoconferencing

Some technical problems were encountered during the study. Communication with the clinician was generally adequate. The sound and the image were judged to be clear by all the participants. With the exception of the lack of internet coverage in the municipality where one patient resided, the technical problems did not prevent the use of mobile videoconferencing or the home evaluation and were not raised as being inconvenient for patients using mobile videoconferencing.

Participant Suggestions on Improving Home Assessment Using Mobile Videoconferencing

One occupational therapist suggested that it would be useful to record the videoconference visit and subsequently review the assessment (as needed or depending on the patient's progress). A caregiver also recommended that the videoconference visit be recorded and available to other professionals. She was again asked about her environment in the rehabilitation unit following her stay at the unit where the initial assessment took place and felt that she was duplicating what had already been done. Another occupational therapist suggested that the interdisciplinary team should be involved in the videoconference. First responders, often the social worker, could explore the possibility of doing the mobile videoconferencing with the patient's family even before the occupational therapist receives the referral in order to address the time constraints of short stay. In addition, it may be pertinent for the physiotherapist to see the walking distances between the home and the parking lot and inside the home, and for the social worker to verify the safety and cleanliness of the home environment. Finally, one occupational therapist conducted the mobile videoconference together with the patient. She explained that the patient was able to provide details of her lifestyle and this experience motivated her in her rehabilitation because she was able to visualize what her return home would be like. This occupational therapist recommended patient participation in mobile videoconferencing.

Discussion

Principal Results

The use of mobile videoconferencing after the interview generally led occupational therapists to modify their initial intervention plan. Most changes were considered by occupational therapists to be minor inasmuch as they were expected to have little impact on a safe return home. However, 3 assistive devices recommended after the interview raised some issues after discharge. In addition, based on mobile videoconferencing, the decision of the interdisciplinary team and that of the patient himself to transfer to a seniors' residence was modified, and the patient returned home upon discharge instead. This is a clinically important point. Unimplementable recommendations (such as the 3 assistive devices mentioned above) can interfere with older adults' ability to age in their homes, and a change in home environment is no small matter in a person's life.

Overall, the perceived advantages of mobile videoconferencing for occupational therapists and caregivers exceeded the disadvantages; however, the nature of the disadvantages—time required to conduct mobile videoconferencing (meeting planning, tablet training, equipment loan, virtual visit) combined with the increased workload perceived by occupational therapists, intervention priorities such as pressure injury, availability of caregivers on working hours, and the short length of stay—do not support its use. More specifically, the perceived reliability of data collected through interviews and the short time required for interviews led occupational therapists to prefer interviews as an evaluation method. This is consistent with the conclusions from a scoping review [13] on the use of information and communication technology for home assessment. Our study highlighted that mobile videoconferencing is considered beneficial by occupational therapists when the patient has a cognitive impairment and a caregiver is not available, both of which reduce the reliability of data collected through interviews. However, for individuals with cognitive impairment, it is very important to observe their interaction with their home environment, and mobile videoconferencing used in the manner described in this study does not allow for this interaction to be seen [23]. Also, in our study, availability and motivation of caregivers were identified as prerequisites for the use of mobile videoconferencing by occupational therapists.

Another clinically relevant finding was that mobile videoconferencing required increased involvement on the part of caregivers in discharge planning. This appears to be an advantage for improving communication between the clinician and caregivers, thereby increasing the probability that the caregiver will implement the occupational therapist's recommendations. In contrast, some occupational therapists, including those who did not recruit a patient, felt uncomfortable burdening caregivers with this task. In fact, some eligible patients were not part of the study because the caregiver declined to participate due to their busy schedule. Knowing that caregivers are at risk of exhaustion [24], clinicians may have been reluctant to add to their burden of care. The family caregivers enrolled in the study, who may arguably be more available and interested in the project, commented that the logistics surrounding mobile videoconferencing had not been a problem. They said that the mobile videoconferencing had reassured them and that they appreciated being guided by the occupational therapist to make the measurements. Holland and colleagues [25] reported that seeing the clinicians on video made caregivers feel as if they were at home with them, which facilitated interactions. Chi and Dimiris [26] also found that caregivers felt more involved in the process. Therefore, mobile videoconferencing can be perceived as a burden by some caregivers and as a facilitator by others.

Some feasibility issues may explain recruitment difficulties and, therefore, will have an impact on the choice to use mobile

XSL•FO
RenderX

videoconferencing or not. Based on the Telehealth Readiness Assessment questionnaire [18], there was a degree of reluctance with respect to telehealth. These findings are not consistent with those of a study [27] in which clinicians were reported to be supportive of more frequent use of the telecommunication system. However, our results may be influenced by clinicians' perceived openness of their workplace to telehealth. Indeed, in the Telehealth Readiness Assessment questionnaire [18], almost half of the points (40 out of 100) pertain to how clinicians perceived the receptivity of the institution. In one study [15], occupational therapists reported that they needed more training in communication technology use but organizational constraints were a barrier [15]. This is consistent with our finding that many occupational therapists did not have the skills to use mobile videoconferencing or to show caregivers how to use the technology. This perception of a lack of technological skills, combined with occupational therapists' perceptions that caregivers who are less familiar with technology would require more time, may explain why they favored the involvement of caregivers who are familiar with the technology. Our conclusions are consistent with those of Ninnis and al [13], who suggested that therapists consider the use of mobile apps to be appropriate for some patients but not those who are less confident or less able to use new technologies. However, in our study. it does not appear that the caregivers' level of comfort with the technology affected its use.

Future Directions

Some occupational therapists and caregivers suggested that the use of caregivers' own smartphones, despite potential confidentiality issues, would allow for a better videoconferencing experience. Smartphones are becoming more and more popular among people aged 65 and over [28]. In addition to precluding the need for mobile videoconferencing training, the use of their own device would eliminate the need for caregivers to come to the hospital to pick up equipment. We are currently working with engineers on making personal smartphones safe and simple to use (only one button to press), with options to measure distances between home facilities through screenshots. Another suggestion made by one occupational therapist was to involve patients in the videoconference, which is in line with shared decision-making and patient-centered approaches [29-31]. The involvement of a social worker and physiotherapist could also help to gather further information (presence of an interior and exterior staircase for example) during the virtual visit and thus optimize hospital discharge planning (such as planning the need for assistance with mobility). We suggest that future studies compare standard assessment (interview), videoconference, and in-person visits of the home environment with the patient in terms of benefits and clinical, ethical, and financial issues [32,33]. It would also be of interest to document the clinical reasoning behind the decision whether or not to assess the home environment, through mobile videoconferencing or otherwise, in order to guide occupational therapists on the best methods to use for this and on how to best use their time [5].

Limitations

This study has some limitations. First, we had fewer participants than desired. The recruitment difficulties encountered during the study underscore the need to make organizational changes to support the use of mobile videoconferencing in routine care. Nevertheless, the added value perceived by participants as well as the opportunity to obtain additional and more appropriate recommendations suggest the relevance of using mobile videoconferencing. Second, it would have been relevant to further document the occupational therapists' and caregivers' level of comfort with technology use in order to better understand how it influenced occupational therapists' receptivity and participant recruitment. Occupational therapists were not asked to recruit the ideal candidate, but a participant selection bias cannot be excluded because of workload concerns. To reduce their workload, they may have been inclined to select patients with family caregivers who were comfortable with the technology or who were motivated to use videoconferencing. Moreover, the analysis was performed by one person (KL). However, the interviews were transcribed verbatim, and 2 co-authors who participated in the interviews (KB, MG) attested to the consistency between themes and interviews. Finally, the start of the COVID-19 pandemic occurred in the period between the study's completion and its publication, which may also impact the results as the pandemic forced occupational therapists and the general population to learn about, if not improve, their technological proficiency and to use mobile video conferencing more frequently.

Conclusions

Clinical feasibility issues were found when using mobile videoconferencing to support hospital discharge planning. Although mobile videoconferencing provides multiple benefits, such as more appropriate occupational therapist recommendations, the inconveniences, such as time constraints, make it difficult to perceive the added value of this method. However, it was suggested that having caregivers use their own smartphone, involvement of the interdisciplinary team, and patient participation in the videoconference would mitigate these inconveniences.

Acknowledgments

We would like to thank the patients, the caregivers, and the occupational therapists who participated in the study. The contributions of the students from the School of Rehabilitation at the University of Sherbrooke (G Fortin, F Gagnon, M Bruneau-Cossette, P Prince) in drafting this manuscript were much appreciated. We also thank the Quebec Network for Research on Aging and Rehabilitation Research Network for their financial contribution to this multicenter project.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Added open questions to the receptivity questionnaire for occupational therapists who did not recruit a patient.
[PDF File (Adobe PDF File), 53 KB - aging_v5i3e24376_app1.pdf]

References

1. Davis AJ, Mc Clure P. An exploratory study of discharge planning home visits within an Irish context – investigating nationwide practice and nationwide perspectives. IJOT 2019 Nov 04;47(2):95-113. [doi: 10.1108/ijot-10-2018-0015]
2. Lockwood KJ, Harding KE, Boyd JN, Taylor NF. Predischarge home visits after hip fracture: a randomized controlled trial. Clin Rehabil 2019 Apr 14;33(4):681-692. [doi: 10.1177/0269215518823256] [Medline: 30642194]
3. Hoffmann T, Russell T. Pre-admission orthopaedic occupational therapy home visits conducted using the internet. J Telemed Telecare 2008 Mar 01;14(2):83-87. [doi: 10.1258/jtt.2007.070808] [Medline: 18348754]
4. Harris S, James E, Snow P. Predischarge occupational therapy home assessment visits: towards an evidence base. Aust Occup Ther J 2008 Jun;55(2):85-95. [doi: 10.1111/j.1440-1630.2007.00684.x] [Medline: 20887443]
5. Godfrey M, Cornwell P, Eames S, Hodson T, Thomas T, Gillen A. Pre-discharge home visits: a qualitative exploration of the experience of occupational therapists and multidisciplinary stakeholders. Aust Occup Ther J 2019 Jun 10;66(3):249-257. [doi: 10.1111/1440-1630.12561] [Medline: 30740713]
6. Robertson L, Blaga L. Occupational therapy assessments used in acute physical care settings. Scand J Occup Ther 2013 Mar 30;20(2):127-135. [doi: 10.3109/11038128.2012.737369] [Medline: 23110547]
7. Daniel H, Oesch P, Stuck A, Born S, Bachmann S, Schoenenberger A. Evaluation of a novel photography-based home assessment protocol for identification of environmental risk factors for falls in elderly persons. Swiss Med Wkly 2013 Nov 12;143:w13884-w13886 [FREE Full text] [doi: 10.4414/smw.2013.13884] [Medline: 24222638]
8. Romero S, Lee MJ, Simic I, Levy C, Sanford J. Development and validation of a remote home safety protocol. Disabil Rehabil Assist Technol 2018 Feb 22;13(2):166-172. [doi: 10.1080/17483107.2017.1300345] [Medline: 28326967]
9. Threapleton K, Newberry K, Sutton G, Worthington E, Drummond A. Virtually home: exploring the potential of virtual reality to support patient discharge after stroke. Br J Occup Ther 2016 Jul 28;80(2):99-107. [doi: 10.1177/0308022616657111]
10. Sanford JA, Pynoos J, Tejral A, Browne A. Development of a comprehensive assessment for delivery of home modifications. Phys Occup Ther Geriatr 2009 Jul 28;20(2):43-55. [doi: 10.1080/j148v20n02_03]
11. Money AG, Atwal A, Young KL, Day Y, Wilson L, Money KG. Using the technology acceptance model to explore community dwelling older adults' perceptions of a 3D interior design application to facilitate pre-discharge home adaptations. BMC Med Inform Decis Mak 2015 Aug 26;15(1):73-15 [FREE Full text] [doi: 10.1186/s12911-015-0190-2] [Medline: 26307048]
12. Kim JB, Brienza DM. Development of a remote accessibility assessment system through three-dimensional reconstruction technology. J Rehabil Res Dev 2006;43(2):257-272 [FREE Full text] [doi: 10.1682/jrrd.2004.12.0163] [Medline: 16847792]
13. Ninnis K, Van Den Berg M, Lannin NA, George S, Laver K. Information and communication technology use within occupational therapy home assessments: a scoping review. Br J Occup Ther 2018 Jul 12;82(3):141-152. [doi: 10.1177/0308022618786928]
14. Sanford JA, Jones M, Daviou P, Grogg K, Butterfield T. Using telerehabilitation to identify home modification needs. Assist Technol 2004 Jun 30;16(1):43-53. [doi: 10.1080/10400435.2004.10132073] [Medline: 15357147]
15. Chedid RJ, Dew A, Veitch C. Barriers to the use of information and communication technology by occupational therapists working in a rural area of New South Wales, Australia. Aust Occup Ther J 2013 Jun 19;60(3):197-205. [doi: 10.1111/1440-1630.12016] [Medline: 23730785]
16. Hung KN G, Fong KN. Effects of telerehabilitation in occupational therapy practice: a systematic review. Hong Kong J Occup Ther 2019 May 27;32(1):3-21. [doi: 10.1177/1569186119849119]
17. Latulippe K, Provencher V, Boivin K, Vincent C, Guay M, Kairy D, et al. Using an electronic tablet to assess patients' home environment by videoconferencing prior to hospital discharge: protocol for a mixed-methods feasibility and comparative study. JMIR Res Protoc 2019 Jan 14;8(1):e11674 [FREE Full text] [doi: 10.2196/11674] [Medline: 31344677]
18. Légaré E, Vincent C, Lehoux P, Anderson D, Kairy D, Gagnon M, et al. Developing and validating the French-Canadian version of the practitioner and organizational telehealth readiness assessment tools. J Telemed Telecare 2010 Apr 12;16(3):140-146. [doi: 10.1258/jtt.2009.090514] [Medline: 20386034]
19. Holmes TH, Rahe RH. The social readjustment rating scale. J Psychosom Res 1967 Aug;11(2):213-218. [doi: 10.1016/0022-3999(67)90010-4]
20. Chan CC, Lee TM. Validity of the Canadian occupational performance measure. Occup Ther Int 2006 Aug;4(3):231-249. [doi: 10.1002/oti.58]
21. Paillé P, Mucchielli A. L'analyse qualitative en sciences humaines et sociales 3e éd. Paris: Armand Colin; 2012.

XSL•FO
RenderX

22. Carswell A, McColl MA, Baptiste S, Law M, Polatajko H, Pollock N. The Canadian Occupational Performance Measure: a research and clinical literature review. Can J Occup Ther 2004 Oct;71(4):210-222. [doi: 10.1177/000841740407100406] [Medline: 15586853]
23. Lai FH, Yan EW, Yu KK. Home-based evaluation of executive function (Home-MET) for older adults with mild cognitive impairment. Arch Gerontol Geriatr 2020 Mar;87:104012. [doi: 10.1016/j.archger.2020.104012] [Medline: 31968279]
24. Borghi AC, Castro VCD, Marcon SS, Carreira L. Overload of families taking care of elderly people with Alzheimer's Disease: a comparative study. Rev. Latino-Am. Enfermagem 2013 Jul;21(4):876-883. [doi: 10.1590/s0104-11692013000400007]
25. Holland DE, Vanderboom CE, Ingram CJ, Dose AM, Borkenhagen LM, Skadahl P, et al. The feasibility of using technology to enhance the transition of palliative care for rural patients. Comput Inform Nurs 2014 Jun;32(6):257-266. [doi: 10.1097/CIN.0000000000000066] [Medline: 24814998]
26. Chi N, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015 Jan;21(1):37-44 [FREE Full text] [doi: 10.1177/1357633X14562734] [Medline: 25475220]
27. Cristancho-Lacroix V, Wrobel J, Cantegreil-Kallen I, Dub T, Rouquette A, Rigaud A. A web-based psychoeducational program for informal caregivers of patients with Alzheimer's disease: a pilot randomized controlled trial. J Med Internet Res 2015 May 12;17(5):e117 [FREE Full text] [doi: 10.2196/jmir.3717] [Medline: 25967983]
28. Vieillir à l'ère numérique. Créneau des entreprises collectives en service aux personnes âgées autonomes ou en perte d'autonomie. 2018. URL: https://creneaupaapa.uqam.ca/wp-content/uploads/2018/10/netendances-2018_veillir_avec_le_numerique.pdf [accessed 2022-06-22]
29. Stiggelbout AM, Van der Weijden T, De Wit MPT, Frosch D, Légaré F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012 Jan 27;344(jan27 1):e256-e256. [doi: 10.1136/bmj.e256] [Medline: 22286508]
30. Roelands M, Van Oost P, Stevens V, Depoorter A, Buysse A. Clinical practice guidelines to improve shared decision-making about assistive device use in home care: a pilot intervention study. Patient Educ Couns 2004 Nov;55(2):252-264. [doi: 10.1016/j.pec.2003.10.002] [Medline: 15530763]
31. Rose A, Rosewilliam S, Soundy A. Shared decision making within goal setting in rehabilitation settings: a systematic review. Patient Educ Couns 2017 Jan;100(1):65-75. [doi: 10.1016/j.pec.2016.07.030] [Medline: 27486052]
32. Atwal A, Caldwell K. Ethics, occupational therapy and discharge planning: four broken principles. Aust Occ Ther J 2003 Dec;50(4):244-251. [doi: 10.1046/j.1440-1630.2003.00374.x]
33. Green S, Lambert R. A systematic review of health economic evaluations in occupational therapy. Br J Occup Ther 2016 Jul 20;80(1):5-19. [doi: 10.1177/0308022616650898]

Abbreviations

CLSC: centres locaux de services communautaires (local community service centers)

Edited by J Wang; submitted 16.09.20; peer-reviewed by K Laver, M Naeemabadi; comments to author 03.11.20; revised version received 29.01.21; accepted 25.03.22; published 05.07.22.

Please cite as:
Latulippe K, Giroux D, Guay M, Kairy D, Vincent C, Boivin K, Morales E, Obradovic N, Provencher V
Mobile Videoconferencing for Occupational Therapists' Assessments of Patients' Home Environments Prior to Hospital Discharge: Mixed Methods Feasibility and Comparative Study
JMIR Aging 2022;5(3):e24376
URL: https://aging.jmir.org/2022/3/e24376
doi:10.2196/24376
PMID:35787486

©Karine Latulippe, Dominique Giroux, Manon Guay, Dahlia Kairy, Claude Vincent, Katia Boivin, Ernesto Morales, Natasa Obradovic, Véronique Provencher. Originally published in JMIR Aging (https://aging.jmir.org), 05.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

A Conceptual Model of Experiences With Digital Technologies in Aging in Place: Qualitative Systematic Review and Meta-synthesis

Mareike Hechinger[1], MScN, RN; Diana Hentschel[1], BSc; Christine Aumer[1], PhD, RN; Christian Rester[1], PhD, RN

Deggendorf Institute of Technology, Faculty of Applied Healthcare Science, Deggendorf, Germany

Corresponding Author:
Mareike Hechinger, MScN, RN
Deggendorf Institute of Technology
Faculty of Applied Healthcare Science
Land-Au 27
Deggendorf, 94469
Germany
Phone: 49 991 3615 156
Email: mareike.hechinger@th-deg.de

Abstract

Background: Older adults with chronic illnesses or dependency on care who strive to age in place need support and care depending on their illness. Digital technology has enabled the possibility of supporting older adults in their wishes to age in place. However, current studies have mainly focused on the solitary evaluation of individual technologies or on evaluating technologies for specific illnesses.

Objective: This study aimed to synthesize research on the experiences of older people from the Western culture with chronic illnesses or care needs and their families with digital technology for aging in place. From the meta-synthesis, a model was derived that can be useful for the development of assistive devices in old age and that can support health care providers and professionals in their work with affected individuals.

Methods: A systematic review and qualitative meta-synthesis was performed using an inductive approach, as proposed by Sandelowski and Barroso. We performed a systematic literature search in 6 databases from 2000 to 2019, with an update in 2021 and, in addition, conducted a hand search in 2 databases, relevant journals, and reference lists. The results of each study were analyzed using initial and axial coding, followed by theoretical coding. A conceptual model was derived.

Results: A total of 7776 articles were identified. Articles were screened independently by 2 authors based on the eligibility criteria. Finally, of the 7776 studies, 18 (0.23%) were included in the meta-synthesis. The derived conceptual model describes older adults with chronic illnesses or dependency on care and their family members in an individual process of reflection and decision-making, starting with the use of a digital device. Older adults live in times of change. They experience stable and unstable times of illness as they are part of a changing digital world. Hence, older adults and their families consider digital technology a solution to their current situation. As they become familiar with a specific digital technology, they refine their needs and demands, gain confidence in its use, and note its advantages and disadvantages. They weigh hopes, needs, demands, and experiences in a process of reflection to decide on convenience and inconvenience. Independent of their decision, they achieve peace of mind either with or without digital technology. This process can restart repeatedly during the illness trajectory of older adults.

Conclusions: This study promotes a differentiated understanding of older adults' experiences with digital technology. The conceptual model can be useful for the development of assistive technology in old age. Moreover, it can guide health care professionals in their work with older adults and their families to provide individual counseling to find the appropriate digital technology for their respective situations.

(JMIR Aging 2022;5(3):e34872) doi:10.2196/34872

KEYWORDS

older adults; old age; assistive device; aging in place; home modification; independent living; telemedicine; assistive technology; ambient assisted living; assisted living; community living; chronic disease; chronic condition; chronic illness; elder; older adult; systematic review; meta-synthesis; digital technology; mobile phone

XSL•FO
RenderX

Introduction

Background

Aging is closely linked to the question of how and where one wants to live in later life. Aging in place is a central wish for many older adults. They associate it with autonomy, continuity in daily life, privacy, and memories that give meaning to life [1,2]. The concept of *aging-in-place* can be viewed through 5 key themes. In addition to place, these are social networks, support, technology, and personal characteristics [3]. On the basis of these key themes, we defined aging in place as living in one's own home for as long as possible while maintaining social networks and respecting older adults' autonomy in deciding what professional or technical assistance is needed. Support, as part of the aging-in-place concept, refers to the likelihood of developing (further) health or care needs in the aging process. The World Health Organization defined healthy aging as "the process of developing and maintaining the functional ability that enables well-being in older age" [4]. Healthy aging at home is increasingly being supported by digital-technical solutions. The diversity of digitally assisted living technologies has grown significantly. Nilsson et al [5] reviewed digital technologies for older adults and their informal caregivers as interventions for healthy aging, using the corresponding World Health Organization framework. Interventions focused "on physical capacity and function, on managing the symptoms of dementia and cognitive impairment, on supporting functioning in daily life and on self-caring with a chronic disease" [5]. Interestingly, the ability to participate in society was not addressed at all, although it is a central aspect of aging in place.

For existing qualitative reviews or meta-syntheses of the experiences of older adults living at home and using digital technologies, the study situation is as follows. In a meta-synthesis, Larsen et al [6] analyzed the process of becoming a user of assistive technologies and identified facilitating factors. They had narrow criteria for technology inclusion and did not focus on linking aging in place with assistive technology. In a qualitative systematic review and meta-synthesis, Moore et al [7] specified whether users integrate devices into everyday life with the degree to which motivation, ease of use, and device purpose match. In a meta-ethnography, Rosenwohl-Mack et al [8] specified the concept of aging in place as a balance of threats, agency in relation to identity, connectedness, and place. Owing to the holistic view of aging in place, the connection between aging in place and the use of technological-digital tools from the user perspective was only marginally addressed.

The systematic review by Pol et al [9] on the use of digital technologies by older adults summarizes studies that mainly examine healthy volunteers without care needs and focus on technical aspects of sensory monitoring rather than on the applicability in the daily life of a person aging in place. In a scoping review, Rodrigues et al [10] examined the extent to which web-based interventions could address loneliness in the context of the COVID-19 pandemic. Both studies show little visibility of the subjective perspective in the use of digital technologies in the aging in place of older adults. Other studies have concentrated on a specific group, such as the study by Chen and Yeh [11] on individuals with diabetes. In a meta-synthesis on self-monitoring of diabetes, they showed the experiences of patients on 5 different topics that can help health care professionals to better communicate with patients. Research on this topic has often focused on benefits, barriers, or ambivalences. The scoping review by Raja et al [12] shows such ambivalences in the context of telehealth, making life easier and the opposite. The scoping review also shows that the absence and presence of social support facilitated the use of digital technologies. A systematic review by Stargatt et al [13] stated the benefits of digital storytelling for older people with dementia living in the community. The improvements were concerned with mood, memory, social engagement, and quality of relationships in older people with dementia.

A research gap that emerges is that the experience of older people living at home when using digital technologies is still included less in the general evidence. The perspectives of both older adults and family members should be addressed when developing technical solutions [5,14]. The experiences of older adults aging in place and their family members are valuable for the development of a conceptual model overarching the varying illnesses, levels of care needs, or tested digital technologies.

Objective

Therefore, we intended to synthesize qualitative research to gain insights into the experiences of older adults and their family members. The following research question guided this study: what experiences do older adults with chronic illness or dependency on care and their family members have with digital technology referring to aging in place? The aim was to gather and synthesize qualitative research on the experiences of older adults and their family members in using digital technology to overcome challenges in aging, chronic illness, and the need for care in daily life. We intended to derive a conceptual model that illustrates the patterns identified within these experiences. As such, the model can be useful for the development of assistive digital technology in old age, as well as support health care providers and health care professionals in their work with specific populations. The term *healthcare professionals* refers to nurses, physicians, and therapists.

Methods

Overview

A qualitative systematic review and meta-synthesis was conducted. Conducting a meta-synthesis allows for the interpretive integration of several studies, resulting in findings based on a larger sample than would be possible in a single qualitative research study. It is a systematic and inductive approach in which the results of the included studies are interpreted as a whole. Consequently, coherent experiences of the research topic can be explained and described as they enable an understanding that goes beyond a mere summary of all findings. The approach by Sandelowski and Barroso [15] was used with the following recommended steps: (1) formulating a purpose, (2) searching and retrieving literature, (3) appraising findings, (4) classifying findings, (5) conducting a

XSL•FO
RenderX

meta-summary, and (6) developing a meta-synthesis. To ensure methodical rigor, we followed the ENTREQ (Enhancing Transparency in Reporting the Synthesis of Qualitative Research) checklist [16].

Searching for and Retrieving Relevant Studies

A sensitive search strategy was deployed for the qualitative systematic review to identify all the studies relevant to our research question. We oriented ourselves on the PICo (population, phenomenon of interest, and context) scheme supplemented by design. The search terms used in this study are listed in Textbox 1. These were adjusted slightly to fit the different search systems, such as using Medical Subject Heading terms in MEDLINE or subject headings in CINAHL. The search was conducted using the following databases: MEDLINE, CINAHL, PsycINFO, SocINDEX, GeroLit, and Bibnet. In addition, a hand search was performed in Google Scholar, Social Science Open Access Repository (SSOAR), and in specific journals: *Technology and Health Care*, *Neurorehabilitation*, *Technology and Disability*, *Journal of Ambient Intelligence and Smart Environments*, *Telemedicine and eHealth*, *Telemedicine and Telecare*, *Journal of Applied Gerontology*, *Archives of Geriatrics and Gerontology*, and the German journals—*Pflegezeitschrift*, *Pflege*, and *Pflege & Gesellschaft*. Subsequently, a literature search of the reference lists of relevant studies was conducted.

Textbox 1. Search strategy.

Category and search terms
Populations
(patients OR resident* OR “in need of care” OR “care needs” OR impaired OR disability OR disabled OR geriatric* OR elderly OR old OR “old age” OR “chronic disease” OR “chronic* ill” OR “chronic condition” OR “long-term condition” OR sick OR ill OR illness OR disease OR outpatients OR “next of kin” OR relative OR relatives OR family OR “family member*”)
Phenomenon of interest
AND (gaming OR exergaming OR app OR smartphone OR computer OR “mobile phone” OR tablet OR “health monitoring” OR telemedicine OR “tele-monitoring” OR telemonitoring OR “tele-homecare” OR telehomecare OR “tele-health” OR “telehealth” OR “tele-care” OR telecare OR “tele-nursing” OR telenursing OR telemetry OR “tele-communication” OR telecommunication OR tracking OR robotic OR “web based” OR “web-based” OR “health informatic” OR AAL OR “ambient assisted living” OR digital divide OR digital* OR virtual OR internet)
Context
AND (home OR house OR flat OR community OR “flat share” OR “flat-share” OR “assisted living” OR “assisted-living”)
Design
AND (qualitative OR “qualitative research” OR “qualitative studies” OR “mixed-methods” OR “mixed methods” OR ethno* OR hermeneutic* OR constructiv* OR constructionis* OR phenomeno* OR “focus group*” OR narration OR observation* OR interview* OR experience* OR “grounded theory” OR “image interpretation”)

The eligibility criteria for the literature search and the study selection process are presented in Table 1. Subsequently, the term *older adults* is used to refer to individuals of older age with chronic illness or dependency on care. We defined *old age* broadly to take into account the individuality of aging trajectories, its social construction, and other influences, such as biological, psychological, and social dimensions [17,18]. We focused on older adults from Western industrialized regions as they are considered similar in terms of cultural norms and values, as well as sociopsychological aging [18]. Studies that involved family members were also included. At the same time, the inclusion of family members in studies was not necessarily required. In this study, the term *family members* will be applied irrespective of eventual involvement in informal caregiving. The term family refers to close people whom the older person includes in this group, regardless of an existing family relationship [19]. We also included additional criteria for study quality. The development of digital technologies is advancing as fast as users’ attitudes and interactions with technologies are changing. Therefore, we limited the inclusion of studies from 2000 to 2021 to focus on current technologies and the experiences of the generation that will become older adults in the near future. We have considered studies in the languages that the authors speak.

XSL•FO
RenderX

Table 1. Inclusion and exclusion criteria of the literature research and the study selection process.

Selection criteria	Inclusion criteria	Exclusion criteria
Population	• Older adults or synonymous expressions with chronic illness and dependency on care aged ≥50 years • Family members • From Europe, North America, or Australia	• Children, adults aged <50 years • Health care professionals • Physicians • From South America or Asia
Setting	• Rented or purchased homes of older adults • Outpatient • Assisted living facilities	• Homelessness • Home of a family member • Inpatient • Nursing home
Digital technology	• Internet- or sensor-based digital technology • Use at home in daily life to overcome challenges in aging, illness, and care needs	• Technology under development • Technology for diagnosis purposes • Tests in laboratory situations
Experiences	• Older adults using digital technology related to aging in place • Family members who use digital technology with the goal of supporting older adults to age in place	• Experiences stated by health care professionals • Exclusive focus on technology without relation to user experience • Exclusive focus on usability and acceptance or nonacceptance
Study designs	• Qualitative design • Mixed methods design with a separate qualitative part	• Quantitative design • Discussion papers • Reviews • Monographs and book chapters • Study protocols
Study quality	• Clear separation of perspectives • Sufficient quality score (≥7) in the CASP[a] checklist	• No clear separation of perspectives • Insufficient quality score (≤7) in the CASP checklist
Year of publication	• Between 2000 and March 2021	• Before 2000 and after March 2021
Language	• English, German, or Spanish	• All other languages

[a]CASP: Critical Appraisal Skills Program.

Appraising and Classifying the Findings

A systematic literature search was performed in March 2019, with an update made in March 2021. All studies were imported into the reference management system. A total of 2 researchers independently screened titles and abstracts for eligibility and examined the full texts of the studies that appeared to meet the inclusion criteria, as far as this was evident from the title and abstract. Discrepancies were discussed with a third researcher. The remaining 26 studies were appraised critically by 2 researchers using checklists for qualitative research of the CASP (Critical Appraisal Skills Program) [20]. This CASP checklist enables the systematic appraisal of qualitative studies to identify strengths and weaknesses [20]. The tool appraises the quality of the studies but not the quality of the appraisal itself. Studies were excluded if they had <7 "yes" points in the 10-point questionnaire. Consequently, 8 studies were excluded because of insufficient quality. An overview of the critical appraisals of the included studies is provided in the Multimedia Appendix 1.

Sandelowski and Barroso [15] recommended classifying the findings with regard to the methods that were used and the manner in which the data were interpreted. Hence, the included studies were classified as one of the following: thematic surveys, conceptual or thematic descriptions, or interpretive explanations.

Conducting a Meta-summary

Two researchers independently extracted the following data from the included studies: authors, location, population, age of older adults, focus of interest, used methodology, and used technical devices. The studies were transferred to MAXQDA 2020 software (VERBI GmbH), which was used to support and manage the analysis process. We considered the results of the included studies to be interpretations of the collected data. Therefore, the results sections were treated as transcripts and used as data. The studies were read several times, and the results sections were analyzed inductively. Parts referring to the perspective of health care professionals were not taken into account. As a first cycle method, we used initial line-by-line coding. We posed the following question to the text: what are the positive and negative experiences of those affected and their family members with digital technology in terms of use, daily life, and their illness? This was followed by axial coding as a second cycle, which led to the first descriptive categories [21]. The meta-summary reflects the contents of the included studies.

Developing a Meta-synthesis

Axial coding was combined with constant comparison [21]. The codes were grouped by constantly comparing similarities and differences. Both helped discover patterns and reassemble the data in categories and subcategories, which facilitated the

XSL•FO
RenderX

formation of a hierarchical tree structure. The leading questions included the following: what added value and fears do affected individuals identify, which processes characterize use, and how did they make a decision? Theoretical coding was used as the third-cycle method to gain a deeper theoretical level of abstraction [21] and develop more generic categories. The main concepts emerged through this interpretive approach, and a conceptual model was derived.

Ethics Approval

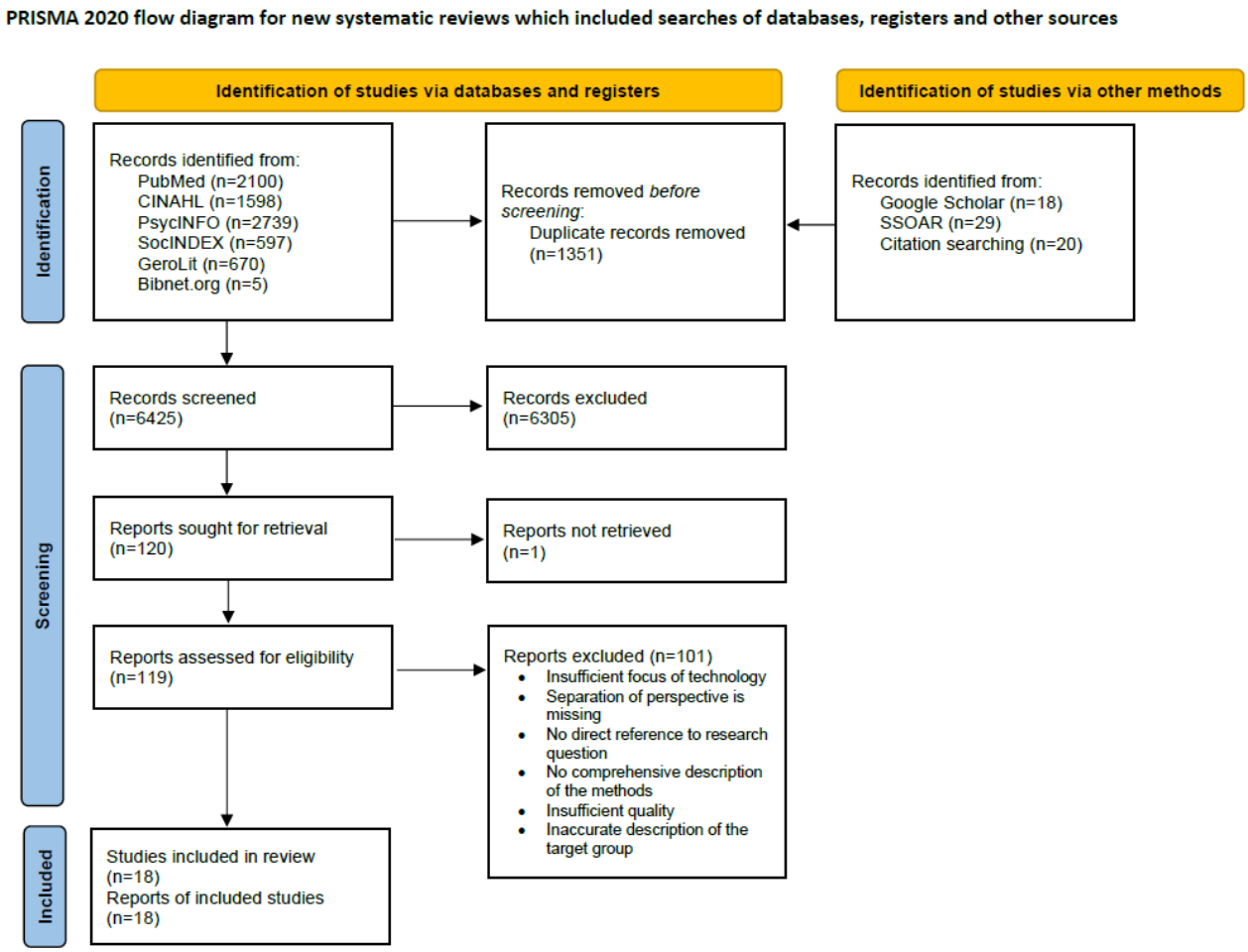
Ethics approval for this study was not necessary as it was a qualitative systematic review and meta-synthesis of published articles and did not involve data collection from participants.

Results

Overview

A total of 7776 studies were identified based on a systematic literature search of databases and a hand search. Furthermore, 119 studies were examined in full. A flow chart of the literature search is shown in Figure 1. Finally, 18 qualitative studies involving 220 older adults and 37 family members were included in the meta-synthesis. The study characteristics are summarized in Table 2. Older adults received care mostly in community-dwelling surroundings in 89% (16/18) and in assisted living residents in 11% (2/18) of studies. Data were collected through interviews in all the studies. Additional observations were part of 17% (3/18) of studies, and focus group discussions were part of 11% (2/18) of studies. The methodologies used were grounded theory, content analysis, thematic analysis, phenomenology, and systematic text condensation analysis. The ages of the older adults ranged from 52 to 101 years. The studies were conducted in seven different countries: Sweden, Denmark, the United Kingdom, the United States, the Netherlands, Scotland, and Australia. The technology included telehealth applications, telecommunication technology, ambient assisted living technology, exergame platforms, and wearables.

Figure 1. Flow diagram of the review process.



XSL•FO
RenderX

Table 2. Characteristics of included studies.

Authors	Location	Population	Age (years) of older adults	Focus of interest	Methodology	Technical device
Chao et al [22]	United States	15 assisted living residents with functional deterioration	78-92	Experiences of assisted living residents with facilitators and barriers with respect to exergames in relation to cognitive, physical, and psychological effects	Content analysis	The Wii-Exergaming intervention includes gaming activities for improving their functional impairment.
Chen et al [23]	United States	13 people with stroke	52-86	Experiences of benefits and barriers when using a telerehabilitation system	Thematic analysis	The telerehabilitation system comprises treatment sessions in the form of daily guided rehabilitation games, exercises, and stroke education.
Emme et al [24]	Denmark	9 people with COPD[a]	Mean 67.6	Coping with physical, emotional, and social problems of individuals with acute exacerbation of COPD before, during, and after virtual admission	Grounded theory	Virtual admission comprises virtual scheduled ward rounds via videoconferencing systems and medical equipment as monitoring devices.
Göransson et al [25]	Sweden	17 people with different comorbid conditions	70-101	Experiences with an app for supporting older people's health and self-care	Thematic analysis	The app "Interaktor" is used to report health problems and receive evidence-based self-care advice and links to relevant websites. Caregivers can access the generated information via a web interface.
Gorst et al [26]	England	8 people with COPD; 5 family members	58-84	Beliefs and perceptions of individuals with COPD in using home telehealth	Phenomenology	Peripheral telemedicine devices were used to monitor one's vital signs such as blood pressure, oxygen level, pulse, temperature, and weight.
Killin et al [27]	Scotland	10 individuals with Alzheimer disease or dementia; 10 family caregivers	66-81	Experiences of families with a diagnosis of dementia using a digital support platform	Thematic analysis	It involved an internet-based support platform that combines three different technologies: Living It Up, Jointly, and Click-Go. It enables families to receive comprehensive knowledge about dementia.
Klompstra et al [28]	Sweden	14 people with chronic cardiac diseases	56-81	Preferences, attitudes, use, and abilities of individuals with heart failure when using an exergame platform	Content analysis	The exergame platform (Nintendo Wii) enables to play games such as basketball, boxing, bowling, tennis, and golf. Using remote control, patients learn to play these sports in a way similar to real life.
LaFramboise et al [29]	United States	13 people with heart failure	Mean 68	Experiences about ease of use, efficacy, and difficulties of individuals with heart failure using a home communication device	Content analysis	"Health Buddy" is a telehealth device that allows people to record the status of their heart failure symptoms and receive information about their health state.
Lie et al [30]	England	21 people with chronic age-related health conditions; 11 family members	≥65	Experiences of older people with a home monitoring system focused on the acceptability, use, design, and trust of the system	Thematic analysis	The "SHel" home monitoring system comprises a home hub that communicates with wireless passive infrared sensors to monitor people's activities.

XSL•FO
RenderX

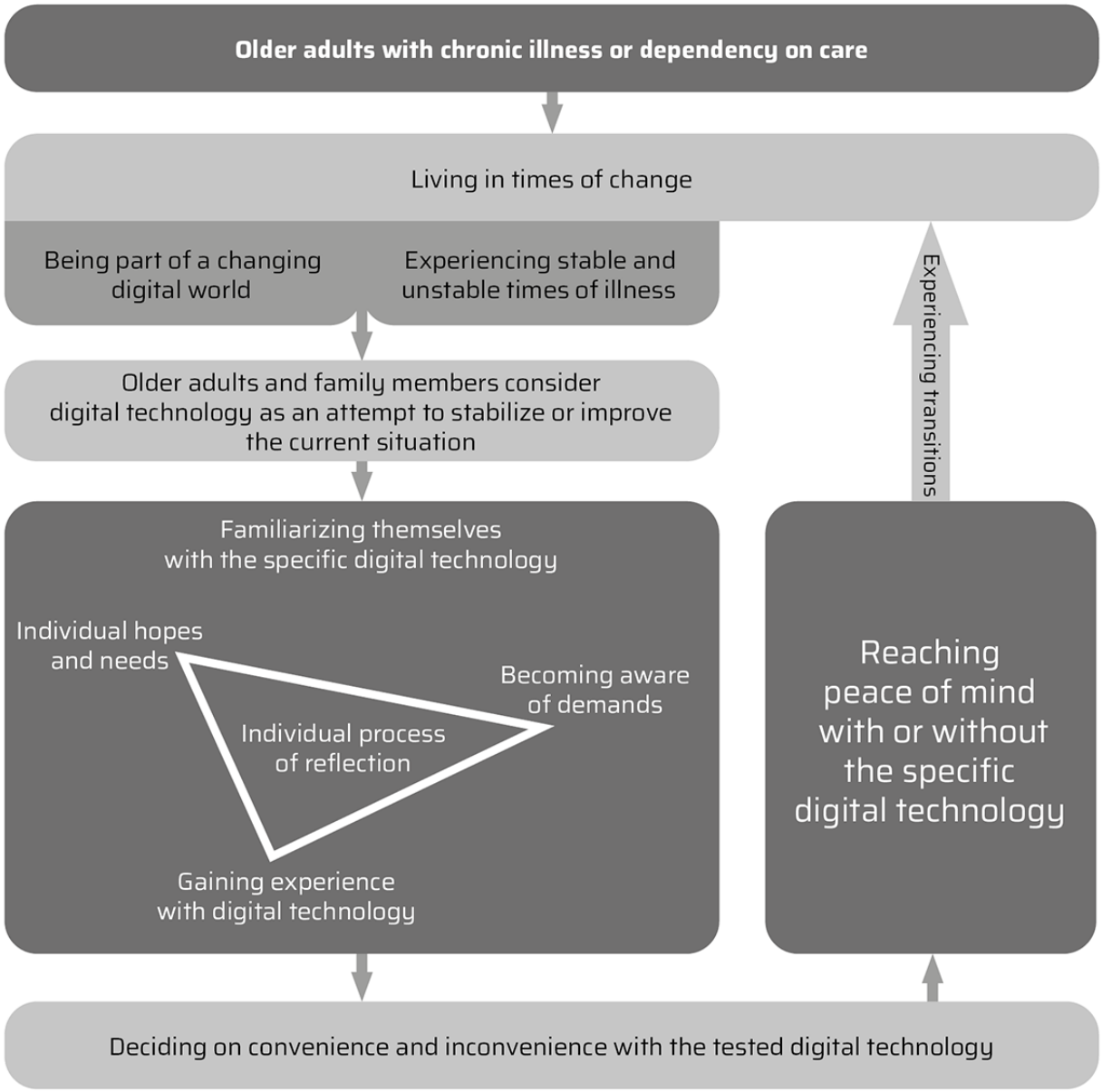
Authors	Location	Population	Age (years) of older adults	Focus of interest	Methodology	Technical device
Lind et al [31]	Sweden	12 individuals receiving palliative home care; 4 spouses	58-79	Experiences of individuals receiving palliative home care with a pain diary and digital pen and understanding their perception of pain control	Cross-case content analysis	The “Digital pen” technology comprises a pen and ordinary paper with a printed close-to-invisible pattern read by a camera inside the digital pen. This tool is intended for follow-up pain treatment, with assessment and documentation of their pain.
Lind and Karlsson [32]	Sweden	14 people with heart failure; 2 spouses	Mean 84	Experiences of older individuals with heart failure and spouses using a digital health diary and a digital pen technology	Content analysis	A telemedicine diary and digital pen are used for daily assessment. With this telehealth equipment, health care professionals can monitor patients’ daily reports via a mobile internet connection.
Mathar et al [33]	Denmark	6 people with COPD	67-83	Experiences and preferences of individuals with COPD using tele–video consultations after discharge from hospital	Systematic text condensation method	The tele–video consultations comprise eight 30-minute live investigations. During these sessions, medical professionals make observations to examine their general well-being and give advice on their medication.
Olsson et al [34]	Sweden	11 people with mild dementia	62-72	Perceptions of a passive positioning alarm for people with dementia	Content analysis	A passive positioning alarm is a GPS that includes a transmitter and a receiver. In addition, communication is possible through a loudspeaker function and getting help by pushing a button.
Selman et al [35]	England	12 people with heart failure and COPD	Mean 71.2	Experiences with a yoga intervention in relation to acceptability, appropriateness, and potential active ingredients for people with COPD and heart failure	Content analysis	The multipoint videoconferencing system enables 1-hour teleyoga live stream classes at home to receive personal instruction from the coach.
Shulver et al [36]	Australia	13 individuals undergoing rehabilitation; 3 spouses	60-92	Experiences of community-dwelling participants with a home-based telerehabilitation program and its acceptability	Thematic analysis	The home-based telerehabilitation program involves off-the-shelf technologies with tracking of activity data from the FitBit and having video calls via iPads with the therapists.
Smaerup et al [37]	Denmark	7 people with vestibular dysfunction	67-86	Experiences with computer-assisted home training for vestibular rehabilitation with a focus on self-efficacy, motivation, and acceptance	Meaning Interpretation Analyses (phenomenology)	The computer-assisted rehabilitation program “Move it to improve it (Mitii)” is personalized for patients exercising at home. Therapists can adapt the program to individual needs.
Starkhammar and Nygard [38]	Sweden	7 people with memory impairment; 7 relatives	66-87	Experiences from older adults and their families with a timer device for the stove	Grounded theory	The timer device was installed in existing electric stoves, which protects the user from fire hazards and allows the user to receive a warning signal if they forget to turn off the stove.
van Hoof et al1 [39]	The Netherlands	18 community-dwelling people with different chronic illnesses	63-87	Experiences, needs, and motives of individuals aging in place with new ambient intelligence technologies	Content analysis	The unattached autonomous surveillance system comprises the following various functions: mobility monitoring, voice output, fire detection, and wander detection and prevention.

XSL•FO
RenderX

[a]COPD: chronic obstructive pulmonary disease.

To answer the research question, a conceptual model was synthesized based on an inductive interpretive analysis process for the included studies. The conceptual model derived from this comprises 3 concepts in an ongoing process that reflect the experiences of older adults aging in place and their family members with digital technology (Figure 2). Table 3 provides an overview of the identified concepts, categories, and subcategories.

Figure 2. Conceptual model of older adults' and family members' experiences with digital technology.



XSL•FO
RenderX

Table 3. Concepts, categories and subcategories (excerpts from data analysis).

Concepts and categories	Subcategories
Living in times of change	
Experiencing stable and unstable times of illness	• Having substantial experience with the illness • Being confronted with worries, insecurities, and insufficient coping strategies
Being part of a changing digital world	• Seeking new ways to age in place • Considering digital technology to stabilize or improve the current situation
Familiarizing oneself with digital technology	
Having initial hopes and needs	• Wishing to age in place • Not wanting to rely on others • Wanting to improve health condition • Wanting to challenge themselves
Becoming aware of demands and further needs	• Noticing further needs • Having to adapt
Gaining experience with digital technology	• Overcoming initial difficulties • Gaining confidence in the use • Recognizing advantages and disadvantages
Individual process of reflection to decide on convenience and inconvenience	
Weighing hopes, needs, demands, and experiences	• Becoming aware of changes • Reflecting (un)consciously • Noticing what is important for oneself
Deciding on convenience and inconvenience	• Being (not) ready to make concessions • Considering digital technology for future use • Gaining peace of mind

First Concept: Living in Times of Change

Overview

The conceptual model shows that older adults live in times of change. As they experience stable and unstable times of illness, they have substantial experiences with the illness and are confronted with worries, insecurities, and insufficient coping strategies. At the same time, they are part of a changing digital world. They seek new ways of managing illness or care needs, which promote their wish for aging in place, and consider digital technology as a possible solution.

Experiencing Stable and Unstable Times of Illness

Chronic diseases often involve hospitalization or contact with health services and rehabilitation schemes. Individuals' expertise with their illness sometimes comprised years of experience, in which they gained knowledge, learned coping strategies, and tested them. When older adults were confronted with illness-related problems, they applied previous coping strategies that they were confident about and had positive experiences with [24,33].

However, individuals with chronic illnesses or dependence on care experience worries, insecurities, and insufficient coping strategies. They also experience loss of control when their coping strategies are insufficient. Thus, dependency on family members and health care professionals could arise [24,33]. Emme et al [24] called this a struggle to be in control of life. Affected individuals worried about whether they could age in place in the future. They anticipated a worsening of their physical condition and feared being lonely. When they noticed insufficient coping strategies, such as during the worsening of illnesses, they experienced feeling powerless and vulnerable [24,33].

Being Part of a Changing Digital World

Older adults and their family members have the fact that they are part of a changing digital world in common. Older adults had different extents of experience with, knowledge about, and attitudes toward digital technology. Their attitudes ranged from being skeptical to being indifferent to being motivated to learn something new. Although some were confident about their ability to learn, others doubted their ability to manage it [22,32,36,37,39]. Some had already used technology before joining the respective studies, mainly to promote feeling safe (eg, burglary alarms or emergency response systems) [30,39].

Nevertheless, older adults seek new ways to age in place because of their chronic illness or dependency on care. They considered digital technology as possibly helpful, some with the intent of just trying, and others with clear aims to improve their situations. Older adults hoped to age in place, whereas family members were glad to have digital technology as a fail-safe system and saw it as a prerequisite to enable aging in place. Family members felt a certain responsibility during the acquisition process [25,27,30,33,38,39]:

XSL•FO
RenderX

Technology was seen as a way to support the wish to age-in-place and, therefore, embraced, accepted or tolerated as a support tool [39].

Second Concept: Familiarizing Oneself With Digital Technology

Overview

Older adults already have individual hopes and needs when they learn digital technology and become aware of their specific demands while using it. During the process of familiarizing themselves with the respective digital technology, older adults and their family members continuously meet challenges and notice advantages and disadvantages.

Having Initial Hopes and Needs

Older adults' intrinsic motivation to test digital technology varied. Motivations differed from not wanting to rely on family members only or to be a burden, hoping for social contact, preventing harm, maintaining or regaining autonomy, and controlling or improving their health condition with the help of digital technology. Some even saw digital technology as a type of last resort [22,24,30,35-37,39]. Others wished that digital technology would take away the decision whether to alarm family members and would support being acknowledged in the illness situation by health care professionals [24,30]. Again, others wanted to challenge themselves with something new or decrease the generation gap between themselves and the younger generation by keeping up with the times, so that they could join conversations as they did not want to be old-fashioned [22,25,28,36].

Well, if the kids can do it, I can do it [36].

Older adults showed self-efficacy toward technology as they believed in their ability to manage it [32,37].

Becoming Aware of Demands and Further Needs

Older adults had individual hopes and needs regarding the use of digital technology. These were complemented by the awareness of further needs and the concretization of demands while gaining experience with their use. An example was the alterations that must be made in the usual surroundings to fit in digital technology. Although some older adults "see the technology with all its implications as a part of the home or as a part of the interior design" [39], others find the rearrangement of the furniture and the lack of flexibility in placing the device inconvenient [29,35]. Older adults "found it bothersome that placement of the Health Buddy had to be where there were both an available phone jack and electrical outlet" [29]. This resulted in older adults having limited opportunities to place the digital technology, which can lead to being annoyed by the location [23,29]. Hence, they became conscious of their demand to maintain their usual surroundings only when digital technology was built in.

Another example of becoming aware of new demands is an adaptation to daily routines. Although some older adults appreciated the flexibility in the use, for others, the development of new routines was helpful. The flexibility to use digital technology at any time during the day was appreciated if other appointments or tasks got in the way [23,28,29,32,33]. However, if the device was not intended to be used flexibly, older adults did not adhere to the recommended assessment times and did so with less frequency or at different times [31]. The establishment of new routines, when successful, supported older adults in their self-management [28,29]. Helpful routines could also be established through scheduled videoconferencing times [23,24,36]. Structured programs with planned and supervised exercises "requiring a commitment on their part motivated them to exercise regularly and adhere to an exercise program" [22]. However, committing oneself to an aim or being motivated did not seem sufficient for all older adults. Reminders from family members or the digital technology itself supported older adults in using it and adhering to their routines [23,28,29,31]. The use of a certain digital technology can be closely connected with the effort to adapt previous routines. Although some older adults appreciated the structured procedure, others required reminders to use it, whereas some appreciated flexibility.

Similarly, family members associated hopes, needs, and demands with the use of digital technology. They wanted to support older adults' wishes regarding aging in place. At the same time, the device satisfied their need for safety as they were less worried and mentally relieved [26,31,34,39]. Family members saw digital technologies as supportive while having the need to support themselves and appreciated the ability to approach the support of the platform when required [23,27,39].

Gaining Experiences With Digital Technology

For older adults and their family members, the use of digital technology was a learning process that required some time. They benefited from past experiences in building acceptance with the respective digital technology and facilitating its use [24,25,27,34,37,38]. Family members wanted to be prepared to support individuals with cognitive impairments for whom it was helpful to learn through several senses. Older adults with cognitive impairments familiarized themselves with the device, for instance, by being instructed verbally by technicians or family members, by reading instructions on little memos, watching other people use it, and experimenting with it [27,38,39]. It was helpful to have a program or digital technology tailored to older adults' needs to motivate and ensure a feeling of safety, such as modifying exercises or adapting digital technology to prevent false alarms caused by pets [22,30,35,39].

Older adults experienced difficulties in handling software and hardware [27,29,31,37]. This resulted in older adults not using parts of the technology or having difficulties in their use, for instance, because of impaired sight [31,32,39]. People with memory impairment felt challenged not to forget charging or where they placed the device [34]. For them, it was demanding to remember the handling, such as necessary actions to reset the stove timer or perform actions on a written memo. Individuals with cognitive impairments continuously required reminders [38]. Older adults tried to solve problems that arose on their own through experimentation or relied primarily on help from family members or friends. Most hesitated to contact IT support in case of hardware or software problems [27,31,32,37,38], whereas others felt reassured over the

XSL•FO
RenderX

availability of technical support and expected rapid responses [23,36,37].

After having managed initial insecurities, older adults gained confidence with respect to use, irrespective of cognitive impairments or terminal illnesses. They could apply the different functions and could even become enthusiastic about being able to use them, increasing skills or the benefits that they noticed [22,23,25,27,28,31,33,36-38]. They enjoyed the digital technology, even looked forward to using it, and noticed an increasing sense of well-being [22,23,25,27-29,34-36,39]. With increasing use, older adults reported improvements in their mental and physical conditions. They noticed that the potentially negative outcomes were diminished. Depending on the tested digital technology, they were more attentive and more relaxed, whereas others reported an increase in physical activity and fitness through physical exercise, for instance, by exergaming or telerehabilitation [22-24,28,29,35,36,38]. Older adults were more easily aware of symptoms and deteriorations with the help of digital technology. They were able to cope with different situations on their own, with medication at hand, or with health care professionals via telehealth. Depending on the tested technology, they were virtually admitted if needed or contacted by health care professionals about relevant issues; thereby, older adults could be cared for at home. Not needing to go somewhere for an appointment and being able to return to everyday life as soon as possible was appreciated and time saving [23,24,26,32,36,37].

Further positive experiences referred to not having to be in control. For example, during virtual admission, older adults were able to hand over responsibility and control to health care professionals who took over disease management [24]. People with cognitive decline were less anxious about the stove timer as they did not need to be in control as before [38].

Older adults also experienced increased, continuous, and improved contact with health care professionals [23,25,26,31,36]. They experienced social support through digital technology, although they did not have more frequent human interactions. For most, contact via devices was equal to face-to-face visits [26,29,36]. Communicating health concerns via digital technology, such as video consultation or a digital pen, gave older adults a sense of security as they felt having a direct connection to and "closer contact" with health care professionals. They felt the need to have positive relationships with health care professionals and did not perceive telehealth as a barrier [23-25,31-33,36]. Older adults wished to receive appreciation and feedback. They perceived the reporting of health concerns via digital technology as appreciation and support [25,31].

However, there were also older adults who wished for more social interaction with other affected persons or health care professionals. Some missed the contact with health care professionals to receive specific feedback or have the opportunity to ask questions. Although for some, it was only important to receive individual feedback, others valued face-to-face contact [26,33,35-37].

For some older adults, the use of the tested digital technology was repetitive or boring, concerning content and exercises that had limited options and were seen as monotonous [28,29,37]. Older adults were annoyed by problems or characteristics of the technical infrastructure or device, such as unreliable internet connections, false alarms, humming, and lighting, as well as negative auditory and visual feedback in exergaming [22,29,35,38,39]. They reported not having time to use the tested device; seeing it as an additional obligation; and prioritizing other things instead, such as holidays. Hence, over the period of use, some older adults became less adherent [27-29,33].

A central wish of older adults was to be self-determined in their decision to use digital technology. At the same time, they and their family members experienced uncertainty about the "right" timing. The discussion around timing is especially evident in individuals with cognitive decline. That is why individuals with cognitive impairments, having tried the device, suggested introducing the device early to promote self-determination of individual persons themselves, so that they can test it, are familiar with the device, and regard it as an aid [34,38].

Third Concept: Individual Process of Reflection

Overview

This concept describes the process of reflection in which older adults and their family members were engaged in deciding on the convenience and inconvenience of the digital technology tested. They consciously or unconsciously reflected on their individual hopes, needs, demands, and experiences. Hence, they decided on the convenience or inconvenience of the tested digital technology. Irrespective of their decision, participants achieved peace of mind, some with digital technology and others without.

Weighing Hopes, Needs, Demands, and Experiences

Overview

Reflecting, in this context, is not a set process of concretely identifying hopes, needs, and demands and weighing pros and cons but a noting of what is important to oneself, incorporating hopes, needs, and demands. A participant highlighted feeling safe because of telehealth:

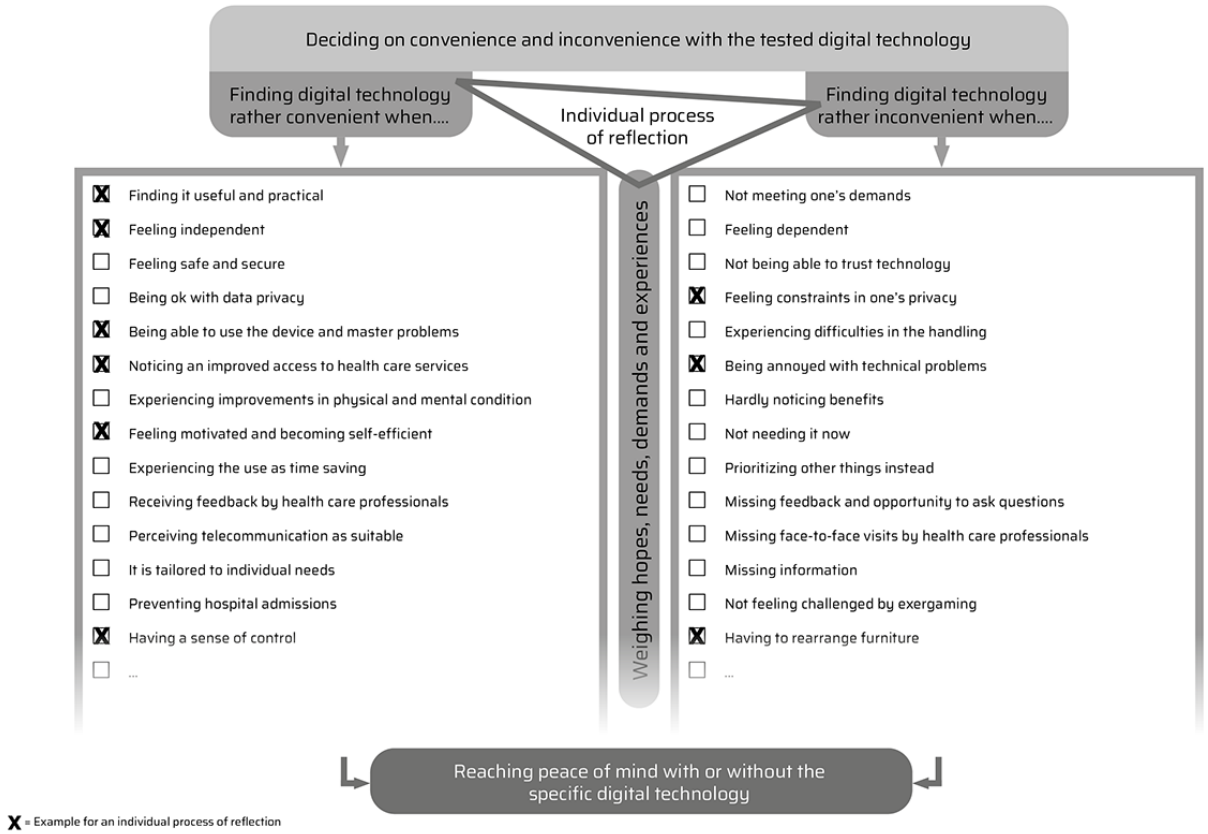
> *Like I say it's reassuring, it's like having another person with you even though it's a machine. I think that's the thing about it. It's because I live on my own, isn't it? I know my son is only a phone call away but I feel more reassured now that's in [26].*

Therefore, reflection in decision-making is more deliberate for one person than for another. The analysis revealed aspects based on the hopes, needs, demands, and experiences of older adults and their family members, which they considered in their individual reflection process. The results are shown in Figure 3. For example, older adults found digital technology convenient when they noticed its usefulness and practicability. They then appreciated it as an alternative to conventional offers such as doing teleyoga alone instead of exposing themselves to a group or doing exergaming at home when there was bad weather [28-30,32,35,36]. By contrast, when older adults and their family members found digital technology inconvenient, they usually noticed that it did not meet their demands. They then did not see the need to use it any further as they hardly noticed benefits in coping, fitness, self-managing behavior, or care participation

XSL•FO
RenderX

[22,24,26,28,30,32,33]. For example, individuals with cognitive impairment reacted with resistance as the digital support platform confronted them with their diagnosis or one felt it was more relevant for family members [27]. A family member explained that he had a close relationship with his father and frequent face-to-face contact where digital technology was not needed [30]. Therefore, every person has different aspects that are important to them personally and that are considered in the reflection process. Hence, the aspects listed in Figure 3 need not apply to every person. Finally, reflection led older adults to decide whether they found the digital technology more convenient or inconvenient.

Figure 3. Aspects considered in the process of reflection.



The reflection process is explained below in terms of 3 key themes within the hopes, needs, and demands identified: wishing to be safe, thriving for independence, and wanting to be in control. In doing so, it becomes clear what this process of reflection by older adults and their family members can look like. In the example, they weighed whether the safety, independence, or control they gained from the digital technology they tested were in relation to the individual concessions they would have to make, such as arising dependency and restricted privacy.

Wishing to Be Safe

Feeling safe at home is a prerequisite for aging in place. Being able to stay at home supported by digital technology provides a feeling of safety, confidence, and comfort [25,31,37,39]. The feeling of safety developed through the 24/7 availability and continuity of digital technology—knowing that health care professionals would be alerted by the system or the older adult, if required [24-26,31,33]. Even family members could increase their feelings of security through tested digital technology [26,34,38,39].

Thriving for Independence

Digital technology also facilitates self-determination and promotes independence, such as being virtually admitted to the hospital while being able to stay in one's usual surroundings [24,39]. Individuals with cognitive impairments noticed increasing freedom and independence through the use of a passive positioning system [34]. Being able to manage digital technology and becoming active in improving one's health condition promoted this feeling of independence. They noticed feeling empowered to care for their own health, being self-sufficient, and taking over responsibility [22,25,26,28-31,35]. An interviewee stated about the tested health buddy device:

> *It taught you how to take care of yourself and do it on your own, because nobody else is gonna do it for you [29].*

Interestingly, taking over a greater part of their own care also mediated a sense of increased security [31,32].

Wanting to Be in Control

Older adults gained greater knowledge about their illness and symptoms through the information provided via their digital

XSL•FO
RenderX

devices. Being able to use digital devices, interpret monitoring data, and contact health care professionals whenever necessary provided older adults with a sense of control over their illness management. Some older adults experience insecurities over when to approach health care professionals when experiencing an exacerbation of chronic obstructive pulmonary disease. When they are unsure of their health status, they can refer to actual and past telehealth data to confirm their symptoms [23-26,29,31-33,35,37]. An interviewee who did tele yoga said the following:

> *When I started to get really short of breath and I saw those numbers bang up, and I was thinking okay, I've got to slow down. So I went into the yoga breathing [...] and just slowed it down. And the nurses were amazed that I could do it, yeah...Because most adults hyperventilate, because they get anxious [35].*

Example of the Process of Weighing

The positive experiences made in relation to the 3 explicated themes must be weighed against further hopes, needs, and demands that older adults and family members have and the experiences they face. They must reflect on whether they want to make concessions or not.

Lie et al [30] "found inherent tensions in maintaining both safety and privacy at the same time, as in some cases, a certain amount of privacy had to be given up for the sake of safety." Privacy is seen as important and closely connected to being in control, as well as independent. Although some valued control over their privacy by exerting data sovereignty, others had no concerns regarding privacy and confidentiality. They did not feel invaded in their privacy or did not see telehealth as intrusive surveillance [26,30,34-36,39]. Older adults have a sense of losing independence by being watched [30,39]. They describe monitoring as "invasion of privacy" [30], or "big brother is watching you" [34]. Some participants felt restricted in their autonomy as they did not want to be watched, either via digital technology or by their families [30]. When using digital technology, older adults must trust the confidentiality of their own data. Being able to trust promotes feeling safe and, thus, acceptance of the system [30,36,38]. An interviewee (daughter of an older adult) explained it as follows:

> *You don't want to go to a care or nursing home, and then you have to make some concessions of course. [...] it is not like Big Brother, it is just a sort of assistive device to stay here for longer [39].*

Modern technologies such as medical equipment can also provoke a certain dependence while simultaneously providing independence [24,39]. An interviewee reflected on seeing the health care professional as responsible for monitoring his health before he used digital technology. This placed him in a dependent position. With the use of digital technology, he was able to self-manage his health [26]. Dependency also became clear through the reluctance of participants to return the device after the study [24,29,39]. For others, dependency became evident through the number of digital technologies in their households [39]. This dependency on assistive devices, in return, creates a feeling of security. Telehealth can mediate "a positive experience of surveillance" [26]. Older adults felt well cared for and watched over because they were connected to health care professionals via digital technology and interacted with the digital technology on a daily basis [25,26,29,31-33,39]. An interviewee explained it as follows:

> *sort of a lifeline; you know that it's going somewhere else [...] Knowing somebody is at the end of the line, that's important [26].*

Older adults noted that using digital technology created additional obligations that were unrelated to the independence they gained,

> *The security that the equipment can give the patients must thus be weighed against the obligations it creates [33].*

Obligations refer to solely using it or being at home for scheduled remote sessions at the agreed time [29,33]. Older adults and family members weighed their experiences with their inherent hopes, needs, and demands to determine whether they wanted to make these concessions.

Deciding on Convenience and Inconvenience

For older adults, achieving peace of mind means that in the process of weighing needs, demands, and experiences, some aspects receive higher priority from the individual and lead to decisions on convenience and inconvenience. Consequently, the necessary individual concessions are accepted as they do not play a leading role. The individually higher-ranked aspects lead to peace of mind. Decisions on convenience and inconvenience and gaining peace of mind also result in further actions, such as wanting to maintain digital technology, buying it, organizing one's situation without digital technology, and maintaining health-promoting measures.

The larger group of participants gained peace of mind *with* the use of their tested digital technology. They found it convenient and were ready to make certain concessions, such as being dependent on the technology and having to give up a certain amount of privacy. Two participants were not keen on the system's false alarms. They arranged with them and kept them out of health concerns [39]. Feeling safe was identified as the aspect receiving individually higher priority, mainly contributing to peace of mind for older adults and family members. Feeling safe not only referred to accessibility to health care professionals but also resulted from the feeling of being watched over by health care professionals. Being able to check the measurements of telehealth devices themselves, and interpret them, also contributed to feeling reassured [25,26,32,33,39]. Older adults "believe that telehealth had given them peace of mind regarding their health" [26]. An interviewee said, "I can reach them [the caregivers] easier and that means a feeling of greater peace for me, which is the main thing" [31]. Reaching peace of mind also referred to family members who felt relieved:

> *it's given him peace of mind completely too. [...] It's made us both have a life really without worrying [26].*

Older adults promoted different actions as a consequence of peace of mind. Some older adults were so convinced about the device that they considered buying it themselves [24,26,30,36]. For some, maintaining trust in their home "could sometimes

XSL•FO
RenderX

result in more devices being installed for peace of mind" [30]. For them, study participation resulted in wanting to continue monitoring one's own health [24,26]. They also started recommending digital technology to others when they were convinced of its functionality [22,34]. For others, testing digital technology during the respective study was a starting point in the change in behavior and coping strategies. Hence, they were motivated by the use of devices to continue with health-promoting measures, even without the device. Older adults reported being able to maintain established helpful routines, even without digital technology, such as continuing pain assessment with paper and pencil and doing exercises they learned [23,29,31,35].

Older adults who gained peace of mind *without* digital technology were not convinced of the convenience the specific digital technology gave them. They had to compromise in the form of having no digital technology. For them, having weighed hopes, needs, demands, and experiences resulted in the decision that for the moment, they were better off without digital technology. For some, it was even a relief to return to digital technology [29,30,39]. A son explained the following:

> *I don't think it will necessarily erm help us [...] we are seeing each other frequently face to face and we still live quite close to each other, so it would tend to be if we thought there was a problem we would call round [30].*

For other users and family members, the experience led to the insight of not needing it as still being autonomous or not having these special needs at the moment but considering it in the future, whereas others suggested that the tested devices were more convenient for users that were far more ill [27,30,33,34].

Hence, peace of mind is not a continuous state; however, with experiencing transitions in the illness trajectories, the process of considering a new digital technology or the same device at a later time can restart.

Discussion

Principal Findings

The conceptual model shows that older adults with chronic illness or dependency on care live in times of change. They experience stable and unstable times of illness and are part of a changing digital world. Hence, older adults and family members consider digital technology as an attempt to stabilize or improve their current situation. While familiarizing themselves with specific digital technologies, they are in an individual process of reflection. They weigh their hopes, needs, demands, and experiences to come to a decision about whether they find digital technology convenient or inconvenient. Independent of their decision, they achieve peace of mind, either with or without digital technology. The whole process can restart as older adults experience transitions in their illness trajectories and may consider a different technology or the same one later. In this section, we focus on three main aspects: first, thriving for independence while having to arrange with dependency; second, family members' process of adaptation; and third, health care professionals' relationships with older adults via telehealth.

Need for a Technical Device Tailored to the Individual Situation

The qualitative systematic review and meta-synthesis revealed that older adults thrive for independence. This was 1 of the 3 main themes that hope, need, and demand were ranked around. Digital technology can necessitate concessions as one experiences dependence on digital technology or being watched. Other studies have similarly stated this dualism. Holmberg et al [40] and Barken [41] highlighted how older adults strive for autonomy and independence while having to accept home care. Barken [41] explained how older adults maintain an independent identity by taking part in their own care. However, the independent self can be limited when older adults experience insufficient support and are unable to care for themselves as they wish. Moreover, Holmberg et al [40] identified that obtaining care implied accepting certain inconveniences to be treated in their ordinary surroundings. Hence, older adults live in a continuous process of adaptation because of experiencing illness, dependency on care, and a changing social environment while balancing their wishes with their current life circumstances [40-42]. Although these studies referred to "classical" caring situations *without* digital technology, we found similar results *with* digital technology. Older adults can make decisions about digital devices and the concessions to be made.

Our results show that family members associate digital technology with the option to enable aging in place. When successfully implemented, they experienced relief and a satisfied need for security. However, depending on the illness or care needs of the older adults, family members face various challenges. They see themselves confronted with competing demands while trying to develop a fitting arrangement [43]. Family members experience transitions with their ill relatives by striving for normalcy, as indicated by studies in palliative care [44,45]. While trying to maintain normalcy, they have to adapt to new life circumstances by balancing their old and new life situations, taking into account their own needs and demands [45,46]. From our results, family members are also confronted with a changing process in arranging with the tested technology. It remains unclear how the decision on convenience and inconvenience between family members and older adults is made. This could be further researched to optimally support the adaptation process of family members and older adults.

The results revealed mainly positive experiences of contact with health care professionals. The results obtained from appreciating the continuous availability and direct interaction with health care professionals point toward improved communication via digital technology. This leads to a closer connection with health care professionals. These positive reports are probably attributable to being part of a study, as health care professionals take more time for older adults than in normal circumstances without being part of a study. By considering the literature on video-based technology, the experiences of telehealth were equated with digital connectedness between older adults and health care professionals and were perceived as strengthening their relationship through better communication [47,48]. However, there are also contrary results in the experiences of the health care professional–patient relationship related to telehealth. Rykkje and Hjorth [49], as well as Steindal et al [47],

XSL•FO
RenderX

illustrated a loss of interpersonal dynamics and replacement of human contact by using chatbots, telephone, and video calls. This could result in a lack of trust, although trust is essential for a nurse-patient relationship [50]. Therefore, our results have to be considered in terms of the positive relationship between health care professionals and older adults. From these results, we cannot make a final statement about the relationship between health care professionals and older adults. Further research on realistic scenarios for the everyday use of digital technology is required.

Strengths and Limitations

The strength of this qualitative systematic review and meta-synthesis is the conceptual model that was derived, which provides a thorough understanding of the experiences of older adults and their family members. A sensitive search strategy was used to identify all the relevant studies. However, a limitation is that the studies focused on the experiences gained in the study context. Another limitation related to the experiences that were analyzed using different methodological approaches in the individual studies. Most studies used content or thematic analysis; only 22% (4/18) of studies used phenomenological or grounded theory methodologies. It should also be taken into consideration that only studies from Western cultures were included. Hence, transferring the results to a different cultural context must be verified beforehand. In addition, only articles published in English, German, or Spanish were considered. Another limitation is that the review was not registered in the international PROSPERO (International Prospective Register of Systematic Reviews) database.

Conclusions

We derived a conceptual model of experiences of older adults with chronic illnesses or dependency on care and their family members using digital technologies that they tested in their homes. The model showed older adults and family members in a reflection process of weighing hopes, needs, demands, and experiences to decide on the convenience and inconvenience of the specific digital technology. Irrespective of their decisions, they attained peace of mind. This is a continuous process; it can restart during individual illness trajectories with different digital technologies or the same ones later.

This meta-synthesis had several implications. In terms of practical implications, the conceptual model reveals the need for individual counseling of older adults with chronic illnesses or care needs and their family members. Their living conditions and illness situations must be taken into account when deciding on a digital technology. Above all, the variety of available technologies must be taken into account to select the right device for the individual. Moreover, the results indicate that older individuals and their family members need to be introduced to or even trained with specific digital technology. These devices should be provided or paid for by the respective public health insurance to prevent social inequality. Thus, the model can be useful for health care providers and health care professionals. The derived conceptual model can also be used to develop digital technology. This can be the basis for communication among different disciplines. When the disciplines of technology, usability, and health sciences collaborate, a common basis can be helpful for developing digital solutions with added value for many potential users. Finally, the conceptual model can be used for the training and education of health care professionals. They could be sensitized to transitions in the illness trajectory to realize how a specific digital technology can support the current situation or need to be changed, respectively, considering older adults' needs and demands. In addition, the conceptual model supports the understanding of health professionals and users. When users are aware of the process, they can participate more consciously to stabilize or improve their current situation with the help of digital solutions.

As an implication for future research, digital technology should be explored in the context of everyday life. This approach can enable an analysis of older adults' experiences in their relationship with health care professionals without being subjected to special conditions in the context of an evaluation study. Moreover, the future development of technological devices should integrate older adults into participative study designs. Further research is required on the decision-making processes of older adults and their family members.

Acknowledgments

This study was funded by the Bavarian State Ministry of Health and Care.

Authors' Contributions

MH and CA drafted the systematic literature search and meta-synthesis. MH performed the systematic search. MH, CA, and DH performed the screening process, critical appraisal, and analysis. MH drafted the manuscript. DH, CA, and CR commented critically on the manuscript. All authors consented to the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Quality assessment using the Critical Appraisal Skills Program (CASP).
[DOCX File , 19 KB - aging_v5i3e34872_app1.docx]

XSL•FO
RenderX

References

1. Löfqvist C, Granbom M, Himmelsbach I, Iwarsson S, Oswald F, Haak M. Voices on relocation and aging in place in very old age--a complex and ambivalent matter. Gerontologist 2013 Dec;53(6):919-927 [FREE Full text] [doi: 10.1093/geront/gnt034] [Medline: 23626372]
2. Stones D, Gullifer J. 'At home it's just so much easier to be yourself': older adults' perceptions of ageing in place. Ageing Soc 2014 Dec 01;36(3):449-481 [FREE Full text] [doi: 10.1017/s0144686x14001214]
3. Pani-Harreman KE, Bours GJ, Zander I, Kempen GI, van Duren JM. Definitions, key themes and aspects of 'ageing in place': a scoping review. Ageing Soc 2020 Mar 18;41(9):2026-2059 [FREE Full text] [doi: 10.1017/s0144686x20000094]
4. World Report on Ageing and Health. Geneva: World Health Organization; Sep 29, 2015.
5. Nilsson MY, Andersson S, Magnusson L, Hanson E. Ambient assisted living technology-mediated interventions for older people and their informal carers in the context of healthy ageing: a scoping review. Health Sci Rep 2021 Mar;4(1):e225 [FREE Full text] [doi: 10.1002/hsr2.225] [Medline: 33392394]
6. Larsen SM, Mortensen RF, Kristensen HK, Hounsgaard L. Older adults' perspectives on the process of becoming users of assistive technology: a qualitative systematic review and meta-synthesis. Disabil Rehabil Assist Technol 2019 Feb;14(2):182-193. [doi: 10.1080/17483107.2018.1463403] [Medline: 29683014]
7. Moore K, O'Shea E, Kenny L, Barton J, Tedesco S, Sica M, et al. Older adults' experiences with using wearable devices: qualitative systematic review and meta-synthesis. JMIR Mhealth Uhealth 2021 Jun 03;9(6):e23832 [FREE Full text] [doi: 10.2196/23832] [Medline: 34081020]
8. Rosenwohl-Mack A, Schumacher K, Fang M, Fukuoka Y. A new conceptual model of experiences of aging in place in the United States: results of a systematic review and meta-ethnography of qualitative studies. Int J Nurs Stud 2020 Mar;103:103496 [FREE Full text] [doi: 10.1016/j.ijnurstu.2019.103496] [Medline: 31884333]
9. Pol MC, Poerbodipoero S, Robben S, Daams J, van Hartingsveldt M, de Vos R, et al. Sensor monitoring to measure and support daily functioning for independently living older people: a systematic review and road map for further development. J Am Geriatr Soc 2013 Dec 25;61(12):2219-2227. [doi: 10.1111/jgs.12563] [Medline: 24479150]
10. Rodrigues NG, Han CQ, Su Y, Klainin-Yobas P, Wu XV. Psychological impacts and online interventions of social isolation amongst older adults during COVID-19 pandemic: a scoping review. J Adv Nurs 2022 Mar 09;78(3):609-644 [FREE Full text] [doi: 10.1111/jan.15063] [Medline: 34625997]
11. Chen C, Chang Yeh M. The experiences of diabetics on self-monitoring of blood glucose: a qualitative metasynthesis. J Clin Nurs 2015 Mar 25;24(5-6):614-626. [doi: 10.1111/jocn.12691] [Medline: 25258005]
12. Raja M, Bjerkan J, Kymre IG, Galvin KT, Uhrenfeldt L. Telehealth and digital developments in society that persons 75 years and older in European countries have been part of: a scoping review. BMC Health Serv Res 2021 Oct 26;21(1):1157 [FREE Full text] [doi: 10.1186/s12913-021-07154-0] [Medline: 34696789]
13. Stargatt J, Bhar S, Bhowmik J, Al Mahmud A. Digital storytelling for health-related outcomes in older adults: systematic review. J Med Internet Res 2022 Jan 12;24(1):e28113 [FREE Full text] [doi: 10.2196/28113] [Medline: 35019845]
14. Abdi S, de Witte L, Hawley M. Emerging technologies with potential care and support applications for older people: review of gray literature. JMIR Aging 2020 Aug 11;3(2):e17286 [FREE Full text] [doi: 10.2196/17286] [Medline: 32780020]
15. Sandelowski M, Barroso J. Handbook for Synthesizing Qualitative Research. New York: Springer Publishing Company; 2007.
16. Tong A, Flemming K, McInnes E, Oliver S, Craig J. Enhancing transparency in reporting the synthesis of qualitative research: ENTREQ. BMC Med Res Methodol 2012 Nov 27;12(1):181 [FREE Full text] [doi: 10.1186/1471-2288-12-181] [Medline: 23185978]
17. Ferraro KF, Shippee TP. Aging and cumulative inequality: how does inequality get under the skin? Gerontologist 2009 Jun 17;49(3):333-343 [FREE Full text] [doi: 10.1093/geront/gnp034] [Medline: 19377044]
18. Powell JL. Social Theory and Aging. Lanham, Maryland: Rowman and Littlefield; 2005.
19. Bozett FW. Family nursing and life-threatening illness. In: Leayhey M, Wright LM, editors. Families and Life-Threatening Illness. Springhouse, PA: Springhouse; 1988:2-25.
20. Singh J. Critical appraisal skills programme. J Pharmacol Pharmacotherapeutics 2022 May 23;4(1):76-77. [doi: 10.4103/0976-500x.107697]
21. Saldaña J. The Coding Manual for Qualitative Researchers (4th edition). London: SAGE Publications; 2021.
22. Chao Y, Lucke KT, Scherer YK, Montgomery CA. Understanding the Wii Exergames use: voices from assisted living residents. Rehabil Nurs 2016 Sep 14;41(5):279-288. [doi: 10.1002/rnj.216] [Medline: 25974657]
23. Chen Y, Chen Y, Zheng K, Dodakian L, See J, Zhou R, et al. A qualitative study on user acceptance of a home-based stroke telerehabilitation system. Top Stroke Rehabil 2020 Mar 04;27(2):81-92 [FREE Full text] [doi: 10.1080/10749357.2019.1683792] [Medline: 31682789]
24. Emme C, Rydahl-Hansen S, Ostergaard B, Schou L, Svarre Jakobsen A, Phanareth K. How virtual admission affects coping - telemedicine for patients with chronic obstructive pulmonary disease. J Clin Nurs 2014 May 26;23(9-10):1445-1458. [doi: 10.1111/jocn.12444] [Medline: 24372676]

XSL•FO
RenderX

25. Göransson C, Eriksson I, Ziegert K, Wengström Y, Langius-Eklöf A, Brovall M, et al. Testing an app for reporting health concerns-Experiences from older people and home care nurses. Int J Older People Nurs 2018 Jun 05;13(2):e12181. [doi: 10.1111/opn.12181] [Medline: 29210218]
26. Gorst SL, Coates E, Armitage CJ. "It's sort of a lifeline": chronic obstructive pulmonary disease patients' experiences of home telehealth. Health Psychol 2016 Jan;35(1):60-68. [doi: 10.1037/hea0000246] [Medline: 26148189]
27. Killin LO, Russ TC, Surdhar SK, Yoon Y, McKinstry B, Gibson G, et al. Digital Support Platform: a qualitative research study investigating the feasibility of an internet-based, postdiagnostic support platform for families living with dementia. BMJ Open 2018 Apr 12;8(4):e020281 [FREE Full text] [doi: 10.1136/bmjopen-2017-020281] [Medline: 29654028]
28. Klompstra L, Jaarsma T, Mårtensson J, Strömberg A. Exergaming through the eyes of patients with heart failure: a qualitative content analysis study. Games Health J 2017 Jun;6(3):152-158 [FREE Full text] [doi: 10.1089/g4h.2016.0087] [Medline: 28422516]
29. LaFramboise L, Woster J, Yager A, Yates BC. A technological life buoy: patient perceptions of the Health Buddy. J Cardiovasc Nurs 2009;24(3):216-224. [doi: 10.1097/JCN.0b013e318199a60f] [Medline: 19390339]
30. Lie ML, Lindsay S, Brittain K. Technology and trust: older people's perspectives of a home monitoring system. Ageing Soc 2015 Jun 09;36(7):1501-1525. [doi: 10.1017/s0144686x15000501]
31. Lind L, Karlsson D, Fridlund B. Patients' use of digital pens for pain assessment in advanced palliative home healthcare. Int J Med Inform 2008 Feb;77(2):129-136. [doi: 10.1016/j.ijmedinf.2007.01.013] [Medline: 17363323]
32. Lind L, Karlsson D. Telehealth for "the digital illiterate"--elderly heart failure patients experiences. Stud Health Technol Inform 2014;205:353-357. [Medline: 25160205]
33. Mathar H, Fastholm P, Sandholm N. A qualitative study of televideo consultations for COPD patients. Br J Nurs 2015 Feb 26;24(4):205-209. [doi: 10.12968/bjon.2015.24.4.205] [Medline: 25723262]
34. Olsson A, Skovdahl K, Engström M. Using diffusion of innovation theory to describe perceptions of a passive positioning alarm among persons with mild dementia: a repeated interview study. BMC Geriatr 2016 Jan 08;16(1):3 [FREE Full text] [doi: 10.1186/s12877-016-0183-8] [Medline: 26745961]
35. Selman L, McDermott K, Donesky D, Citron T, Howie-Esquivel J. Appropriateness and acceptability of a Tele-Yoga intervention for people with heart failure and chronic obstructive pulmonary disease: qualitative findings from a controlled pilot study. BMC Complement Altern Med 2015 Feb 07;15:21 [FREE Full text] [doi: 10.1186/s12906-015-0540-8] [Medline: 25887324]
36. Shulver W, Killington M, Morris C, Crotty M. 'Well, if the kids can do it, I can do it': older rehabilitation patients' experiences of telerehabilitation. Health Expect 2017 Feb;20(1):120-129 [FREE Full text] [doi: 10.1111/hex.12443] [Medline: 26890025]
37. Smaerup M, Grönvall E, Larsen SB, Laessoe U, Henriksen JJ, Damsgaard EM. Exercise gaming - a motivational approach for older adults with vestibular dysfunction. Disabil Rehabil Assist Technol 2017 Feb 04;12(2):137-144. [doi: 10.3109/17483107.2015.1104560] [Medline: 26727034]
38. Starkhammar S, Nygård L. Using a timer device for the stove: experiences of older adults with memory impairment or dementia and their families. Technol Disability 2008 Oct 08;20(3):179-191. [doi: 10.3233/tad-2008-20302]
39. van Hoof J, Kort H, Rutten P, Duijnstee M. Ageing-in-place with the use of ambient intelligence technology: perspectives of older users. Int J Med Inform 2011 May;80(5):310-331 [FREE Full text] [doi: 10.1016/j.ijmedinf.2011.02.010] [Medline: 21439898]
40. Holmberg M, Valmari G, Lundgren SM. Patients' experiences of homecare nursing: balancing the duality between obtaining care and to maintain dignity and self-determination. Scand J Caring Sci 2012 Dec 16;26(4):705-712. [doi: 10.1111/j.1471-6712.2012.00983.x] [Medline: 22420464]
41. Barken R. 'Independence' among older people receiving support at home: the meaning of daily care practices. Ageing Soc 2017 Oct 03;39(3):518-540. [doi: 10.1017/s0144686x17001039]
42. Hechinger M, Mayer H, Fringer A. Kenneth Gergen's concept of multi-being: an application to the nurse-patient relationship. Med Health Care Philos 2019 Dec 15;22(4):599-611 [FREE Full text] [doi: 10.1007/s11019-019-09897-4] [Medline: 30989571]
43. Büscher A. Negotiating Helpful Action: A Substantive Theory on the Relationship Between Formal & Informal Care. Tampere (Finland) : Tampere University Press; 2007.
44. Penrod J, Hupcey JE, Shipley PZ, Loeb SJ, Baney B. A model of caregiving through the end of life: seeking normal. West J Nurs Res 2012 Mar 14;34(2):174-193 [FREE Full text] [doi: 10.1177/0193945911400920] [Medline: 21403056]
45. Fringer A, Hechinger M, Schnepp W. Transitions as experienced by persons in palliative care circumstances and their families - a qualitative meta-synthesis. BMC Palliat Care 2018 Feb 05;17(1):22 [FREE Full text] [doi: 10.1186/s12904-018-0275-7] [Medline: 29402242]
46. Steinvall K, Johansson H, Berterö C. Balancing a changed life situation: the lived experience from next of kin to persons with inoperable lung cancer. Am J Hosp Palliat Care 2011 Mar 08;28(2):82-89. [doi: 10.1177/1049909110375246] [Medline: 20826491]
47. Steindal SA, Nes AA, Godskesen TE, Dihle A, Lind S, Winger A, et al. Patients' experiences of telehealth in palliative home care: scoping review. J Med Internet Res 2020 May 05;22(5):e16218 [FREE Full text] [doi: 10.2196/16218] [Medline: 32369037]

XSL•FO
RenderX

48. Rubeis G, Schochow M, Steger F. Patient autonomy and quality of care in telehealthcare. Sci Eng Ethics 2018 Feb 9;24(1):93-107. [doi: 10.1007/s11948-017-9885-3] [Medline: 28281148]
49. Rykkje L, Hjorth GH. "safety at home": experiences from testing of video communication between patients and home health care personnel. SAGE Open 2017 Dec 10;7(4):215824401774490 [FREE Full text] [doi: 10.1177/2158244017744900]
50. Dinç L, Gastmans C. Trust in nurse-patient relationships: a literature review. Nurs Ethics 2013 Aug 20;20(5):501-516. [doi: 10.1177/0969733012468463] [Medline: 23426234]

Abbreviations

CASP: Critical Appraisal Skills Program
ENTREQ: Enhancing Transparency in Reporting the Synthesis of Qualitative Research
PICo: Population, Phenomenon of Interest, and Context
PROSPERO: International Prospective Register of Systematic Reviews
SSOAR: Social Science Open Access Repository

Edited by T Leung; submitted 25.11.21; peer-reviewed by A Nes, SA Steindal, S Bhattacharjya; comments to author 07.02.22; revised version received 25.04.22; accepted 20.06.22; published 09.09.22.

Please cite as:
Hechinger M, Hentschel D, Aumer C, Rester C
A Conceptual Model of Experiences With Digital Technologies in Aging in Place: Qualitative Systematic Review and Meta-synthesis
JMIR Aging 2022;5(3):e34872
URL: https://aging.jmir.org/2022/3/e34872
doi:10.2196/34872
PMID:36083625

©Mareike Hechinger, Diana Hentschel, Christine Aumer, Christian Rester. Originally published in JMIR Aging (https://aging.jmir.org), 09.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Advance Planning for Technology Use in Dementia Care: Development, Design, and Feasibility of a Novel Self-administered Decision-Making Tool

Clara Berridge[1], MSW, PhD; Natalie R Turner[1], MSW; Liu Liu[2], MA, MEd; Sierramatice W Karras[3]; Amy Chen[3]; Karen Fredriksen-Goldsen[1], MSW, PhD; George Demiris[4], PhD

[1]School of Social Work, University of Washington, Seattle, WA, United States

[2]College of Education, University of Washington, Seattle, WA, United States

[3]Clinical Informatics Research Group, Biobehavioral Nursing and Health Informatics, University of Washington, Seattle, WA, United States

[4]School of Nursing and Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

Corresponding Author:
Clara Berridge, MSW, PhD
School of Social Work
University of Washington
4101 15th Ave NE
Seattle, WA, 98105
United States
Phone: 1 (206) 543 5640
Email: clarawb@uw.edu

Abstract

Background: Monitoring technologies are used to collect a range of information, such as one's location out of the home or movement within the home, and transmit that information to caregivers to support aging in place. Their surveilling nature, however, poses ethical dilemmas and can be experienced as intrusive to people living with Alzheimer disease (AD) and AD-related dementias. These challenges are compounded when older adults are not engaged in decision-making about how they are monitored. Dissemination of these technologies is outpacing our understanding of how to communicate their functions, risks, and benefits to families and older adults. To date, there are no tools to help families understand the functions of monitoring technologies or guide them in balancing their perceived need for ongoing surveillance and the older adult's dignity and wishes.

Objective: We designed, developed, and piloted a communication and education tool in the form of a web application called Let's Talk Tech to support family decision-making about diverse technologies used in dementia home care. The knowledge base about how to design online interventions for people living with mild dementia is still in development, and dyadic interventions used in dementia care remain rare. We describe the intervention's motivation and development process, and the feasibility of using this self-administered web application intervention in a pilot sample of people living with mild AD and their family care partners.

Methods: We surveyed 29 mild AD dementia care dyads living together before and after they completed the web application intervention and interviewed each dyad about their experiences with it. We report postintervention measures of feasibility (recruitment, enrollment, and retention) and acceptability (satisfaction, quality, and usability). Descriptive statistics were calculated for survey items, and thematic analysis was used with interview transcripts to illuminate participants' experiences and recommendations to improve the intervention.

Results: The study enrolled 33 people living with AD and their care partners, and 29 (88%) dyads completed the study (all but one were spousal dyads). Participants were asked to complete 4 technology modules, and all completed them. The majority of participants rated the tool as having the right length (>90%), having the right amount of information (>84%), being very clearly worded (>74%), and presenting information in a balanced way (>90%). Most felt the tool was easy to use and helpful, and would likely recommend it to others.

Conclusions: This study demonstrated that our intervention to educate and facilitate conversation and documentation of preferences is preliminarily feasible and acceptable to mild AD care dyads. Effectively involving older adults in these decisions and informing care partners of their preferences could enable families to avoid conflicts or risks associated with uninformed or disempowered use and to personalize use so both members of the dyad can experience benefits.

XSL•FO
RenderX

(JMIR Aging 2022;5(3):e39335) doi:10.2196/39335

KEYWORDS

Alzheimer disease; advance care planning; dyadic intervention; technology; remote monitoring; artificial intelligence; older adult; seniors; human-computer interaction; aging; elderly population; digital tool; educational tool; dementia care; ethics; informed consent

Introduction

Background

Reducing unnecessary care transitions while enabling aging in place is widely considered a priority in the context of a shortage of human and financial resources for elder care [1-3]. To address this challenge, policy makers, providers, and family caregivers are looking to technological solutions and investing in internet-connected devices that monitor the activity and safety of older adults with Alzheimer disease (AD) and AD-related dementias (ADRD), including technologies that involve artificial intelligence and natural language processing [4-12]. A problem that accompanies this shift is that the dissemination of technologies that passively collect and transmit personal data is outpacing our understanding of how to help families think about and involve older adults in decisions about if and how they want to be monitored. When families are not equipped to make informed decisions about technology use and to match a device to their needs and values, they are unlikely to use or benefit from it [4,13], and ill-informed decisions expose them to the technologies' risks.

Helping families navigate the complex technological landscape is a timely goal. Researchers often highlight the core ethical problem of achieving informed consent with an individual living with dementia [14-18]. They have demonstrated the need for tools to support education, awareness, and decision-making about technologies used to support care [16,19-21], including forward-looking consent processes before dementia undermines informed preference formation or expression [22]. This article reports on a novel self-administered intervention, Let's Talk Tech, to address this problem, which is modeled on advance care planning interventions. The goal is to meaningfully engage people living with mild dementia in planning for the use of technology in their care and to enable understanding of the implications of technology use and communication about it, so families are not left to navigate this complex space alone.

Let's Talk Tech supports decision-making about the use of these technologies and will advance the scientific understanding of how to engage people with early stage AD/ADRD in these decisions to enable their personalized use. To evaluate the intervention's feasibility, usability, and acceptability, we conducted a pilot study with 29 mild AD dementia care dyads. In this paper, we describe the development of the intervention to engage older adults in decisions about technology used in their care, report findings on study endpoints of feasibility and acceptability, and discuss key insights to support successful online intervention development with dementia care dyads.

Problems This Intervention Targets

The passive collection of location, audio, video, movements, and activities in elder care, and dementia home care in particular, is becoming more common [14,23]. Empirical and ethics research have established that the potential benefits of technologies with remote monitoring functions come with potential risks, and these have not been presented for consumers to understand. The dominant risks and challenges that are innate to passive monitoring are in conflicts among privacy, autonomy, freedom, and safety and risk management [24]. Potential risks include isolation through reduced human interaction and hands-on care, privacy invasion, loss of control, data inaccuracy, and reduced behavioral autonomy [14,16,17,22,25-36]. Research indicates that it is not easy for older adults on their own to appreciate what it will be like to be monitored [37]. Pragmatically, it is wasteful when families invest in technologies that do not ultimately work for them. Preventing this and maximizing the potential benefits require the right balance and respect for boundaries that are specific to each family or care partnership.

Potential for Dyadic Incongruence, Conflict, and Stress

The use of technologies that collect data, such as visual, activity, location, and audio, in dementia care may contribute to familial stress and conflict due to their surveilling nature. Studies also indicate the potential for incongruent preferences and difficulty navigating these decisions in a way that respects the values held by the older adult who the technology would be used on or with [17,27,30,37-40]. For example, in a dyadic study of Meals on Wheels clients and their primary family support person, adult children and their parents expressed conflicting views about how and when location tracking technology, in-home sensors, and web cameras should be used [27]. Adult children felt that involving their parents in conversations and decisions about whether to use these kinds of technologies would be complicated, underestimating their capacity to understand the technologies, and most felt that they would involve parents minimally [27].

Disagreement about treatment preferences has been associated with caregiver depression [41]. Dyadic strain has been associated with lower quality of life for African American dementia caregivers, and involvement in decision-making and concordance has been associated with quality of life for people living with dementia [42]. Incongruent appraisal of care values may contribute to worse quality of life for both members of the dementia care dyad [43].

Our hope, therefore, is that this communication tool will prepare care partners to make decisions that they feel confident about, support ethical application, and uphold the dignity and rights of people living with dementia. Involvement in planning recognizes the personhood of people living with dementia and the fact that they have preferences that can be expressed [44,45]. Care dyads require support to identify person-centered values

XSL•FO
RenderX

in relation to technologies and practical tools to select the options that align with those values as needs change.

There is no tool available to families to facilitate conversation, decisions, or planning for technology used to support elder care. Web-based behavioral interventions for older adults can be made freely available and have been found to be feasible and acceptable [46], yet there are few interventions that support dyadic dementia care partners to plan for care [47]. Dementia care apps have the potential to improve quality of life for people living with dementia and care partners [46], but their development requires consideration of a range of needs [48]. The research on usability and needs with regard to apps to support dementia caregiving or dyads is at an early stage [16,49-51]. The development of Let's Talk Tech was informed by the limited research on human-computer interaction–related best practices and techniques to support people living with dementia through online interventions [52]. As such, these pilot feasibility findings should help inform future design directions, particularly for web application development.

Methods

The Intervention

The intervention is in the form of a self-administered web application completed by a person with early stage AD (here, "person living with dementia") together with their primary care partner in an active collaborative process that involves education, discussion questions, and the documentation of the preferences expressed by the person living with dementia. Let's Talk Tech guides them through a discussion about what technologies they might want and under what conditions they would want them to be used. It requires no trained professional but requires that at least two people sit down together and use an internet-connected device to complete it.

The intervention's purpose is to facilitate communication and sharing of preferences similar to what has been developed for decisions about advance care planning, for the benefit of both members of the care dyad. Its components were modeled on elements of established advance care planning tools, including Your Life, Your Choices [53], and PREPARE [54]. Let's Talk Tech's flexibility to be retaken, referred to at a later date, and edited is informed by the research conducted by Sudore and Fried, in which decision-making was conceptualized as a dynamic process of communicating values [55]. An aim of the intervention, therefore, is to improve, as potential surrogates, care partners' knowledge of the technology preferences of the person living with dementia and related values to prepare them to make the best decisions in the future should the person lose capacity to participate [55].

Development of Let's Talk Tech

The web application intervention's content was developed from a 2-study process that generated cross-stakeholder input from key groups who had not previously been engaged in the same knowledge production process. First, in order to identify the technologies that should be included in the tool and information that could help families understand the implications of use, we employed the Delphi approach to achieve consensus from gerontechnology domain experts in the United States and Canada, to identify the salient risks and benefits associated with specific technologies predicted to be commonly used in home dementia care in the near future [14]. Domain experts also ranked these technologies according to those most likely to warrant a conversation with the person living with dementia to ensure acceptable use. We selected 4 technology categories from this list with an eye on the collective variability of data type (location, audio, visual, etc). The technologies featured in the web application are location tracking outside of the home, 4 activity sensors inside the home, web cameras, and artificial companion robots that use artificial intelligence and voice to interact with a person.

In that domain expert study, specific ways to mitigate prominent risks these technologies pose were also identified. These risk mitigation strategies applied to such diverse realms as design, policy, and regulation, and to interpersonal care practices, such as ensuring the ability to pause a device when one wants privacy or to be reminded about what information a given technology is collecting about them [14]. Five of the most commonly endorsed risk mitigation strategy options were incorporated into a survey for 825 people aged 21-92 years, with a mean age of 64 years (SD 13.13 years). The sample included a significant proportion of people who had memory problems or had been seen by doctors about memory concerns (n=201) [56]. The survey assessed the importance participants placed on the 5 actionable risk mitigation strategies for the use of these kinds of technologies in elder care [56]. Findings from the survey confirmed the very high importance and relevance of these 5 options to an older sample of people, including those with and those without reported memory problems [56].

The cumulative findings from the expert study and large survey of older adults were the building blocks of the intervention. The main components of the Let's Talk Tech web application are 4 featured technology modules (location tracking, in-home sensors, web cameras, and artificial companion robots). The goals of each module are to (1) clearly communicate the function of each technology, (2) clearly communicate the research-based prominent risks and benefits of using each, (3) prompt discussion between dyad members about their feelings, (4) document the preferences of the person living with dementia for use, nonuse, or conditioned use of each technology, and (5) document preferences for the use of alternatives to the featured technologies. Alternatives are offered to ensure that the option to use a given technology is presented as a true choice rather that the only acknowledged option to support care. Participants are presented with clear descriptions of the 4 data-diverse technology categories and prompted to discuss with each other their feelings about them. In order to help the dyad members appreciate what it might be like to use each, the web application presents prominent positive and negative implications for each technology, derived from the expert study [14], and assesses which are of most importance to the person living with dementia. The person living with dementia is then prompted to document their use preference for each, as well as the options that use may be contingent on. After the technology modules are completed, participants are guided through a series of general questions regarding the options that may be important to them, derived

XSL•FO
RenderX

from the survey research described above [56]. The web application provides a summary document that summarizes their choices and discussion. It can be accessed and edited any time at a later date.

For accessibility, the web application complies with the criteria of the Web Content Accessibility Guidelines (WCAG) 2.1 at Level AA. The WCAG represent the standard for web accessibility developed by the World Wide Web Consortium. The web application includes an automatic audio option to hear the content read aloud and is screen-reader compatible. To ensure that all aspects of accessibility for those living with mild dementia were considered, the design and content of the tool were shared and discussed with 10 human-computer interaction design experts and clinicians who work with people living with dementia. The clinicians included a neurologist, a neuropsychologist, physicians (eg, geriatricians and palliative care experts), gerontological social workers, and nurses. For example, Let's Talk Tech's content was designed to avoid abstractions, and sentences were reviewed for clarity, singular constructs, and shorter length to enable comprehension, in addition to word choice. The verbiage used to introduce the web application on its home page is provided in Textbox 1.

Textbox 1. An introductory message of Let's Talk Tech.

Why talk about technology now?

There are many ways to add support to help someone live independently. One is to use technology.

Some technologies collect information about a person to help a family member or caregiver monitor them. The only person who knows if you're comfortable about any of these technologies is you. Information about what it can be like to use these technologies could help you decide how you feel.

That's why this can help. Having conversations about technology choices can help you think about what you want or don't want. Letting people know how you feel can make it easier for them to follow your choices when deciding about using technologies in the future.

The reason to do this together is this can be a shared decision.

Participants

The inclusion criteria for older adults were as follows: (1) enrollment in the University of Washington Alzheimer's Disease Research Center (UW ADRC) clinical core or research registry with a diagnosis of mild AD dementia; (2) age 55+ years; (3) English speaking; and (4) having a care partner identified as a primary support person willing to participate in the study. The inclusion criteria for care partners were as follows: (1) co-participant of an ADRC clinical core patient or research registry patient who has been diagnosed with mild AD dementia; (2) identification by a study participant aged 55+ years as someone who is their primary support person; (3) age 18+ years; and (4) English speaking. Between the 2 potential dyad participants, one had to have access to a device (such as a computer, laptop, or tablet) that they could use together, which had an internet connection. Twenty-nine dyads participated in the study. Each individual participant received a Visa gift card for US $150 for their time upon completion of the 3 steps described below.

Ethics Approval

The study received approval from the University of Washington Division of Human Subjects (study number: STUDY00014226). Informed consent was obtained from each participant.

Procedures

Reported in this paper are the feasibility findings for recruitment, enrollment, and retention, and the survey questions and interviews that assess the acceptability of Let's Talk Tech. Acceptability was assessed at time 2 (T2) after use of the intervention, using 7 survey questions to measure satisfaction, quality, and usability. The specific items are presented with their outcomes in the Results section. T2 surveys were followed immediately by dyadic interviews that probed further about participants' responses to the acceptability questions. The interviews allowed us to learn about specific components of the web application that worked or did not work well for each dyad, and to identify areas for improvement. The interview portion lasted an average of 33 min (range 15-75 min). Among the dyads, 65% completed these interviews by Zoom video and 35% completed in person. All interviews were audio recorded with permission.

Procedures for the pilot study as a whole involved the following 3 steps: (1) time 1 (T1) study questionnaire completion individually with the researcher present to support the person living with dementia, if needed; (2) web application completion together as a dyad without the researcher present, and (3) T2 questionnaire completion individually with the researcher present to support the person living with dementia, followed immediately by an interview with the dyad. Questionnaires were administered via REDCap, and printed copies were used for those who requested it. The web application was self-administered, and no researcher was present or assisted dyads with it, apart from showing them how to access it during T1. However, the set of T1 and T2 surveys relied on a researcher to administer the questions to the people living with dementia. Care partners independently completed their surveys in REDCap and a researcher stayed with the person living with dementia to answer clarifying questions as they completed their own surveys in REDCap. In our case, the researcher was a licensed master social worker with clinical experience working with people living with dementia and their care partners.

The study outcome measures, which are not reported here, included 27 questions for care partners and 7 questions for people living with dementia unique to this study to assess knowledge, understanding, and preparedness to make decisions about technology use (primary efficacy outcomes). Both participant groups also completed 2 subscales of the Dyadic Relationship Scale to measure positive dyadic interaction and strain [57] (secondary outcomes) and the Decision-Making

XSL•FO
RenderX

Involvement Scale assessing the level of involvement of people living with dementia in daily decisions [58] for descriptive purposes. Care partners were administered the General Anxiety Disorder-7 (GAD-7) [59] to confirm that the intervention would not increase anxiety, and the Stetz Inventory to describe this participant group's level of involvement with caregiving tasks [60,61].

Analysis

Analyses for descriptive statistics and frequency counts were performed in R (R Core Team). Frequency counts were used to summarize participant T2 feasibility and acceptability results, and T2 transcribed interviews were coded in Dedoose (Version 9.0.17; SocioCultural Research Consultants, LLC). Two coders used a process of thematic analysis to identify themes regarding participants' experiences with the web application and suggestions for improvement [62,63]. A codebook was developed based on the interview guide followed by initial coding by a primary coder who developed inductive codes in the process. The new codes were incorporated, and a secondary coder then reviewed the coding decisions and the 2 discussed discrepancies and reached consensus about them [63]. The pair then read the coded excerpts across interviews and identified themes related to outcomes of feasibility and acceptability.

Results

Feasibility

Recruitment, Enrollment, and Retention

Recruitment was conducted through 2 existing university research volunteer pools who had consented to be contacted about potential participation in other studies. As part of the UW ADRC's operations, both the clinical core patient participants and their co-participants have annual visits with UW ADRC, and the status of mild AD dementia is reassessed. The ADRC prescreened participants in their research registry group to identify those with mild AD dementia and those with mild AD who also had a co-participant (here, "care partner") volunteer for our recruitment list. Because the UW ADRC diagnoses the patients and reassesses them annually and because the center has diagnosis and severity information for the participants, there was no further assessment to determine cognitive impairment status.

Those who were identified by the ADRC as having a diagnosis of mild AD dementia were invited to participate with their care partner by phone or email according to their preferences for a total of 110 people living with dementia invited. Thirty did not respond to the invitation, and we do not know the reasons for their nonparticipation. The reasons for nonparticipation among respondents were as follows: care partners determined that the people living with dementia had dementia too far advanced (n=11), not interested (n=11), not a good time (n=9), and lack of a device or comfort using a computer (n=2). Thirty-three people living with dementia enrolled with their care partners, and 29 (88%) dyads completed the study. Of the 4 dyads who did not complete the study, 1 dropped out before T1 because of difficulties with a recent move to memory care, 2 dropped out during T1 because the standardized survey scales were too difficult for the people living with dementia, and 1 dropped out after T1 because of computer difficulties generally and because the care partner had an overwhelming health change.

Age, gender, race, and ethnicity reported by both people living with dementia and care partners are presented in Table 1. Care partners were mainly spouses, and 1 care partner was an adult daughter. Participants wrote in their gender identity. Among the participants, 38% (11/29) of care partners and 62% (18/29) of people living with dementia were male. The age of care partners ranged from 55 to 83 years (mean 68 years, SD 6.73 years), and the age of people living with dementia ranged from 59 to 82 years (mean 70 years, SD 7.06 years). Only 3 participants did not identify as non-Hispanic white (2 Asian American care partners and 1 African American person living with dementia). Data on Hispanic/Latino ethnicity were missing for 7 care partners.

Table 1. Participant characteristics.

Demographics	Care partner (N=29)	People living with dementia (N=29)
Age (years), mean (SD); range	68 (6.73); 55-83	70 (7.06); 59-82
Gender, n (%)		
Male	11 (38)	18 (62)
Female	18 (62)	11 (38)
Race, n (%)		
White	27 (93)	28 (97)
African American	0 (0)	1 (3)
Asian	2 (7)	0 (0)
Hispanic/Latino ethnicity, n (%)	0 (0)	0 (0)

Completion

Our a priori cut point, at which the intervention is considered complete, was if the dyad completed at least three of the four technology modules. Participants were asked to complete all modules, but this was a self-administered intervention in which the researcher was not present while the dyads worked through the web application. As such, we did not expect the high completion rate of 100% for the 4 modules. As a group, participants completed 98.4% of the primary 17 questions asked

XSL•FO
RenderX

in the web application (a total of 485 of 493 nonskip logic follow-up questions). Two participants did not answer 1 question each, 1 did not answer 2 questions, and 1 did not answer 4 questions. Only 2 of the dyads reported spreading the web application over 2 sessions. On average, the time between T1 and T2 was 16 days. The average time between web application completion and T2 was 4 days.

Acceptability: Satisfaction, Quality, and Usability

Satisfaction, quality, and usability were measured with Likert response item questions and follow-up interviews to probe responses. Satisfaction was measured using the following questions answered on 5-item Likert scales: "How helpful was the tool?" ("Extremely unhelpful" to "Extremely helpful") and "How likely would you be to recommend this tool to others living with dementia or their caregivers?" ("Extremely unlikely" to "Extremely likely"). Quality was assessed with the following questions: "Was the tool balanced?" ("Slanted in favor of using the technology," "Slanted against using the technology," and "Balanced") and "Was there enough information to help you decide about how to answer the questions?" ("Too much information," "Too little information," and "Just right"). Questions about ease of use, clarity, and length describe usability as follows: "How easy was it to use this tool?" ("0 [very easy]" to "10 [very hard]"), "Were the descriptions clearly worded?" ("Very clearly," "Somewhat clearly," and "Not clearly"), and "Please rate the tool's length" ("Too long," "Too short," and "Just right"). These findings are presented in Table 2. Semistructured dyadic interviews immediately followed this questionnaire to probe these responses and to learn about participants' experiences with the intervention. We also present interview themes that provide greater insight into survey responses about feasibility and acceptability.

As depicted in Table 2, all care partners answered all satisfaction, quality, and usability questions, and depending on the question, 4 to 5 people living with dementia did not answer because they reported difficulty remembering the web application experience well enough to answer the questions. Both participant groups generally reported that Let's Talk Tech's length was just right. Overall, 80% (23/29) of care partners and 68% (17/25) of people living with dementia who answered the question felt that the descriptions were very clearly worded. One care partner reported that the descriptions were not clearly worded. Moreover, 86% (25/29) of care partners and 83% (20/24) of people living with dementia said that the amount of information was just right, while 14% (4/29) of care partners and 13% (3/24) of people living with dementia said there was too little information, with 1 person living with dementia reporting too much information. Some dyads specifically noted the need for more concrete and visual examples, particularly about what an artificial companion robot could do. Some participants suggested ways to enable a deeper dive into the technologies in each module for those who wanted to learn even more, including how to find a device or product on the market.

XSL•FO
RenderX

Table 2. Feasibility measures of satisfaction, quality, and usability for care partners and people living with dementia at time 2.

Question and responses	Care partners (N=29), n (%)		People living with dementia (N=29), n (%)	
	Completed	Missing	Completed	Missing
Please rate the tool's length	29 (100)	0 (0)	24 (83)	5 (17)
Too long	2 (7)		1 (4)	
Too short	0 (0)		1 (4)	
Just right	27 (93)		22 (92)	
Were the descriptions clearly worded?	29 (100)	0 (0)	25 (86)	4 (14)
Not clearly	1 (3)		0 (0)	
Somewhat clearly	5 (17)		8 (32)	
Very clearly	23 (80)		17 (68)	
Was there enough information to help you decide about how to answer the questions?	29 (100)	0 (0)	24 (83)	5 (17)
Too much information	0 (0)		1 (4)	
Too little information	4 (14)		3 (13)	
Just right	25 (86)		20 (83)	
Was the tool balanced?	29 (100)	0 (0)	24 (83)	5 (17)
Slanted in favor of using the technology	2 (7)		3 (13)	
Slanted against using the technology	0 (0)		0 (0)	
Balanced	27 (93)		21 (87)	
How helpful was the tool?	29 (100)	0 (0)	24 (83)	5 (17)
Extremely unhelpful	0 (0)		0 (0)	
Unhelpful	0 (0)		1 (4)	
Neutral	2 (7)		3 (13)	
Helpful	25 (86)		18 (75)	
Extremely helpful	2 (7)		2 (8)	
How likely would you be to recommend this tool to others living with dementia or their caregivers?	29 (100)	0 (0)	25 (86)	4 (14)
Extremely unlikely	1 (4)		0 (0)	
Unlikely	0 (0)		1 (4)	
Neutral	5 (17)		5 (20)	
Likely	18 (62)		14 (56)	
Extremely likely	5 (17)		5 (20)	

Participants felt Let's Talk Tech was balanced, except for 7% (2/29) of care partners and 13% (3/24) of people living with dementia who felt it was slanted in favor of using the technology. None felt it was slanted against use. A couple of participants noted that having an intervention that has a purpose to encourage discussion about technology options causes bias toward technology (eg, "maybe it's the fact that here's some offer of technology to help. You know, not that you're pushing it but it's there. So it feels like it's an automatic pro for the technology"). Others appreciated the neutralizing features of the tool, specifically, presentation of nontechnology alternative options to support care, as well as both positive and negative aspects of each technology. One care partner explained, "the pros and cons examples were very good and I think those are very important. Because otherwise it can be very leading…I thought you did a good job, because otherwise, if you just list all the pros your brain goes that way." Another care partner elaborated, "the format's conducive to being honest with it. It doesn't promote trying to gain anything. It's pretty neutral that way."

On a scale of extremely unhelpful to extremely helpful, 86% (25/29) of care partners and 75% (18/24) of people living with dementia rated the intervention as helpful, with 2 in each group rating it as extremely helpful, and 2 care partners and 3 people living with dementia selecting neutral. One person living with dementia rated it as unhelpful. Additionally, 79% (23/29) of care partners and 76% (19/25) of people living with dementia were likely or extremely likely to recommend Let's Talk Tech, while 1 person living with dementia was unlikely and 1 care

XSL•FO
RenderX

partner was extremely unlikely to recommend Let's Talk Tech. Interviews revealed that most of the individuals who reported that they would be likely to recommend it to others cited the benefit of awareness gained about technological tools that may be helpful and the support with having conversations about them, and some who were not inclined felt it would not be their business to make such a recommendation.

Roughly half of the dyads reported some discomfort in completing Let's Talk Tech, noting that thinking about the need for technologies is scary or unsettling, that any disagreement is hard, or that it can bring up worries about being a burden for people living with dementia. However, all stated that it was worth the discomfort. For example, a dyad explaining that it makes people living with dementia very sad to talk about advance planning for care support, discussed why that was worthwhile as follows:

> *And by having these conversations, makes it easier for both of us, because then we're not guessing.* [Care partner]
>
> *It's true and, and the more we're able to talk about it, the more comfortable it is, that okay, this is just how things are now and it's okay.* [Person living with dementia]
>
> *And we can joke about it.* [Care partner]

When asked directly if it was worth the sadness people living with dementia felt, this person living with dementia responded, "Oh absolutely yes. Yes, because it's something to get through. And the only way to get to the other side is to talk about it and yeah absolutely."

Participants were asked how easy it was to use Let's Talk Tech. Six of the people living with dementia did not remember it well enough to answer. Figure 1 presents a visual comparison of the 2 participant group ratings of ease of use. It was harder for people living with dementia as compared with care partners, though both groups primarily reported it as somewhat easy. Four participants (3 people living with dementia and 1 care partner) rated the tool difficult to use (score of 6 or greater). When asked about this rating, 1 person living with dementia said it was because this was the first time she was thinking about this topic and she was trying to wrap her head around the technology. Another had trouble remembering that he had used the web application and was in pain during the interview, so he did not expand on the reason for his rating. The dyad that rated the tool difficult to use reported difficulty in relating scenarios specified in the tool to their own lives and felt that they were too broad.

Interviews confirmed the survey findings that the intervention is most usable and useful during early/mild stages of dementia when using the tool is not too onerous for people living with dementia, and it is easily navigated with questions well understood. Some people living with dementia felt that they were not at a stage of their disease that warranted the use of the featured technologies and thus had difficulty relating to the questions about their preferences for them as they felt they were not needed. Not all struggled with this, but participants from 12 dyads (11 care partners and 4 people living with dementia) recommended including more scenarios to enable people to imagine times in the future when their responses or preferences may change.

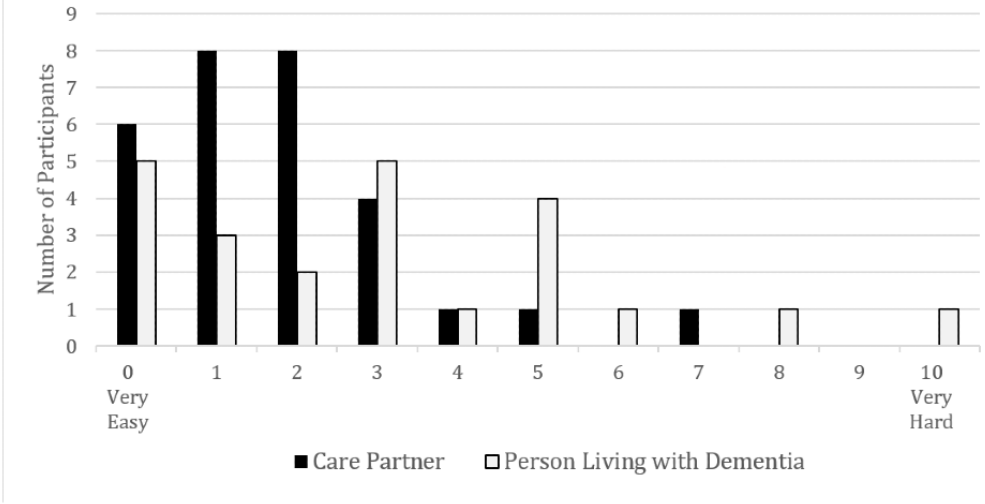
Caregivers also indicated that they felt that the disease stage would impact the person's answers to the questions posed in the web application. While recruitment was conducted with those identified by the ADRC as in a mild stage of AD, we did not conduct additional tests to confirm the current status. Two care partners explained in their T2 interviews that they believed the patients were in the middle stages of the disease. One care partner explained why she thought the ideal time to use Let's Talk Tech would be at an early stage:

> *I feel like we're like moderate like in the middle stages like right in the middle of the middle stage, and so I almost think that in the early stages of, of Alzheimer's or like right in the beginning of the moderate stage. I mean he can still answer the questions now. It just takes a lot of like rephrasing.*

This person felt that had the patient still been at an early stage, he would have been able to answer with better judgment, a more accurate understanding of his own condition, and greater consideration of the demands on her as a care partner, and would have felt less worried about being judged (amplified via a camera, for example) than he was at this moderate stage. Care partners who doubted their partners' comprehension often also doubted the validity of their responses, making the intervention less helpful as a planning tool for those participants whose AD had advanced beyond the early stage.

XSL•FO
RenderX

Figure 1. Web application tool ease of use. Frequency responses to "How easy was it to use this tool?".



Discussion

Principal Findings

The findings of this study represent promising feasibility data for a self-administered web application intervention designed for people living with mild AD/ADRD and a care partner. Participants were able to navigate through the entirety of Let's Talk Tech and perceived value in the discussions it facilitated despite some discomfort with advance care planning. The high completion of all aspects of the web application was particularly encouraging because we anticipated that working through all modules in one sitting could be a challenge for people living with dementia. Only 2 dyads reported splitting their session with the application into 2 sittings. Participants may have completed all web application forms because they were asked to as part of the introduction of the study, and we should thus not expect such a high rate of completion outside of a study context. It is likely that in a real-world nonstudy context, participants may only complete those modules that seem of particular interest or relevance to them. Still, the successful completion of the Let's Talk Tech intervention that dyads achieved, primarily in 1 sitting, indicates that the intervention is not too strenuous for care partners or people living with mild AD and is well targeted for this group.

Having difficult conversations was not reportedly a problem for our sample. The interviews described that it was uncomfortable for some, but not so uncomfortable that it outweighed the benefits of having these conversations. This is an important element of feasibility and a promising finding that people may accept this intervention as an opportunity to have conversations they feel are important, though difficult to facilitate on one's own without such a tool.

Limitations and Implications for Future Work

While the feasibility and acceptability ratings were all high, some limitations of the intervention were illuminated by participants through interviews. First, while our findings clearly indicated that Let's Talk Tech is very well targeted to people living with early stages of AD, a difficulty for people living with mild dementia is that they may not yet feel that there is need for the technologies featured in the intervention. Sometimes there may be disagreement with a care partner about this if they assess their condition differently. Adding future-oriented scenarios would be a clear response to this issue, and participants suggested this directly; however, research also shows that people have a very difficult time projecting themselves into future scenarios with accuracy [64]. Another complication of this potential approach is that a common symptom of AD/ADRD is difficulty with abstract thinking, which makes advance planning and imagining oneself in future or imaginary scenarios challenging [65]. Still, the need to enable the expression of preferences for future scenarios in addition to current use was a strong interview theme, indicating that more research is needed on how to enable this in a way that meets the needs of both members of the care dyad.

Second, 2 care partners described their partners as being at mid-stage and no longer at the early stages of dementia, and those individuals had difficulty with comprehension. These participants still completed Let's Talk Tech, but care partners reported more work to navigate it to the point where it could become too onerous and where the responses of people living with dementia could be deemed less reliable by care partners. This underscores our finding that this self-administered web application is well suited for people who have not yet progressed to moderate stages of AD dementia. This also suggests that more research is needed to find ways to engage dyads at moderate stages, such as additional support to answer questions.

Third, the finding that the intervention's bias toward technology use was mitigated by not naming specific products or devices was not consistent with the finding that dyads would have considered photographs useful for comprehension, and many care partners desired links and next steps to find devices for

XSL•FO
RenderX

purchase. While they often had enough information to form preferences, some people living with dementia and care partners reported that they lacked clarity about the scope of what an artificial companion robot could do. This is unsurprising given the relatively low levels of algorithmic awareness [21] and lower familiarity with a more recently developed technology, such as artificial companion robots, relative to location tracking and other featured technologies. It indicates that focus is needed on how to describe the capabilities of such a device, and possibly others using artificial intelligence and natural language processing specifically, in ways that are more likely to be clear to people living with mild dementia and their care partners. Because algorithmic awareness is also associated with categories of socioeconomic status and may be associated with race and ethnicity, it will be important in future studies to collect education, income, and wealth data, and to ensure racial diversity in study samples. A limitation of this pilot study is that we were unable to examine potential associations by these categories.

Fourth, obtaining feedback on the web application days after completion from people living with dementia was sometimes challenging owing to their difficulty with short-term memory. Because it is critical that researchers understand this participant group's experience with the intervention, creative solutions, such as soliciting real-time feedback during or immediately after use of the intervention, are required.

Finally, having a dementia-trained researcher to administer the survey portions of this study was helpful to guide people living with dementia through a long set of surveys, clarify interview questions, and be sensitive to signals that it was time to stop or pause. Because of this researcher's clinical experience, we were able to closely observe the points at which participants living with dementia reached their limit with regard to answering research questions. We found that speaking beyond 30 to 35 minutes was sometimes difficult for people living with dementia, at which point answering questions started to become cognitively taxing. They often reported a lack of attention after that point or feeling tired. This observation may be informative for other intervention studies involving people living with mild AD.

Conclusion

The use of in-home monitoring technologies to predict health problems and support aging in place is growing faster than our understanding of how to help families make decisions about how and when to use them. Our pilot study findings demonstrate strong preliminary feasibility and acceptability of the Let's talk Tech intervention for promoting informed shared decision-making about technologies used in dementia care. Successful recruitment, enrollment, and retention, and 100% completion of the web application intervention demonstrate strong feasibility. Good ratings were given for the satisfaction, quality, and usability measures of acceptability. Our findings also revealed useful considerations for other self-administered web application interventions for people living with mild AD and care partners, including optimal exit interview time and the potential need for immediate feedback processes upon intervention completion. Most importantly, this pilot study demonstrated that a self-administered dyadic intervention in the form of a web application can be successfully independently completed in 1 sitting by mild AD care dyads. This research advances the scientific understanding of how to engage people living with dementia in decisions while helping families navigate a complex technology landscape.

Acknowledgments

We thank Jeffrey Kaye, Karen Lyons, Carolyn Parsey, Elizabeth Lindley, Karen Clay, Batya Friedman, and David Hendry for providing insightful feedback on the intervention's content and design, as well as members of the development team at the University of Washington Clinical Informatics Research Group, William B Lober, Justin McReynolds, and Winter Roberts. This research was supported by the National Institute of Aging (principal investigator: CB; NIA K01AG062681) and the University of Washington Alzheimer's Disease Research Center (P30AG066509). The intervention Let's Talk Tech has been developed by the author CB (©Clara Berridge 2021).

Data Availability

The data sets generated during this study are not publicly available due to the limitations on data reuse as part of the consenting procedures.

Conflicts of Interest

None declared.

References

1. Seegert L. White House conference on aging: Aging in place, falls, and elder abuse. Am J Nurs 2015 Oct;115(10):13. [doi: 10.1097/01.NAJ.0000471922.64720.fd] [Medline: 26402269]
2. Stone R, Harahan MF. Improving the long-term care workforce serving older adults. Health Aff (Millwood) 2010;29(1):109-115. [doi: 10.1377/hlthaff.2009.0554] [Medline: 20048368]
3. Wiles JL, Leibing A, Guberman N, Reeve J, Allen RES. The meaning of "aging in place" to older people. Gerontologist 2012 Jun;52(3):357-366. [doi: 10.1093/geront/gnr098] [Medline: 21983126]

XSL•FO
RenderX

4. Ienca M, Wangmo T, Jotterand F, Kressig RW, Elger B. Ethical design of intelligent assistive technologies for dementia: A descriptive review. Sci Eng Ethics 2018 Aug;24(4):1035-1055. [doi: 10.1007/s11948-017-9976-1] [Medline: 28940133]
5. Frennert S. Moral distress and ethical decision-making of eldercare professionals involved in digital service transformation. Disabil Rehabil Assist Technol 2020 Nov 05:1-10. [doi: 10.1080/17483107.2020.1839579] [Medline: 33151763]
6. Kaye J. Public Policy Aging Rep 2017;27(2):53-61 [FREE Full text] [doi: 10.1093/ppar/prx005] [Medline: 31148911]
7. IP Video Surveillance and VSaaS Market Expected to Reach $57.3 Billion by 2020 - Allied Market Research. PRNewswire. 2014. URL: http://www.securityhive.com/ip-video-surveillance-vsaas-market-expected-reach-57-3-billion-2020-allied-market-research/ [accessed 2016-10-13]
8. Lukkien DRM, Nap HH, Buimer HP, Peine A, Boon WPC, Ket JCF, et al. Towards responsible artificial intelligence in long-term care: A scoping review on practical approaches. Gerontologist 2021 Dec 06:gnab180. [doi: 10.1093/geront/gnab180] [Medline: 34871399]
9. Rubeis G. The disruptive power of Artificial Intelligence. Ethical aspects of gerontechnology in elderly care. Arch Gerontol Geriatr 2020 Jul 15;91:104186. [doi: 10.1016/j.archger.2020.104186] [Medline: 32688106]
10. Schulz R, Wahl H, Matthews JT, De Vito Dabbs A, Beach SR, Czaja SJ. Advancing the aging and technology agenda in gerontology. Gerontologist 2015 Oct;55(5):724-734 [FREE Full text] [doi: 10.1093/geront/gnu071] [Medline: 25165042]
11. Steventon A, Bardsley M, Billings J, Dixon J, Doll H, Beynon M, et al. Effect of telecare on use of health and social care services: findings from the Whole Systems Demonstrator cluster randomised trial. Age Ageing 2013 Jul;42(4):501-508 [FREE Full text] [doi: 10.1093/ageing/aft008] [Medline: 23443509]
12. Peine A, Marshall B, Martin W, Neven L. Socio-gerontechnology. Abingdon-on-Thames, Oxfordshire, England, UK: Routledge; 2021.
13. Sun M, Burke LE, Mao Z, Chen Y, Chen H, Bai Y, et al. eButton: A wearable computer for health monitoring and personal assistance. Proc Des Autom Conf 2014;2014:1-6 [FREE Full text] [doi: 10.1145/2593069.2596678] [Medline: 25340176]
14. Berridge C, Demiris G, Kaye J. Domain experts on dementia-care technologies: Mitigating risk in design and implementation. Sci Eng Ethics 2021 Feb 18;27(1):14 [FREE Full text] [doi: 10.1007/s11948-021-00286-w] [Medline: 33599847]
15. Mahoney DF, Purtilo RB, Webbe FM, Alwan M, Bharucha AJ, Adlam TD, Working Group on Technology of the Alzheimer's Association. In-home monitoring of persons with dementia: ethical guidelines for technology research and development. Alzheimers Dement 2007 Jul;3(3):217-226. [doi: 10.1016/j.jalz.2007.04.388] [Medline: 19595941]
16. Meiland F, Innes A, Mountain G, Robinson L, van der Roest H, García-Casal JA, et al. Technologies to support community-dwelling persons with dementia: A position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil Assist Technol 2017 Jan 16;4(1):e1 [FREE Full text] [doi: 10.2196/rehab.6376] [Medline: 28582262]
17. Novitzky P, Smeaton AF, Chen C, Irving K, Jacquemard T, O'Brolcháin F, et al. A review of contemporary work on the ethics of ambient assisted living technologies for people with dementia. Sci Eng Ethics 2015 Jun;21(3):707-765. [doi: 10.1007/s11948-014-9552-x] [Medline: 24942810]
18. Robillard JM, Cleland I, Hoey J, Nugent C. Ethical adoption: A new imperative in the development of technology for dementia. Alzheimers Dement 2018 Sep;14(9):1104-1113. [doi: 10.1016/j.jalz.2018.04.012] [Medline: 29937247]
19. Berridge C. Medicaid becomes the first third-party payer to cover passive remote monitoring for home care: Policy analysis. J Med Internet Res 2018 Feb 21;20(2):e66 [FREE Full text] [doi: 10.2196/jmir.9650] [Medline: 29467120]
20. Mulvenna M, Hutton A, Coates V, Martin S, Todd S, Bond R, et al. Views of caregivers on the ethics of assistive technology used for home surveillance of people living with dementia. Neuroethics 2017;10(2):255-266 [FREE Full text] [doi: 10.1007/s12152-017-9305-z] [Medline: 28725288]
21. Gran A, Booth P, Bucher T. To be or not to be algorithm aware: a question of a new digital divide? Information, Communication & Society 2020 Mar 09;24(12):1779-1796. [doi: 10.1080/1369118x.2020.1736124]
22. Thorstensen E. Privacy and Future Consent in Smart Homes as Assisted Living Technologies. In: Zhou J, Salvendy G, editors. Human Aspects of IT for the Aged Population. Applications in Health, Assistance, and Entertainment. ITAP 2018. Lecture Notes in Computer Science, vol 10927. Cham: Springer; 2018:415-433.
23. Fabricatore C, Radovic D, Lopez X, Grasso-Cladera A, Salas CE. When technology cares for people with dementia: A critical review using neuropsychological rehabilitation as a conceptual framework. Neuropsychol Rehabil 2020 Sep 21;30(8):1558-1597. [doi: 10.1080/09602011.2019.1589532] [Medline: 30896282]
24. Niemeijer AR, Frederiks BJM, Depla MFIA, Legemaate J, Eefsting JA, Hertogh CMPM. The ideal application of surveillance technology in residential care for people with dementia. J Med Ethics 2011 May 02;37(5):303-310. [doi: 10.1136/jme.2010.040774] [Medline: 21292695]
25. Berridge C. Breathing room in monitored space: The impact of passive monitoring technology on privacy in independent living. Gerontologist 2016 Oct;56(5):807-816. [doi: 10.1093/geront/gnv034] [Medline: 26035900]
26. Berridge C. Active subjects of passive monitoring: responses to a passive monitoring system in low-income independent living. Ageing Soc 2017 Mar;37(3):537-560 [FREE Full text] [doi: 10.1017/S0144686X15001269] [Medline: 28239211]
27. Berridge C, Wetle TF. Why older adults and their children disagree about in-home surveillance technology, sensors, and tracking. Gerontologist 2020 Jul 15;60(5):926-934. [doi: 10.1093/geront/gnz068] [Medline: 31102442]

XSL•FO
RenderX

28. Chung J, Demiris G, Thompson HJ. Ethical considerations regarding the use of smart home technologies for older adults: An integrative review. Annu Rev Nurs Res 2016;34:155-181. [doi: 10.1891/0739-6686.34.155] [Medline: 26673381]
29. Deliverable 7: Final Research Report. EFORTT: Ethical Frameworks for Telecare technologies for older people at home. Lancaster University. URL: https://www.lancaster.ac.uk/efortt/documents/Deliverable%207%20Final%20Research%20report.pdf [accessed 2022-07-18]
30. Ho A. Are we ready for artificial intelligence health monitoring in elder care? BMC Geriatr 2020 Sep 21;20(1):358 [FREE Full text] [doi: 10.1186/s12877-020-01764-9] [Medline: 32957946]
31. Jaschinski C, Allouch SB. Ambient assisted living: Benefits and barriers from a user-centered perspective. In: AMBIENT 2014 : The Fourth International Conference on Ambient Computing, Applications, Services and Technologies. 2014 Presented at: Fourth International Conference on Ambient Computing, Applications, Services and Technologies; August 2014; Rome, Italy.
32. Mortenson WB, Sixsmith A, Woolrych R. The power(s) of observation: Theoretical perspectives on surveillance technologies and older people. Ageing Soc 2015 Mar;35(3):512-530 [FREE Full text] [doi: 10.1017/S0144686X13000846] [Medline: 29307944]
33. Sanders C, Rogers A, Bowen R, Bower P, Hirani S, Cartwright M, et al. Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: a qualitative study. BMC Health Serv Res 2012 Jul 26;12:220 [FREE Full text] [doi: 10.1186/1472-6963-12-220] [Medline: 22834978]
34. Wrede C, Braakman-Jansen A, van Gemert-Pijnen L. Requirements for unobtrusive monitoring to support home-based dementia care: Qualitative study among formal and informal caregivers. JMIR Aging 2021 Apr 12;4(2):e26875 [FREE Full text] [doi: 10.2196/26875] [Medline: 33843596]
35. Lariviere M, Poland F, Woolham J, Newman S, Fox C. Placing assistive technology and telecare in everyday practices of people with dementia and their caregivers: findings from an embedded ethnography of a national dementia trial. BMC Geriatr 2021 Feb 15;21(1):121 [FREE Full text] [doi: 10.1186/s12877-020-01896-y] [Medline: 33588768]
36. Zwijsen SA, Niemeijer AR, Hertogh CMPM. Ethics of using assistive technology in the care for community-dwelling elderly people: an overview of the literature. Aging Ment Health 2011 May;15(4):419-427. [doi: 10.1080/13607863.2010.543662] [Medline: 21500008]
37. Berridge C. Selling Passive Monitoring to Manage Risk in Independent Living: Frontline Workers in a Bind. In: Adams S, Purtova N, Leenes R, editors. Under Observation: The Interplay Between eHealth and Surveillance. Law, Governance and Technology Series, vol 35. Cham: Springer; 2017:73-90.
38. Godwin B. The ethical evaluation of assistive technology for practitioners: a checklist arising from a participatory study with people with dementia, family and professionals. Journal of Assistive Technologies 2012;6(2):123-135. [doi: 10.1108/17549451211234975]
39. Kenner AM. Securing the elderly body: Dementia, surveillance, and the politics of "aging in place". Surveillance & Society 2008;5(3):252-269. [doi: 10.24908/ss.v5i3.3423]
40. Mort M, Roberts C, Callén B. Ageing with telecare: care or coercion in austerity? Sociol Health Illn 2013 Jul;35(6):799-812 [FREE Full text] [doi: 10.1111/j.1467-9566.2012.01530.x] [Medline: 23094945]
41. Tsai C, Lee Y, Lee W, Hwang J, Wang S, Fuh J. Depression of family caregivers is associated with disagreements on life-sustaining preferences for treating patients with dementia. PLoS One 2015;10(7):e0133711 [FREE Full text] [doi: 10.1371/journal.pone.0133711] [Medline: 26230958]
42. Bonds K, Whitlatch CJ, Song M, Lyons KS. Factors influencing quality of life in African-American dementia dyads. Aging Ment Health 2021 Apr;25(4):703-710 [FREE Full text] [doi: 10.1080/13607863.2020.1711865] [Medline: 31920090]
43. Moon H, Townsend AL, Whitlatch CJ, Dilworth-Anderson P. Quality of life for dementia caregiving dyads: Effects of incongruent perceptions of everyday care and values. Gerontologist 2017 Aug 01;57(4):657-666 [FREE Full text] [doi: 10.1093/geront/gnw055] [Medline: 27048708]
44. Whitlatch CJ, Piiparinen R, Friss Feinberg L. How well do family caregivers know their relatives' care values and preferences? Dementia 2009 Apr 22;8(2):223-243. [doi: 10.1177/1471301209103259]
45. Whitlatch CJ, Feinberg LF, Tucke SS. Measuring the values and preferences for everyday care of persons with cognitive impairment and their family caregivers. Gerontologist 2005 Jun;45(3):370-380. [doi: 10.1093/geront/45.3.370] [Medline: 15933277]
46. Kozlov E, McDarby M, Duberstein P, Carpenter BD. The feasibility and acceptability of an intergenerational, web-based intervention to enhance later-life family care planning. Gerontologist 2021 Sep 13;61(7):1153-1163. [doi: 10.1093/geront/gnaa137] [Medline: 32946554]
47. Orsulic-Jeras S, Whitlatch CJ, Szabo SM, Shelton EG, Johnson J. The SHARE program for dementia: Implementation of an early-stage dyadic care-planning intervention. Dementia (London) 2019 Jan 13;18(1):360-379. [doi: 10.1177/1471301216673455] [Medline: 27738110]
48. Øksnebjerg L, Woods B, Waldemar G. Designing the ReACT app to support self-management of people with dementia: An iterative user-involving process. Gerontology 2019;65(6):673-685 [FREE Full text] [doi: 10.1159/000500445] [Medline: 31266031]

XSL•FO
RenderX

49. Choi SK, Yelton B, Ezeanya VK, Kannaley K, Friedman DB. Review of the content and quality of mobile applications about Alzheimer's disease and related dementias. J Appl Gerontol 2020 Jun;39(6):601-608 [FREE Full text] [doi: 10.1177/0733464818790187] [Medline: 30049239]
50. Werner NE, Brown JC, Loganathar P, Holden RJ. Quality of mobile apps for care partners of people with Alzheimer disease and related dementias: Mobile app rating scale evaluation. JMIR Mhealth Uhealth 2022 Mar 29;10(3):e33863 [FREE Full text] [doi: 10.2196/33863] [Medline: 35348467]
51. Ye B, How T, Chu CH, Mihailidis A. Dementia care apps for people with dementia and informal caregivers: A systematic review protocol. Gerontology 2021;67(5):633-638 [FREE Full text] [doi: 10.1159/000514838] [Medline: 33774646]
52. Brankaert R, Kenning G. HCI and Design in the Context of Dementia. Cham: Springer; 2020.
53. Pearlman R, Starks H, Cain K, Cole W, Rosengren D, Patrick D. Your Life Your Choices. Planning for Future Medical Decisions: How to Prepare a Personalized Living Will. Elder Guru. URL: https://www.elderguru.com/downloads/your_life_your_choices_advance_directives.pdf [accessed 2022-07-18]
54. Sudore RL, Boscardin J, Feuz MA, McMahan RD, Katen MT, Barnes DE. Effect of the PREPARE website vs an easy-to-read advance directive on advance care planning documentation and engagement among veterans: A randomized clinical trial. JAMA Intern Med 2017 Aug 01;177(8):1102-1109 [FREE Full text] [doi: 10.1001/jamainternmed.2017.1607] [Medline: 28520838]
55. Sudore RL, Fried TR. Redefining the "planning" in advance care planning: preparing for end-of-life decision making. Ann Intern Med 2010 Aug 17;153(4):256-261 [FREE Full text] [doi: 10.7326/0003-4819-153-4-201008170-00008] [Medline: 20713793]
56. Berridge C, Zhou Y, Lazar A, Porwal A, Mattek N, Gothard S, et al. Control Matters in Elder Care Technology: Evidence and Direction for Designing It In. In: DIS '22: Designing Interactive Systems Conference. 2022 Presented at: Designing Interactive Systems Conference; June 13-17, 2022; Virtual Event Australia. [doi: 10.1145/3532106.3533471]
57. Sebern MD, Whitlatch CJ. Dyadic relationship scale: a measure of the impact of the provision and receipt of family care. Gerontologist 2007 Dec;47(6):741-751. [doi: 10.1093/geront/47.6.741] [Medline: 18192628]
58. Menne HL, Tucke SS, Whitlatch CJ, Feinberg LF. Decision-making involvement scale for individuals with dementia and family caregivers. Am J Alzheimers Dis Other Demen 2008;23(1):23-29 [FREE Full text] [doi: 10.1177/1533317507308312] [Medline: 18276955]
59. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
60. Stetz K. The Experience of Spouse Caregiving for Persons with Advanced Cancer. elibrary. Seattle, WA: University of Washington; 1986. URL: https://www.elibrary.ru/item.asp?id=7507678 [accessed 2022-07-18]
61. Wallhagen M. Perceived control and adaptation in elderly caregivers. Sigma Repository. Seattle, WA: University of Washington; 1988. URL: https://sigma.nursingrepository.org/handle/10755/18350 [accessed 2022-07-18]
62. Deterding N, Waters M. Flexible coding of in-depth interviews: A twenty-first-century approach. Sociological Methods & Research 2018 Oct 01;50(2):708-739. [doi: 10.1177/0049124118799377]
63. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis. International Journal of Qualitative Methods 2017 Oct 02;16(1):160940691773384. [doi: 10.1177/1609406917733847]
64. Irish M, Piolino P. Impaired capacity for prospection in the dementias--Theoretical and clinical implications. Br J Clin Psychol 2016 Mar;55(1):49-68. [doi: 10.1111/bjc.12090] [Medline: 26014112]
65. Berridge C, Turner A, Zaslavsky O, Parsey C. Design for dementia. Interactions 2022 Mar;29(2):46-49. [doi: 10.1145/3512346]

Abbreviations

AD: Alzheimer disease
ADRC: Alzheimer’s Disease Research Center
ADRD: Alzheimer disease–related dementias
UW: University of Washington
WCAG: Web Content Accessibility Guidelines

XSL•FO
RenderX

Edited by T Leung; submitted 06.05.22; peer-reviewed by A Garcia; comments to author 03.06.22; revised version received 17.06.22; accepted 28.06.22; published 27.07.22.

Please cite as:
Berridge C, Turner NR, Liu L, Karras SW, Chen A, Fredriksen-Goldsen K, Demiris G
Advance Planning for Technology Use in Dementia Care: Development, Design, and Feasibility of a Novel Self-administered Decision-Making Tool
JMIR Aging 2022;5(3):e39335
URL: https://aging.jmir.org/2022/3/e39335
doi:10.2196/39335
PMID:35896014

©Clara Berridge, Natalie R Turner, Liu Liu, Sierramatice W Karras, Amy Chen, Karen Fredriksen-Goldsen, George Demiris. Originally published in JMIR Aging (https://aging.jmir.org), 27.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Web-Based Platform (CareVirtue) to Support Caregivers of People Living With Alzheimer Disease and Related Dementias: Mixed Methods Feasibility Study

Justin J Boutilier[1], PhD; Priya Loganathar[1], MS; Anna Linden[1], MSc; Eleanore Scheer[2], BSc; Sofia Noejovich[1], BSc; Christian Elliott[3], BSc; Matthew Zuraw[3], MBA; Nicole E Werner[4], PhD

[1]Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, United States

[2]Department of Engineering Systems and Environment, University of Virginia, Charlottesville, VA, United States

[3]Whiplash Technology, Inc, Palm Springs, CA, United States

[4]Department of Health and Wellness Design, Indiana University School of Public Health-Bloomington, Bloomington, IN, United States

Corresponding Author:
Nicole E Werner, PhD
Department of Health and Wellness Design
Indiana University School of Public Health-Bloomington
1025 E 7th St
Bloomington, IN, 47405
United States
Phone: 1 703 340 7177
Email: newerner@iu.edu

Related Article:

This is a corrected version. See correction statement: https://aging.jmir.org/2022/3/e41912

Abstract

Background: People living with Alzheimer disease and related dementias (ADRD) require prolonged and complex care that is primarily managed by informal caregivers who face significant unmet needs regarding support for communicating and coordinating across their informal care network. To address this unmet need, we developed CareVirtue, which provides (1) the ability to invite care network members; (2) a care guide detailing the care plan; (3) a journal where care network members can document, communicate, and coordinate; (4) a shared calendar; and (5) vetted geolocated caregiver resources.

Objective: This study aims to evaluate CareVirtue's feasibility based on: (1) Who used CareVirtue? (2) How did caregivers use CareVirtue? (3) How did caregivers perceive the acceptability of CareVirtue? (4) What factors were associated with CareVirtue use?

Methods: We conducted a feasibility study with 51 care networks over a period of 8 weeks and used a mixed methods approach that included both quantitative CareVirtue usage data and semistructured interviews.

Results: Care networks ranged from 1 to 8 members. Primary caregivers were predominantly female (38/51, 75%), White (44/51, 86%), married (37/51, 73%), college educated (36/51, 71%), and were, on average, 60.3 (SD 9.8) years of age, with 18% (9/51) living in a rural area. CareVirtue usage varied along 2 axes (total usage and type of usage), with heterogeneity in how the most engaged care networks interacted with CareVirtue. Interviews identified a range of ways CareVirtue was useful, including practically, organizationally, and emotionally. On the Behavioral Intention Scale, 72% (26/36) of primary caregivers reported an average score of at least 3, indicating an above average intention to use. The average was 81.8 (SD 12.8) for the System Usability Scale score, indicating "good" usability, and 3.4 (SD 1.0) for perceived usefulness, suggesting above average usefulness. The average confidence score increased significantly over the study duration from 7.8 in week 2 to 8.9 in week 7 (P=.005; r=0.91, 95% CI 0.84-0.95). The following sociodemographic characteristics were associated with posting in the journal: retired (mean 59.5 posts for retired caregivers and mean 16.9 for nonretired caregivers), income (mean 13 posts for those reporting >US $100K and mean 55.4 for those reporting <US $100K), relationship to care recipient (mean 18.7 posts for child and mean 56.4 for partners/spouses), and living situation (mean 44.7 for those who live with the care recipient and mean 13.1 for those who do not). Older care recipients were associated with fewer posts (r=–0.33, 95% CI –0.55 to –0.06).

XSL•FO
RenderX

Conclusions: This study establishes the acceptability and feasibility of CareVirtue among ADRD care networks and highlights the importance of designing flexible, multicomponent interventions that allow care networks to tailor their engagement according to their needs. The results will be used to improve CareVirtue feasibility and acceptability in preparation for a subsequent randomized trial to test CareVirtue's effectiveness in improving caregiver outcomes.

(JMIR Aging 2022;5(3):e36975) doi:10.2196/36975

KEYWORDS

Alzheimer disease and related dementias; mHealth; caregivers; dementia caregiving; eHealth; telehealth

Introduction

Background

More than 6 million individuals in the United States are living with Alzheimer disease and related dementias (ADRD) and it is attributed as the cause of death for 1 in 3 individuals over the age of 65 [1]. In the past 2 years, deaths attributed to ADRD have increased by 16% and research has projected that the number of people living in the United States with ADRD will triple by 2060 [1].

Individuals living with ADRD require prolonged and complex care that is primarily managed by informal caregivers. Informal caregivers are unpaid, nonprofessionals that provide care and typically include family and friends. There are an estimated 11 million caregivers providing care for people living with ADRD in the United States and they provide approximately 15.3 billion hours of unpaid care valued at nearly US $257 billion [2]. Caregivers report being undertrained, under-supported, and under-resourced to perform their caregiving role. Although caregivers can experience positive outcomes related to caregiving, the imbalance of caregiving demands and supports is often associated with mental, physical, and economic challenges that can lead to significant consequences for caregivers and the individual living with ADRD, such as caregiver stress, burden, depression, and morbidity [3-5].

To address these suboptimal caregiver outcomes, the US National Institute on Aging and other national advisory panels have highlighted the development and testing of technology-based interventions for caregivers of people living with ADRD as a key priority [6-8]. For example, the 2015 Alzheimer's Research Summit highlighted the need to "test the use of technology to overcome the workforce limitations in the care of older adults with dementia as well as providing caregiver support and education." [7]. In response, researchers have developed numerous information technology interventions such as mobile apps and websites to support ADRD caregivers across a range of domains including caregiver education, self-care support, support for managing behavioral symptoms of dementia, and virtual peer support groups [9-11]. Several systematic reviews and recent meta-analyses report that these technology interventions can improve outcomes for caregivers, such as increased self-efficacy and reduced ADRD caregiver burden, stress, and depression [3,12-19]. These reviews also suggest that effective interventions offer multiple components, tailored options, and social support [9-11].

However, a significant gap in existing interventions is that most focus only on the primary caregiver, even though most people living with ADRD receive care from more than 1 caregiver—*a care network*—with varying degrees of involvement [20-24]. Currently, caregivers face significant unmet needs regarding support for communicating and coordinating across the care network including sharing information, maintaining situation awareness, distributing responsibilities, scheduling, and managing caregiver hand offs [20-24].

Although some mobile apps exist that allow caregivers to share information, they are limited in their functionality, quality, and potential to meet the specific needs of ADRD caregivers [25]. A recent review of mobile apps for caregivers of people living with ADRD available on the US market identified 2 mobile apps that support shared communication and coordination [25]. According to study findings, one of those apps did not function consistently and received a quality rating of inadequate as indicated by the Mobile Application Rating Scale (MARS). The second app received an overall quality rating of minimally acceptable quality according to the MARS but scored lower than average on subjective quality. A similar study conducted in 2018 by Wozney et al [26] identified 3 mobile apps that connect a primary caregiver to other members of the care network. One of the apps identified is no longer available on the US market. The other 2 apps identified are not specifically designed to meet the needs of caregivers for people living with ADRD and are limited in their function (eg, only provide a shared calendar).

To address the current gaps in existing interventions, we developed CareVirtue, a progressive web application developed in React to support and connect ADRD care networks that can be accessed via a web browser on any device with a data connection. CareVirtue seeks to address the current gaps in existing ADRD caregiver support technologies through a high-quality, user-centered, ADRD caregiver–specific multicomponent technology to support communication, coordination, and connection among care networks. CareVirtue was initially inspired by an online support group for people newly diagnosed with ADRD and their caregivers in which support group members expressed an unmet need for tools to support communication and coordination among the care network. The need for CareVirtue was further supported by several findings from foundational research on care networks. First, findings from Block et al [27] suggested that primary caregivers require technologies to communicate and coordinate among the care network, that they try to adapt existing technologies (eg, email, messaging) to meet their needs, and that adaptation requires additional time and effort. Further, Ponnala et al [20] found that for primary caregivers, the currently under-supported communication and coordination

XSL•FO
RenderX

among the care network increases their caregiving demands. Moreover, Tang et al [24] highlighted the consequences caregivers experience with under-supported care network communication and coordination, including maintaining situation awareness among caregivers, missing care information leading to potential patient harm (eg, missing a medication dose), and miscommunication leading to care network tensions or conflict.

Collectively, this prior research provides the foundation for CareVirtue. CareVirtue's design honors the person-centered care model for people living with ADRD and their caregivers by (1) treating people living with ADRD as individuals with unique needs; (2) seeing the world from their perspective; and (3) creating a positive social environment in which the person living with ADRD and caregiver can experience relative well-being and quality of life. CareVirtue was designed and developed through consistent, iterative user input across multiple stages of usability testing coupled with expert evaluation. CareVirtue was specifically designed to encapsulate the foundational principles of person-centered care through the following features. Multimedia Appendix 1 provides a detailed walkthrough of CareVirtue features, which are given in brief in Textbox 1.

Textbox 1. CareVirtue features.

CareVirtue Dashboard

Acts as a centralized hub to document and share important information with the care team. The dashboard includes a view presenting upcoming care appointments and events, linked to the care calendar; a list of current and pending care team members; and a journal where care network members can document, communicate, and coordinate about daily care events (Figure 1).

Journal Reports

Search and filter options to explore trends and gain insights into care recipient needs. The care journal or portions of the care journal can be exported to PDF to share as necessary.

Care Guide Template

Includes a table of abilities related to specific activities of daily living and instrumental activities of daily living, and sections for needs and preferences, with a focus on quality-of-life details to help any caregiver understand the care recipient as a whole person (Figure 2). See Multimedia Appendix 2 for a detailed version of the care guide template.

Care Team Management

The ability to invite care network members to use the account with the primary caregiver with security permissions assigned at each invitation (Figure 3).

Shared Calendar

Supports scheduling and sharing recurring care events, reminders, and appointments (Figure 3).

Geolocated Resources List

For the current study, resources were limited to the Alzheimer's Association 24×7 helpline, contact details for CareVirtue support, and contact details for the research team. The subsequent version of CareVirtue will include caregiver and person living with ADRD resources specific to their specific location such as the local area agency on aging (Figure 4).

XSL•FO
RenderX

Figure 1. The CareVirtue Dashboard, a centralized hub to document and share important information with the care team.

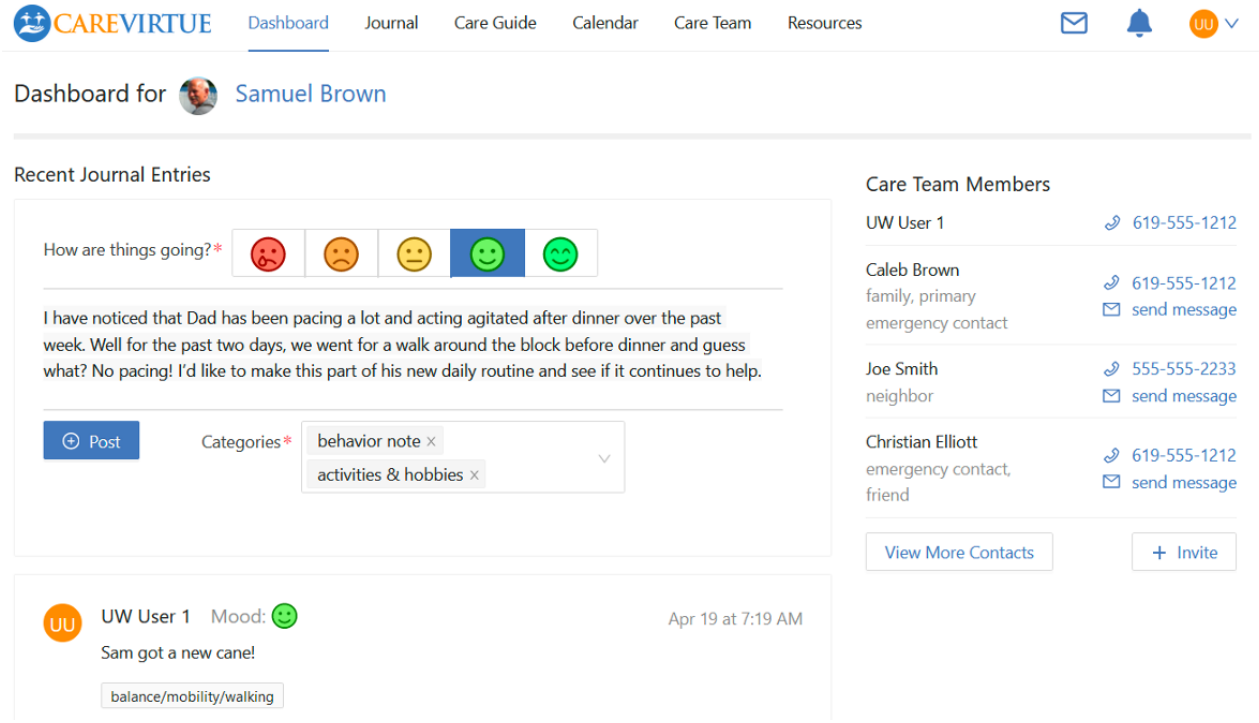
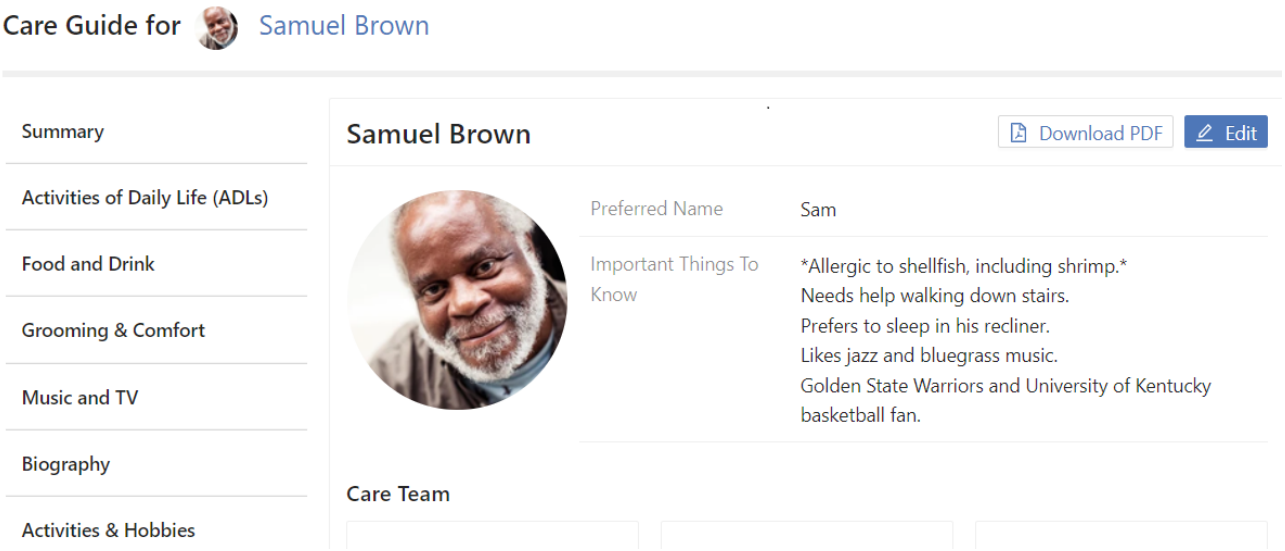


Figure 2. The CareVirtue Care Guide.



XSL•FO
RenderX

Figure 3. The CareVirtue Care Team management feature and form to create a new calendar event.

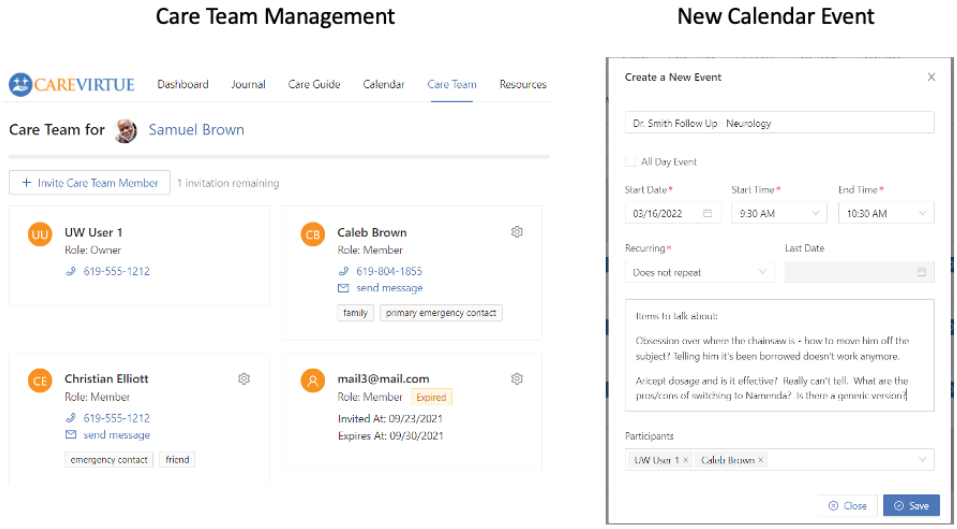
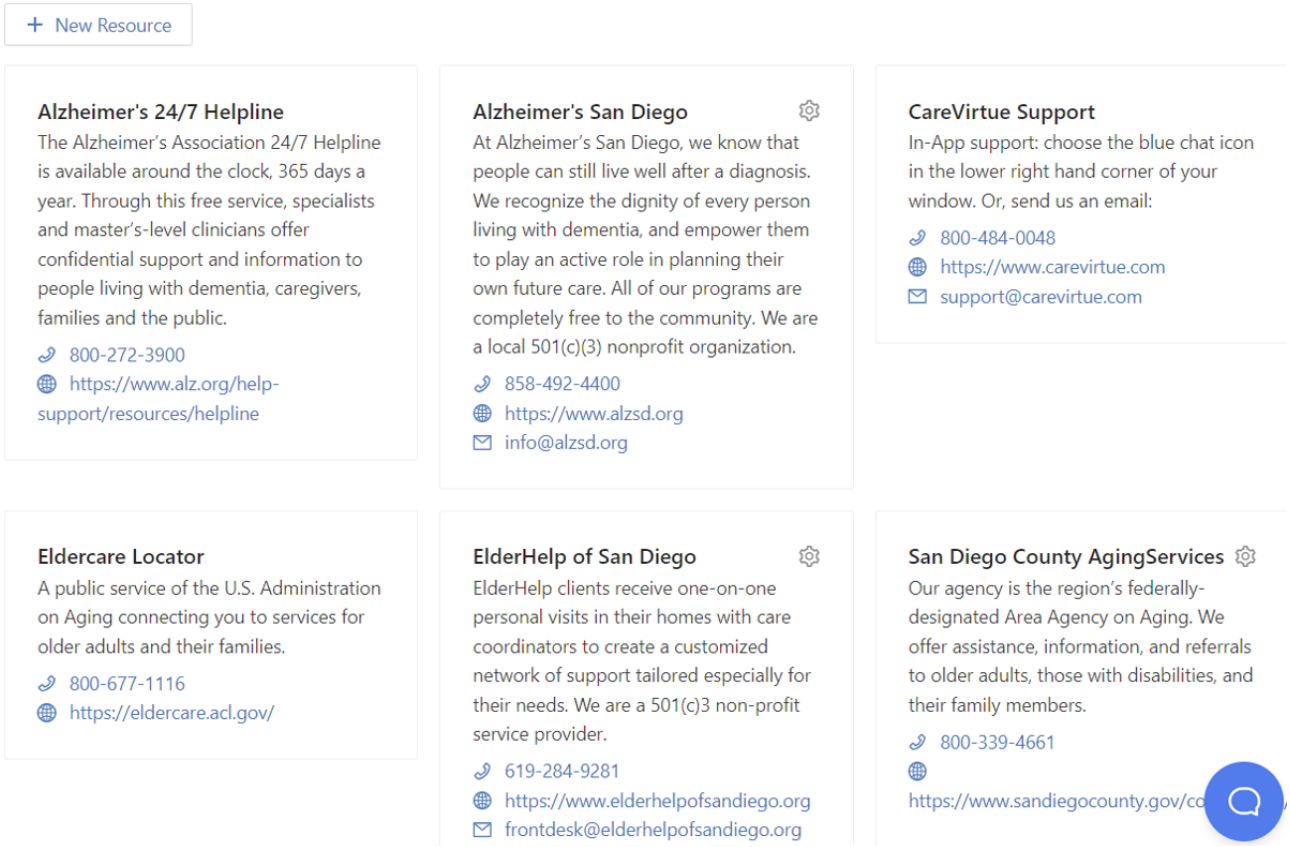


Figure 4. Example of the CareVirtue resources list for a caregiver living in San Diego, California.



Objectives

In this study, we used a mixed methods approach to evaluate CareVirtue's feasibility across the following research questions:

- Who used CareVirtue?
- How did caregivers use CareVirtue?
- How did caregivers perceive the acceptability of CareVirtue?
- What factors were associated with CareVirtue use?

Methods

Design

We conducted a feasibility study over a period of 8 weeks with the purpose of demonstrating CareVirtue's acceptability and feasibility among care networks of people living with ADRD. This study reports one aim of a larger project, which has 2 specific aims. The first aim is the focus of this study. The second aim is to leverage the CareVirtue data generated by this feasibility study to develop an intelligent caregiver assistant (R41AG069607). The larger sample size and longer study

XSL•FO
RenderX

duration than is typical for feasibility studies is due to adjustments made to achieve the second goal [28].

Setting and Sample

Participants were recruited between February and June 2021 through multiple community sites in Wisconsin and Southern California as well as through the Wisconsin Alzheimer's Disease Research Center. Advertisements for study participation were distributed via email, social media, and newsletter posts. Interested individuals contacted the study team via email or phone and were subsequently phone screened for the following eligibility criteria: self-identified primary caregiver of a person living with ADRD, at least 18 years of age, English speaking, daily internet access, and shares caregiving information/responsibility with other caregivers.

Procedures

Eligible participants were scheduled for a 1-hour enrollment visit via videoconferencing software. During the enrollment visit, a study team member obtained informed consent from the primary caregiver and from the associated person living with ADRD (ie, the care recipient). If the person living with ADRD did not have decisional capacity to consent, the primary caregiver could consent on his/her behalf if he/she was the legally authorized representative. Next, the study team member administered a pretrial demographic survey, helped create the CareVirtue account, and provided a walk-through of CareVirtue's functionality. Primary caregivers selected and invited secondary caregivers (to form a care network) at their own discretion. Once secondary caregivers were invited, they were separately contacted via email to electronically obtain informed consent.

Following enrollment, participants used CareVirtue for 8 weeks. During the use period, we administered a weekly survey starting 1 week after enrollment to assess caregiver workload and confidence using CareVirtue. If the survey was not completed, a follow-up reminder was sent the following day. At the completion of 8 weeks of use, we conducted a posttrial visit with the primary caregiver participant via Zoom where we administered posttrial surveys and conducted a semistructured interview as described in the following section. Primary caregivers were provided with a US $150 e-gift card at the end of the study period. Secondary caregivers contacted the study team if they were interested in participating in the postuse survey and interview. Secondary caregivers received US $15 for completion of the postuse survey and US $25 for completion of the 30-minute postuse interview.

Data Sources

Demographic Data

We collected primary caregiver characteristics including age, gender, race and ethnicity, income, education, marital status, location, and employment. We also collected demographics for the care receiver including age, gender, ethnicity, living situation, and relationship to the primary caregiver (Table 1). Demographic data were not collected from the secondary caregivers.

CareVirtue Use

We collected data on 8 CareVirtue platform usage metrics across the study period: number of log-ins, journal posts, journal post replies, calendar events, secondary caregiver invites sent, secondary caregiver invites accepted, care guide sections created, and resources accessed. The log-in data are not fully representative of actual use because users could remain logged in to CareVirtue depending on their preference to log out.

Acceptability Measures

Quantitative Acceptability Measures

To assess acceptability, we measured primary caregiver perceptions of usability and usefulness. We measured usability using a weekly confidence survey with a single question: "rate your confidence in using CareVirtue on a scale from 1 (not at all confident) to 10 (very confident)" and using the System Usability Scale (SUS), which includes 10 statements (eg, "Learning to use CareVirtue was quick for me") and a 5-point response scale (1=strongly disagree to 5=strongly agree) [29]. We measured usefulness with 3 surveys. First, we used the National Aeronautics and Space Administration-Task Load Index (NASA-TLX) to assess caregiver workload on a 6-item subscale with a 100-point range (1=very low to 5=very high). The purpose of this measure was to understand the relationship between caregiver workload and CareVirtue use. Second, we used the Behavioral Intention Scale, which included 4 statements (eg, "If it were up to you, to what extent would you want to use CareVirtue?") and 5-point response scale from 1 (not at all) to 5 (a great deal) [30,31]. Third, we used the perceived usefulness survey, which includes 4 statements (eg, "Using CareVirtue would make it easier to perform my caregiving role") and a 5-point response scale (1=strongly disagree to 5=strongly agree) [32]. Quantitative acceptability measures were collected from all primary caregiver participants. Secondary caregivers could opt in to complete these assessments by contacting the study team.

Qualitative Interviews

To provide context to the quantitative measures of acceptability, we conducted semistructured interviews with the primary caregivers, which focused on caregivers' experiences with CareVirtue during the study period. The interview guide was developed with input from the full research team (Multimedia Appendix 3).

Analysis Plan

Overview

We used Python 3.8 (Python Software Foundation) to compute descriptive statistics and conduct statistical analyses for all quantitative data. Qualitative data were coded using Microsoft Excel. Analyses related to research questions are described in detail below.

Who Used CareVirtue?

To determine participant characteristics, we computed descriptive statistics from demographic survey responses.

XSL•FO
RenderX

How Did Caregivers Use CareVirtue?

To assess usage heterogeneity, we computed descriptive statistics for each of the 8 usage characteristics. We then used k-means clustering to cluster care networks into "user types," where each care network is represented by an 8-dimensional usage vector. The number of clusters, k, was varied from 1 to 20 and the elbow method was used to select the final k value: 8. See Multimedia Appendix 4 for more details.

How Did Caregivers Perceive CareVirtue Acceptability?

To assess perceptions of usability and usefulness, we computed descriptive statistics for the SUS, Behavioral Intention Scale, and Perceived Usefulness Scale. We also computed Pearson correlation coefficient for the weekly NASA-TLX to assess the change in overall caregiver workload over the study period and for the confidence survey to assess if confidence changed over the study period.

To further explore perceptions of usability and usefulness, we conducted a general content analysis of the interview transcripts.[53] Three members of the research team (PL, SN, and AL) with training in human factors engineering reviewed all transcripts and identified initial categories related to CareVirtue usability and usefulness, with 2 team members coding each transcript. Coders met weekly to discuss codes and resolve discrepancies, which were also discussed in a biweekly meeting with a senior research team member (NEW) with expertise in qualitative research and human factors engineering. The codebook was refined iteratively across the team-based discussions and the final codebook was applied across all transcripts using a team-based consensus process [33].

What Factors Were Associated With CareVirtue Use?

To explore the factors associated with CareVirtue use, we conducted a series of univariate analyses to assess the correlation between each of the 8 usage characteristics and each of the 14 variables from the demographic survey: the NASA-TLX score from each week, the confidence score from each week, the SUS score, the average behavioral intention score, and the average perceived usefulness score. For continuous variables, we computed Pearson correlation coefficient and the corresponding 95% CI to assess if the correlation was statistically significant (P=.05). For discrete variables, we first converted them into binary variables (if not already) by merging classes to ensure suitable sample sizes. Then, we used an unpaired t test to assess whether the difference between the average from each class was statistically significant. We were unable to perform multivariate analyses due to our limited sample size.

Ethics Approval

Research ethics approval was granted by the Institutional Review Board at the University of Wisconsin-Madison (Protocol #2020-1035).

Results

Who Used CareVirtue?

We enrolled 51 primary caregivers of people living with ADRD (Table 1) and 61 secondary caregivers to use CareVirtue during the study period. Care networks ranged from 1 to 8 members. Primary caregivers were predominantly female (38/51, 75%), White (44/51, 86%), married (37/51, 73%), college educated (36/51, 71%), and were, on average, 60.3 (SD 9.8) years of age. Care recipients were also primary female (34/51, 67%) and White (45/51, 88%), with an average age of 79.2 (SD 10.6). Care networks were located in both Wisconsin (29/51, 57%) and California (19/51, 37%), with 18% (9/51) living in a rural area. During the study period, 4 primary caregivers dropped out because of care recipient death (n=2) and personal situations (n=2). We were unable to reach 6 primary caregiver participants for the posttrial visit. A total of 12 secondary caregiver participants completed the postuse survey.

XSL•FO
RenderX

Table 1. Summary of primary caregiver and care recipient characteristics.

Characteristic	Primary caregivers (n=51)	Care recipients (n=51)
Female gender, n (%)	38 (75)	34 (67)
Age (years), mean (SD)	60.3 (9.8)	79.2 (10.6)
Race/ethnicity, n (%)		
Asian	2 (4)	2 (4)
Black or African American	1 (2)	1 (2)
Hispanic or Latinx	2 (4)	2 (4)
Native American or American Indian	1 (2)	0 (0)
Not reported	1 (2)	1 (2)
White	44 (86)	45 (88)
Marital status, n (%)		N/A[a]
Married or domestic partnership	37 (73)	
Divorced	11 (22)	
Single, never married	2 (4)	
Widowed	1 (2)	
Education, n (%)		N/A
Postcollege	19 (37)	
4-year college	17 (33)	
Technical school, vocational training, community college	10 (20)	
High school diploma or equivalent	5 (10)	
Employment, n (%)		N/A
Full-time	21 (41)	
Retired	19 (37)	
Part-time	7 (14)	
Not working	4 (8)	
Income, n (%)		N/A
≤US $19,000	1 (2)	
US $20,000-39,000	2 (4)	
US $40,000-59,000	8 (16)	
US $60,000-79,000	4 (8)	
US $80,000-99,000	6 (12)	
≥US $100,000	18 (35)	
Do not wish to answer	8 (16)	
Location, n (%)		N/A
Wisconsin	29 (57)	
California	19 (37)	
Illinois	2 (4)	
Virginia	1 (2)	
Location type, n (%)		N/A
Urban	42 (82)	
Rural	9 (18)	
Relationship to caregiver, n (%)	N/A	

XSL•FO
RenderX

Characteristic	Primary caregivers (n=51)	Care recipients (n=51)
Parent		28 (55)
Spouse/Partner		20 (39)
Other relative		3 (6)
Distance to caregiver, n (%)	N/A	
In household		34 (67)
<20 minutes		12 (24)
20-60 minutes		2 (4)
>2 hours		3 (6)
Living situation, n (%)	N/A	
In a house		40 (78)
In a nursing home, retirement community, or other assisted living facility		9 (18)

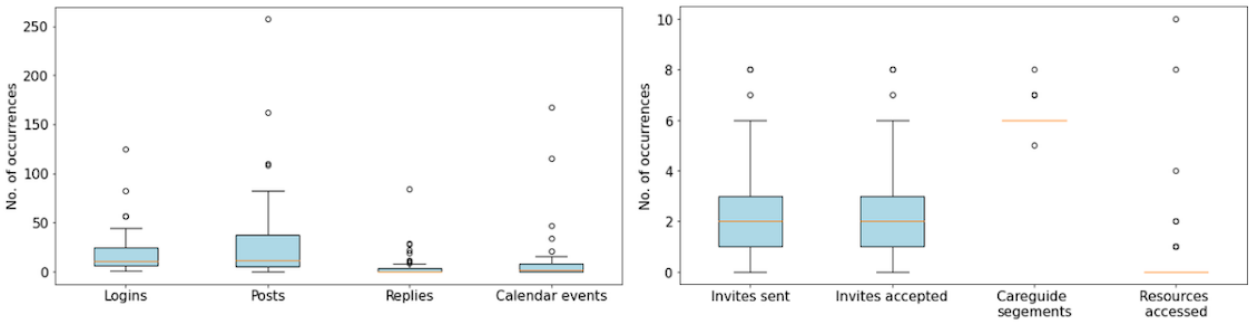
[a]N/A: not applicable.

How Did Caregivers Use CareVirtue?

Figure 5 displays boxplots (across care networks) for each of the 8 usage characteristics. The average (SD) was 18.3 (22.4) log-ins, 32.5 (46.5) journal posts, 5.3 (13.2) journal post replies, 10.6 (28.5) calendar events, 2.2 (2.1) secondary caregiver invites sent, 2.2 (2.1) secondary caregiver invites accepted, 6.1 (0.4) care guide sections created, and 0.6 (1.9) resources accessed. The log-in data are not fully representative of actual use because users could remain logged in to CareVirtue depending on their preference to log out.

Table 2 presents the centroid (the mean values across all care networks in that cluster), cluster size, and a cluster label for each of the 8 clusters. The 8 clusters could be further reduced into 3 primary groups according to the degree of engagement with the platform. There was heterogeneity in how the most engaged care networks interacted with the platform; for example, 2 care networks made heavy use of the calendar feature (average of 141 events) with few posts (average of 41), while another care network made heavy use of the journal (257 posts) with only 1 calendar event.

Figure 5. Box plots for each of the eight usage characteristics. They are separated into two plots due to differences in scale.



XSL•FO
RenderX

Table 2. Cluster centroids for the 8 usage clusters identified by k-means. Each centroid component (eg, invites sent) represents the average across all care networks within that cluster.

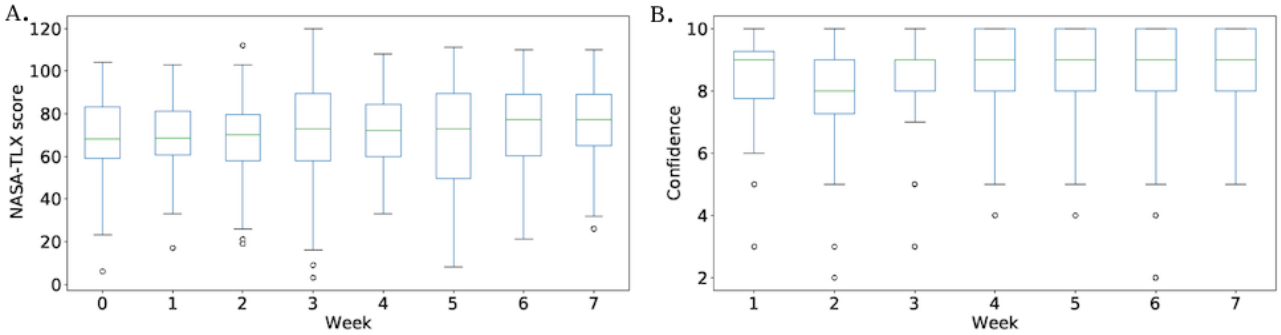
Cluster description	Cluster size	Cluster centroid							
		Log-ins	Posts	Replies	Invites sent	Invites accepted	Calendar events	Care guide sections	Resources accessed
Low engagement group									
	23	10	13	1	1	1	3	6	0
Moderate engagement group									
Small care networks	12	13	25	2	3	2	6	6	0
Large care networks	7	19	40	11	6	6	3	6	0
High engagement group									
Log-in heavy	3	88	56	9	3	3	18	6	0
Balanced usage	2	30	53	14	4	4	18	6	9
Calendar focused	2	33	41	0	2	2	141	6	0
Posts and replies	1	15	162	84	3	3	16	8	2
Posts only	1	12	257	1	2	2	1	6	0

How Did Caregivers Perceive CareVirtue Acceptability?

We used the NASA TLX score to assess usability and usefulness in terms of caregivers' perceptions of their caregiver workload. Figure 6A displays boxplots of the total NASA TLX score for each week from week 0 (before the study began) to 7 (the final week of the study). The average NASA TLX score increased over the duration of the study (P=.02; r=0.79, 95% CI 0.65-0.87). However, at an individual level the NASA TLX score only increased over the duration of the study for 3 primary caregivers, decreased for 2 caregivers, and did not change for the remaining 29 primary caregivers (17 were excluded due to missingness).

We used a confidence scale to assess CareVirtue usability. Figure 6B displays boxplots of the total confidence score for each week of the study. The average confidence score increased significantly over the duration of the study from a low of 7.8 in week 2 to a high of 8.9 in week 7 (P=.004; r=0.91, 95% CI 0.84-0.95). At an individual level, 7 primary caregivers saw a statistically significant increase (P<.05 in all cases; see Multimedia Appendix 5 for precise P values) in confidence, while the remaining 17 remained stable (27 were excluded due to missingness).

Figure 6. (A) Box plot for the weekly NASA TLX score. (B) Box plot for the weekly confidence (in using CareVirtue) survey.



We also used the SUS to assess CareVirtue usability and the Behavioral Intention Scale and Perceived Usefulness Scale to assess usefulness. Multimedia Appendix 6 displays histograms for the Behavioral Intention Scale score, SUS score, and the perceived usefulness score for primary caregivers. The average (SD) was 3.3 (1.2) for the Behavioral Intention Scale and 72% (26/36) of primary caregivers reported an average score of at least 3, indicating an above average intention to use. The average (SD) was 81.8 (12.8) for the SUS score, indicating "good" usability and 3.4 (1.0) for perceived usefulness, suggesting above average perceived usefulness. For secondary caregivers the average (SD) was 2.9 (1.1) for the Behavioral Intention Scale (6/11, 55%, had an average score of at least 3), 77.9 (13.7) for the SUS, and 3.4 (0.8) for perceived usefulness (Multimedia Appendix 7).

XSL•FO
RenderX

Through analysis of the qualitative interviews, we identified 10 categories related to usefulness and 4 categories related to opportunities to improve usefulness (Table 3). Participants described CareVirtue as facilitating connection, exploration, and awareness. In addition, CareVirtue allowed for documentation and tracking of daily experiences and enabled emotional catharsis by facilitating the capture and review of significant moments in their relationship with the person living with ADRD. Additional facilitators of usefulness described by participants included centralization of information, coordination across the care network, and the privacy afforded by controlling user permissions. Participants further explained that using CareVirtue made them feel supported and reduced feelings of being overwhelmed. We also identified opportunities for improvement, which included increasing engagement such as creating a CareVirtue user support group; adding customization such as additional emoji options for journal posts; refining navigation such as expanding search types; and additional functionality such as document upload and storage.

XSL•FO
RenderX

Table 3. Categories of usefulness and opportunities for improvement with illustrative quotations.

Categories and subcategories	Description	Illustrative quotation
Usefulness		
Connection	Interaction across the care network	*The biggest thing was it [CareVirtue] allowed me to bring my brother and sister, both of whom live a thousand or more miles away, closer into the loop...at crucial times...there, you know, were some things going on, and they were actually getting the email reminders, and they were actually logging in and looking at my journal comments and responding in the journal.* [P45]
Documentation and Tracking	View, explore, and understand trends over time	*It helped me in terms of measuring my spouse's progression over those two whole months. Because, you know, when you're with someone every day, 24/7 almost, you may not notice the differences other people do. So, you know, I found that useful to go back and look at what I wrote, you know, a month ago. Because it appears the stage of the disease is, and...that it progresses could go, slow down and could speed up. So, it seemed like the progression was increasing and just help me quantify it to some degree in terms of what her capabilities were.* [P50]
Emotional Catharsis	Capture, share, and recollect experiences and important moments in the relationship with the person living with ADRD	*But the other day he said, you know, I love you. You're my favorite. And he hadn't said I love you to me in like, I don't know, a few years...but to document a moment like that as a way for me to kind of cathartically capture those moments and have those to look back on.* [P13]
Awareness	Real-time understanding of daily care experiences and status of the caregiver and person living with ADRD	*[CareVirtue] also allowed me to let my kids know what was going on the same time it was happening, as far as what, you know, when I was doing the journaling. It also helped my kids to be able to see probably a really good picture of all the different aspects of what their mom is going through.* [P15]
Centralization and Organization	Communication, coordination, documentation, and tracking in one location	*So, with the CareVirtue, having one spot, like if I'm going to communicate, I'm going to put it in there, and then everybody can just go to that spot to look for the information...people can log-in and just have their notifications and know that stuff was going on. And it would be one step there versus me trying to figure out how to get, you know, am I in the right [text message] thread, which information needs to go to who?* [P26]
Coordination	Seamless, high-quality transitions of care	*But the app [CareVirtue] was really helpful, because before a person came in for their quote, unquote, shift time, they can have advance information about how things were before they came, so they could be kind of prepared.* [P32]
Introspection	Self-exploration of feelings, care strategies, and goals	*[CareVirtue] just helped me to identify, I guess, where, how I was feeling and what my plan was for going forward.* [P16]
Privacy	Customizable permissions on a secure platform focused only on care	*[Before CareVirtue] we kind of send stuff through Messenger, which is not necessarily a secure, you know, thing, and this one [CareVirtue] is. So, yeah, it has just created something that was specifically and exclusively for her care, and that, you know, that was good.* [P1]
Reduce Burden	Reduce demands associated with communication, coordination, and documentation	*[CareVirtue] was really useful in terms of keeping a journal and telling everybody what was going on without having to call every single family member, so they could read [in CareVirtue] what was going on.* [P7]
Support	Accessible and responsive customer support	*When I had a question, I just hit the little blue bubble and it sent a note. And in absolutely no time somebody [from CareVirtue support]...answered the question or told me how to do what I was, needed to do...It just made it really helpful, really easy to reach out.* [P23]
Opportunities for improving usefulness		
Engagement	Provide additional interactive content such as private journaling space, a support group across CareVirtue users, and a daily checklist of care activities	*It's very personal. Even though they are family members, it's kind of like do I really want them to read about my inner thoughts about this, you know, because it could frighten them.* [P37]
Customization	Include additional customization options such as for reminders, and including more emoji options	*The [emoji] smiley or the sad or whatever, it makes you really think about...I would like it that you could put a couple options in there though [instead of only one]. Because, you know, I might start the entry out in one way and then key in another because it turned.* [P44]
Navigation	Expand search feature	*A filter to search, you know, for key words, or maybe not even key words, any word search, you know, any text search, I find immensely valuable.* [P12]

XSL•FO
RenderX

Categories and subcategories	Description	Illustrative quotation
Functionality	Add functionality to include allowing for an account with multiple care receivers and document upload and storage	*It still lacks the ability for it to be the all-inclusive kind of filing cabinet that I personally need it to be.* [P26]

What Factors Were Associated With CareVirtue Use?

The following sociodemographic characteristics were associated with the number of posts: retired (average of 59.5 posts for retired caregivers as compared with 16.9 for nonretired caregivers), income (13 posts for those reporting >US $100K and 55.4 for those reporting <US $100K), relationship to care recipient (18.7 posts for child and 56.4 for partners/spouses), and living situation (44.7 for those who live with the care recipient and 13.1 for those who do not). Older care recipients were associated with fewer posts (r=–0.33, 95% CI –0.55 to –0.06).

The following workload characteristics were associated with the number of posts: NASA TLX score representing the perceived workload associated with the caregiving role from weeks 2 to 7 (r=0.37-0.46) and the total hours caregiving (r=0.38, 95% CI 0.11-0.60). In other words, higher perceived mental workload associated with the caregiving role and a greater number of hours spent caregiving were associated with more journal posts.

Regarding usability, we found that a higher SUS score (r=0.40, 95% CI 0.10-0.64) and a higher behavioral intention score (r=0.378, 95% CI 0.11-0.59) were associated with an increased number of posts.

We found that 3 demographic characteristics were associated with the number of log-ins. Retired caregivers had an average of 27.4 log-ins as compared with 11.3 for nonretired caregivers. Both primary caregiver age (r=0.31, 95% CI 0.03-0.54) and the total behavioral intention score (r=0.38, 95% CI 0.11-0.59) were associated with more log-ins. We found that a higher SUS score was associated with an increased number of secondary caregiver invites sent (r=0.33, 95% CI 0.05-0.57).

Discussion

Principal Findings

This study establishes the acceptability and feasibility of CareVirtue use among care networks of individuals living with dementia. The results indicate that CareVirtue was perceived as highly usable and useful, with caregivers indicating a range of ways CareVirtue was useful to them, including practically, organizationally, and emotionally. We found that retirees, spouses/partners of the care recipient, and those who live with the care recipient were more likely to post more frequently in the journal. CareVirtue use was not correlated with caregiver age or education level.

Participants were confident in using CareVirtue, with confidence increasing over time, which aligns with their reported perceptions of CareVirtue as highly useful and usable. The qualitative analysis revealed that CareVirtue is useful across multiple dimensions including reducing burden associated with logistics and organization; providing emotional and social support; and facilitating documentation, tracking, and awareness across the care network.

Interestingly, participants' perceived workload associated with their caregiving role increased over the study period. It is possible that this result was reflective of the burden of participating in the study. However, we also found that higher caregiving workload was associated with more frequent use of the CareVirtue journal. It is also possible that this consistent increase in perceived caregiving workload was related to the increased burden and isolation associated with the COVID-19 pandemic during which our study took place [34,35]. A recent survey study [33] found that caregiver burden was not associated with caregivers' reported intention to adopt a mobile health (mHealth) intervention. Our finding based on the engagement with the technology intervention, that increased caregiver workload was associated with increased journal posts, provides additional insight into factors that may be influencing caregivers' adoption and use of technology interventions. This finding also points to specific components, such as the ability to journal about daily experiences and emotions, that may be more useful during times of higher workload.

We found that usage varied along 2 axes: total usage and type of usage. There was heterogeneity in how the most engaged care networks interacted with the platform; for example, 2 care networks made heavy use of the calendar event feature (average of 141 events) with few posts (average of 41), while another care network made heavy use of the journal (257 posts) with only 1 calendar event. These results confirm the importance of technology interventions that can account for caregiver heterogeneity [36,37]—that caregivers are diverse individuals who have wide-ranging experiences, needs, and contexts. The need for interventions responsive to caregiver heterogeneity has been further supported by the identified importance of multiple component interventions that are tailorable to specific caregiver needs [5,9-11]. Our findings expand on this by providing insights into the acceptability and feasibility of a technology intervention at the care network level. Like studies focused on primary caregiver technology interventions, our results highlight the importance of designing flexible, multicomponent interventions that allow care networks to tailor their engagement according to their needs and what is most helpful to them. Importantly, this could also allow for tailoring over time, in which care networks can shift engagement as their needs change across disease progression.

This study used a sample size that is much larger than is typical for a feasibility study [38]. Doing so gave us the opportunity to leverage data analytics to provide insight into how care networks engaged with the platform and provided depth to our understanding of what components caregivers find useful [28]. Survey studies have been useful in capturing self-reported

XSL•FO
RenderX

perceptions related to adoption and intention to use [39,40]. Findings such as ours along with others such as Øksnebjerg et al [41] can complement and expand upon findings from self-report studies by exploring caregivers' engagement behaviors. Given caregiver and care network heterogeneity along with the evidence demonstrating the importance of flexible and tailorable multicomponent interventions [3,12-19], there is a need for future research to continue to explore caregiver and care network engagement behavior with technology interventions to provide additional insight and begin to build an evidence base regarding how to optimally tailor interventions and support engagement according to individual caregiver and care network needs. To do this, future feasibility trials could strive to engage larger sample sizes, enabling the use of artificial intelligence and machine learning to increase personalization. Further, efficacy/effectiveness trials that typically enroll larger samples could be used to explore engagement. Doing so may allow for an increased understanding of the relationship between engagement with the technology intervention and the health outcomes.

Although the purpose of this study was not to determine the effectiveness of CareVirtue on caregiver outcomes, previous research has demonstrated the potential of technology interventions to improve caregiver outcomes [3,12-19]. Research suggests that the significant unmet needs associated with support for communication and coordination among the care network may contribute to the often-suboptimal outcomes experienced by caregivers such as increased stress and burden [20,24,37]. Further, our findings related to caregiving workload and journal use combined with our qualitative findings provide some initial indication that caregivers may experience positive effects such as reduced burden, increased social support, and increased quality of life from using CareVirtue. Our immediate plan for future research is to conduct a randomized clinical trial to test the hypothesis that CareVirtue reduces caregiver stress and burden and increases caregiver quality of life.

Limitations

Our results should be considered in light of certain limitations. First, although our sample size was much larger than is typical for a feasibility study [38], the sample size should be considered when interpreting the univariate results, as it is possible that a small group of people could be driving our findings. Second, although we achieved enrollment of diversity in terms of income and location within the United States, participants were primarily White, married, college-educated women of 60 years of age who lived with the care recipient. Future work will endeavor to achieve a more sociodemographically diverse sample in terms of race/ethnicity, education, age, and distance from the person living with dementia. Third, although this study found a broad range of care network sizes consistent with previous literature [20], it is likely that the context of the COVID-19 pandemic may have reduced the number of in-home supports, which may have influenced the number of care network members using CareVirtue.

The results of this study establish the acceptability and feasibility of CareVirtue use among care networks of people living with ADRD. This study also highlights the importance of designing flexible, multicomponent interventions that allow care networks to tailor their engagement according to their needs and what is most useful to them. The results of this feasibility study will be used to improve CareVirtue feasibility and acceptability in preparation for a subsequent randomized trial to test CareVirtue's effectiveness in improving caregiver outcomes.

Acknowledgments

The authors acknowledge and thank the many caregivers who voluntarily supported this project. We also thank our community partners in Wisconsin and California who supported this work, including the Alzheimer's and Dementia Alliance of Wisconsin and Alzheimer's San Diego. This work was supported by grants (R41AG069607 and P30AG062715) from the NIH National Institute on Aging.

Conflicts of Interest

CE is the Founder and CEO of Whiplash Technology and the developer of CareVirtue. MZ is the Caregiver Support Officer for Whiplash Technology and supports development of CareVirtue; is an Associate for HFC, which is a 501c3 with a mission of Bringing Light to Alzheimer's; and is a member of the Alzheimer's Impact Movement, which is the advocacy affiliate of the Alzheimer's Association.

Multimedia Appendix 1
Walkthrough of CareVirtue features.
[MP4 File (MP4 Video), 10989 KB - aging_v5i3e36975_app1.mp4]

Multimedia Appendix 2
CareVirtue Care Guide template.
[PDF File (Adobe PDF File), 61 KB - aging_v5i3e36975_app2.pdf]

Multimedia Appendix 3
Semi-structured interview guide.

XSL•FO
RenderX

[DOCX File , 13 KB - aging_v5i3e36975_app3.docx]

Multimedia Appendix 4
Elbow plot for k-means clustering. Each care team is represented by an eight-dimensional vector with the following components: the number of logins, the number of journal posts, the number of journal post replies, the number of calendar events, the number of secondary caregiver invites sent, the number of secondary caregiver invites accepted, the number of Care Guide sections created, and the number of resources accessed through CareVirtue. We clustered all observations using k-means clustering, with a Euclidian distance metric and 10 random initializations. We varied k from 1 to 20 and used the elbow method (on the intra cluster distance) to choose the final k value. Figure A1 displays the intra cluster distance for each k value.
[PNG File , 65 KB - aging_v5i3e36975_app4.png]

Multimedia Appendix 5
P values for caregiver-reported confidence using CareVirtue over the duration of the study for the 24 participants included in the analysis (27 participants were excluded due to missingness).
[DOCX File , 14 KB - aging_v5i3e36975_app5.docx]

Multimedia Appendix 6
Histograms for primary caregiver responses to the behavioral intention scale, system usability scale, and perceived usefulness.
[PNG File , 52 KB - aging_v5i3e36975_app6.png]

Multimedia Appendix 7
Histograms for secondary caregiver responses to the behavioral intention scale, system usability scale, and perceived usefulness scale.
[PNG File , 54 KB - aging_v5i3e36975_app7.png]

References

1. Alzheimer's Disease and Related Dementias. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/aging/publications/features/alzheimers-disease-dementia.html [accessed 2022-07-16]
2. Monica MM, Díaz-Santos M, Vossel K. Alzheimer's Association 2021: Facts and Figures Report. Mary S. Easton Center for Alzheimer's Disease Research at UCLA. 2021. URL: https://www.eastonad.ucla.edu/images/E-Newsletter/Summer_2021/ucla-easton-center-summer-2021-e-newsletter-final.pdf [accessed 2022-07-17]
3. Boots LMM, de Vugt ME, van Knippenberg RJM, Kempen GIJM, Verhey FRJ. A systematic review of Internet-based supportive interventions for caregivers of patients with dementia. Int J Geriatr Psychiatry 2014 Apr 20;29(4):331-344. [doi: 10.1002/gps.4016] [Medline: 23963684]
4. Boots LMM, Wolfs CAG, Verhey FRJ, Kempen GIJM, de Vugt ME. Qualitative study on needs and wishes of early-stage dementia caregivers: the paradox between needing and accepting help. Int. Psychogeriatr 2015 Jan 08;27(6):927-936. [doi: 10.1017/s1041610214002804]
5. Werner NE, Stanislawski B, Marx K, Watkins D, Kobayashi M, Kales H, et al. Getting what they need when they need it. Appl Clin Inform 2017 Dec 20;26(01):191-205. [doi: 10.4338/aci-2016-07-ra-0122]
6. Next Steps for Research on Informal Caregiving. National Institute on Aging. 2012. URL: https://www.nia.nih.gov/sites/default/files/gerald-summary_11-21-14_0.pdf [accessed 2022-07-17]
7. Recommendations from the NIH AD Research Summit 2015. National Institute on Aging. 2015. URL: https://www.nia.nih.gov/research/recommendations-nih-ad-research-summit-2015 [accessed 2022-07-17]
8. Alzheimer's Association National Plan CareSupport Milestone Workgroup, Borson S, Boustani MA, Buckwalter KC, Burgio LD, Chodosh J, et al. Report on milestones for care and support under the U.S. National Plan to Address Alzheimer's Disease. Alzheimers Dement 2016 Mar;12(3):334-369 [FREE Full text] [doi: 10.1016/j.jalz.2016.01.005] [Medline: 26868060]
9. Hopwood J, Walker N, McDonagh L, Rait G, Walters K, Iliffe S, et al. Internet-Based Interventions Aimed at Supporting Family Caregivers of People With Dementia: Systematic Review. J Med Internet Res 2018 Jun 12;20(6):e216 [FREE Full text] [doi: 10.2196/jmir.9548] [Medline: 29895512]
10. Etxeberria I, Salaberria K, Gorostiaga A. Online support for family caregivers of people with dementia: a systematic review and meta-analysis of RCTs and quasi-experimental studies. Aging Ment Health 2021 Jul 02;25(7):1165-1180. [doi: 10.1080/13607863.2020.1758900] [Medline: 32363901]
11. Deeken F, Rezo A, Hinz M, Discher R, Rapp MA. Evaluation of Technology-Based Interventions for Informal Caregivers of Patients With Dementia-A Meta-Analysis of Randomized Controlled Trials. Am J Geriatr Psychiatry 2019 Apr;27(4):426-445. [doi: 10.1016/j.jagp.2018.12.003] [Medline: 30642650]

XSL•FO
RenderX

12. Chiu T, Marziali E, Colantonio A, Carswell A, Gruneir M, Tang M, et al. Internet-based caregiver support for Chinese Canadians taking care of a family member with alzheimer disease and related dementia. Can J Aging 2009 Dec;28(4):323-336. [doi: 10.1017/S0714980809990158] [Medline: 19925698]
13. Glueckauf RL, Ketterson TU, Loomis JS, Dages P. Online support and education for dementia caregivers: overview, utilization, and initial program evaluation. Telemed J E Health 2004;10(2):223-232. [doi: 10.1089/tmj.2004.10.223] [Medline: 15319052]
14. Bateman DR, Srinivas B, Emmett TW, Schleyer TK, Holden RJ, Hendrie HC, et al. Categorizing Health Outcomes and Efficacy of mHealth Apps for Persons With Cognitive Impairment: A Systematic Review. J Med Internet Res 2017 Aug 30;19(8):e301 [FREE Full text] [doi: 10.2196/jmir.7814] [Medline: 28855146]
15. Godwin KM, Mills WL, Anderson JA, Kunik ME. Technology-driven interventions for caregivers of persons with dementia: a systematic review. Am J Alzheimers Dis Other Demen 2013 May;28(3):216-222 [FREE Full text] [doi: 10.1177/1533317513481091] [Medline: 23528881]
16. Martínez-Alcalá CI, Pliego-Pastrana P, Rosales-Lagarde A, Lopez-Noguerola J, Molina-Trinidad EM. Information and Communication Technologies in the Care of the Elderly: Systematic Review of Applications Aimed at Patients With Dementia and Caregivers. JMIR Rehabil Assist Technol 2016 May 02;3(1):e6 [FREE Full text] [doi: 10.2196/rehab.5226] [Medline: 28582258]
17. Marziali E, Garcia LJ. Dementia caregivers' responses to 2 Internet-based intervention programs. Am J Alzheimers Dis Other Demen 2011 Feb;26(1):36-43 [FREE Full text] [doi: 10.1177/1533317510387586] [Medline: 21282276]
18. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec;25(8):674-679 [FREE Full text] [doi: 10.1177/1533317510385812] [Medline: 21131674]
19. van der Roest HG, Meiland FJM, Jonker C, Dröes RM. User evaluation of the DEMentia-specific Digital Interactive Social Chart (DEM-DISC). A pilot study among informal carers on its impact, user friendliness and, usefulness. Aging Ment Health 2010 May;14(4):461-470. [doi: 10.1080/13607860903311741] [Medline: 20455122]
20. Ponnala S, Werner NE. Exploring Informal Caregiver Workload using a Macroergonomics Lens on Multiple Resources. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2021 Feb 09;64(1):856-860. [doi: 10.1177/1071181320641198]
21. Friedman E, Kennedy D. Typologies of Dementia Caregiver Support Networks: A Pilot Study. Gerontologist 2021 Nov 15;61(8):1221-1230 [FREE Full text] [doi: 10.1093/geront/gnab013] [Medline: 33585929]
22. Esandi N, Nolan M, Alfaro C, Canga-Armayor A. Keeping Things in Balance: Family Experiences of Living With Alzheimer's Disease. Gerontologist 2018 Mar 19;58(2):e56-e67. [doi: 10.1093/geront/gnx084] [Medline: 28633376]
23. Neubert L, Gottschalk S, König HH, Brettschneider C. Dementia care-giving from a family network perspective in Germany: A typology. Health Soc Care Community 2022 Feb 16;30(2):579-591. [doi: 10.1111/hsc.13161] [Medline: 32939908]
24. Tang C, Chen Y, Cheng K, Ngo V, Mattison JE. Awareness and handoffs in home care: coordination among informal caregivers. Behaviour & Information Technology 2017 Nov 27;37(1):66-86. [doi: 10.1080/0144929x.2017.1405073]
25. Werner NE, Brown JC, Loganathar P, Holden RJ. Quality of Mobile Apps for Care Partners of People With Alzheimer Disease and Related Dementias: Mobile App Rating Scale Evaluation. JMIR Mhealth Uhealth 2022 Mar 29;10(3):e33863 [FREE Full text] [doi: 10.2196/33863] [Medline: 35348467]
26. Wozney L, Freitas de Souza LM, Kervin E, Queluz F, McGrath PJ, Keefe J. Commercially Available Mobile Apps for Caregivers of People With Alzheimer Disease or Other Related Dementias: Systematic Search. JMIR Aging 2018 Dec 07;1(2):e12274 [FREE Full text] [doi: 10.2196/12274] [Medline: 31518255]
27. Block L, Gilmore-Bykovskyi A, Jolliff A, Mullen S, Werner NE. Exploring dementia family caregivers' everyday use and appraisal of technological supports. Geriatr Nurs 2020 Nov;41(6):909-915 [FREE Full text] [doi: 10.1016/j.gerinurse.2020.06.019] [Medline: 32665095]
28. Boutilier JJ, Craig T, Sharpe MB, Chan TCY. Sample size requirements for knowledge-based treatment planning. Med Phys 2016 Mar 09;43(3):1212-1221. [doi: 10.1118/1.4941363] [Medline: 26936706]
29. Bangor A, Kortum PT, Miller JT. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
30. Holden RJ, Karsh B. A theoretical model of health information technology usage behaviour with implications for patient safety. Behaviour & Information Technology 2009 Jan;28(1):21-38. [doi: 10.1080/01449290601138245]
31. Asan O, Holden RJ, Flynn KE, Murkowski K, Scanlon MC. Providers' assessment of a novel interactive health information technology in a pediatric intensive care unit. JAMIA Open 2018 Jul 12;1(1):32-41 [FREE Full text] [doi: 10.1093/jamiaopen/ooy020] [Medline: 31528847]
32. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
33. Barry CA, Britten N, Barber N, Bradley C, Stevenson F. Using reflexivity to optimize teamwork in qualitative research. Qual Health Res 1999 Jan 01;9(1):26-44. [doi: 10.1177/104973299129121677] [Medline: 10558357]
34. Aledeh M, Habib Adam P. Caring for Dementia Caregivers in Times of the COVID-19 Crisis: A Systematic Review. AJNR 2020 Jan 24;8(5):552-561. [doi: 10.12691/ajnr-8-5-8]

XSL•FO
RenderX

35. Bacsu J, O'Connell ME, Cammer A, Azizi M, Grewal K, Poole L, et al. Using Twitter to Understand the COVID-19 Experiences of People With Dementia: Infodemiology Study. J Med Internet Res 2021 Feb 03;23(2):e26254 [FREE Full text] [doi: 10.2196/26254] [Medline: 33468449]
36. Young HM, Bell JF, Whitney RL, Ridberg RA, Reed SC, Vitaliano PP. Social Determinants of Health: Underreported Heterogeneity in Systematic Reviews of Caregiver Interventions. Gerontologist 2020 Feb 14;60(Suppl 1):S14-S28 [FREE Full text] [doi: 10.1093/geront/gnz148] [Medline: 32057083]
37. Marcum C, Ashida S, Koehly L. Primary Caregivers in a Network Context. J Gerontol B Psychol Sci Soc Sci 2020 Jan 01;75(1):125-136 [FREE Full text] [doi: 10.1093/geronb/gbx165] [Medline: 29304203]
38. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharmaceut. Statist 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
39. Wang J, Fu Y, Lou V, Tan S, Chui E. A systematic review of factors influencing attitudes towards and intention to use the long-distance caregiving technologies for older adults. Int J Med Inform 2021 Sep;153:104536. [doi: 10.1016/j.ijmedinf.2021.104536] [Medline: 34325206]
40. Mendez KJW, Budhathoki C, Labrique AB, Sadak T, Tanner EK, Han HR. Factors Associated With Intention to Adopt mHealth Apps Among Dementia Caregivers With a Chronic Condition: Cross-sectional, Correlational Study. JMIR Mhealth Uhealth 2021 Aug 31;9(8):e27926 [FREE Full text] [doi: 10.2196/27926] [Medline: 34463637]
41. Øksnebjerg L, Woods B, Ruth K, Lauridsen A, Kristiansen S, Holst HD, et al. A Tablet App Supporting Self-Management for People With Dementia: Explorative Study of Adoption and Use Patterns. JMIR Mhealth Uhealth 2020 Jan 17;8(1):e14694 [FREE Full text] [doi: 10.2196/14694] [Medline: 31951217]

Abbreviations

ADRD: Alzheimer disease and related dementias
MARS: Mobile Application Rating Scale
mHealth: mobile health
NASA-TLX: National Aeronautics and Space Administration-Task Load Index
SUS: System Usability Scale

Edited by J Wang; submitted 01.02.22; peer-reviewed by AS Rigaud, A Montilha; comments to author 23.02.22; revised version received 29.04.22; accepted 27.06.22; published 04.08.22.

Please cite as:
Boutilier JJ, Loganathar P, Linden A, Scheer E, Noejovich S, Elliott C, Zuraw M, Werner NE
A Web-Based Platform (CareVirtue) to Support Caregivers of People Living With Alzheimer Disease and Related Dementias: Mixed Methods Feasibility Study
JMIR Aging 2022;5(3):e36975
URL: https://aging.jmir.org/2022/3/e36975
doi:10.2196/36975
PMID:35925654

©Justin J Boutilier, Priya Loganathar, Anna Linden, Eleanore Scheer, Sofia Noejovich, Christian Elliott, Matthew Zuraw, Nicole E Werner. Originally published in JMIR Aging (https://aging.jmir.org), 04.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Short Paper

Automatically Identifying Twitter Users for Interventions to Support Dementia Family Caregivers: Annotated Data Set and Benchmark Classification Models

Ari Z Klein[1], PhD; Arjun Magge[1], PhD; Karen O'Connor[1], MS; Graciela Gonzalez-Hernandez[2], PhD

[1]Department of Biostatistics, Epidemiology, and Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, United States

[2]Department of Computational Biomedicine, Cedars-Sinai Medical Center, Los Angeles, CA, United States

Corresponding Author:
Ari Z Klein, PhD
Department of Biostatistics, Epidemiology, and Informatics
Perelman School of Medicine
University of Pennsylvania
Blockley Hall, 4th Fl.
423 Guardian Dr.
Philadelphia, PA, 19104
United States
Phone: 1 310 423 3521
Email: ariklein@pennmedicine.upenn.edu

Abstract

Background: More than 6 million people in the United States have Alzheimer disease and related dementias, receiving help from more than 11 million family or other informal caregivers. A range of traditional interventions has been developed to support family caregivers; however, most of them have not been implemented in practice and remain largely inaccessible. While recent studies have shown that family caregivers of people with dementia use Twitter to discuss their experiences, methods have not been developed to enable the use of Twitter for interventions.

Objective: The objective of this study is to develop an annotated data set and benchmark classification models for automatically identifying a cohort of Twitter users who have a family member with dementia.

Methods: Between May 4 and May 20, 2021, we collected 10,733 tweets, posted by 8846 users, that mention a dementia-related keyword, a linguistic marker that potentially indicates a diagnosis, and a select familial relationship. Three annotators annotated 1 random tweet per user to distinguish those that indicate having a family member with dementia from those that do not. Interannotator agreement was 0.82 (Fleiss kappa). We used the annotated tweets to train and evaluate support vector machine and deep neural network classifiers. To assess the scalability of our approach, we then deployed automatic classification on unlabeled tweets that were continuously collected between May 4, 2021, and March 9, 2022.

Results: A deep neural network classifier based on a BERT (bidirectional encoder representations from transformers) model pretrained on tweets achieved the highest F_1-score of 0.962 (precision=0.946 and recall=0.979) for the class of tweets indicating that the user has a family member with dementia. The classifier detected 128,838 tweets that indicate having a family member with dementia, posted by 74,290 users between May 4, 2021, and March 9, 2022—that is, approximately 7500 users per month.

Conclusions: Our annotated data set can be used to automatically identify Twitter users who have a family member with dementia, enabling the use of Twitter on a large scale to not only explore family caregivers' experiences but also directly target interventions at these users.

(JMIR Aging 2022;5(3):e39547) doi:10.2196/39547

KEYWORDS

natural language processing; social media; data mining; dementia; Alzheimer disease; caregivers

XSL•FO
RenderX

Introduction

More than 6 million people in the United States have Alzheimer disease and related dementias, and the burden is projected to double by 2060 [1]. Alzheimer disease is the sixth leading cause of death in the United States [2], and only 8% of people with dementia do not receive help from family members or other informal care providers [3], amounting to more than 11 million family or other unpaid caregivers in 2020 [4]. Caregivers of people with dementia are impacted physically, cognitively, socially, mentally, and financially. For instance, compared with noncaregivers, they are more vulnerable to disease due to chronic stress [5] and have lower durations and quality of sleep [6]. Compared with non–dementia caregivers, they are more likely to experience a decline in cognition [7] and social network size [8]. They are also more likely to experience depression compared with noncaregivers [9] and non–dementia caregivers [10], and depressive symptoms in dementia caregivers are associated with increased health care use and costs [11]. In addition to the increased costs of their personal health care, family caregivers of people with dementia pay for much of the recipient's total care costs, with the costs being significantly higher for people with dementia than without dementia [12].

A range of traditional interventions has been developed to support family caregivers of people with dementia [13]; however, most of them have not been implemented in practice and remain largely inaccessible [14]. Recent systematic reviews have concluded that internet-based interventions are valued by family caregivers of people with dementia for their easy access [15] and can have beneficial effects on caregivers' health [16]. While recent studies [17-23] have shown that family caregivers of people with dementia use Twitter to discuss their experiences, to the best of our knowledge, methods have not been developed to enable the use of Twitter as a platform for internet-based interventions. Given that nearly 1 of every 4 adults in the United States uses Twitter [24], Twitter may present a novel opportunity to reach family caregivers on a large scale, such as through user-targeted advertisements providing information about dementia, caregiving, resources, or services. The objective of this study was to develop an annotated data set and benchmark classification models for automatically identifying a cohort of Twitter users who have a family member with dementia.

Methods

Ethical Considerations

The data used in this study were collected in accordance with the Twitter Terms of Service. The Institutional Review Board of the University of Pennsylvania reviewed this study (protocol number: 828972) and deemed it exempt human subjects research under 45 CFR §46.101(b)(4) for publicly available data sources.

Data Collection and Annotation

Between May 4 and May 20, 2021, we collected 67,060 publicly available tweets from the Twitter streaming application programming interface (API) that are in English, are not retweets, and include both a dementia-related keyword (eg, *dementia*, *youngdementia*, *#yod*, *#ftd*, *alzheimer's*, *alz*, *alzheimersdisease*, *mild cognitive impairment*) and a linguistic marker that potentially indicates a diagnosis (eg, *diagnosed*, *diagnosis*, *has*, *got*, *developed*, *with*, *from*). The full list of API search terms is available in Multimedia Appendix 1. We then searched these tweets for references to select familial relationships (Multimedia Appendix 2), identifying 10,733 (16%) of the 67,060 tweets. We randomly sampled 1 tweet per user—8846 (82%) of the 10,733 tweets—and developed annotation guidelines (Multimedia Appendix 3) to help 3 annotators distinguish tweets that indicate having a family member with dementia from those that do not. Among the 8846 annotated tweets, 8346 (94%) were dual annotated, and 500 (6%) were annotated by all 3 annotators. Interannotator agreement, based on the 500 tweets annotated by all 3 annotators, was 0.82 (Fleiss kappa). Upon resolving the disagreements, it was determined that 5946 (67%) of the tweets indicate that the user has a family member with dementia, and 2900 (33%) of the tweets do not.

Automatic Classification

We performed benchmark supervised machine learning experiments to assess the utility of the annotated data set for automatically identifying Twitter users who have a family member with dementia. For the classifiers, we used the LibSVM [25] implementation of support vector machine (SVM) in Weka and SVM and 6 deep neural network classifiers based on BERT (bidirectional encoder representations from transformers): the BERT-Base-Uncased [26], DistilBERT-Base-Uncased [27], RoBERTa-Large [28], BioBERT-Large-Cased [29], Bio+ClinicalBERT [30], and BERTweet-Large [31] pretrained models in the *Flair* Python library. We split the 8846 tweets into 80% (7077 tweets) and 20% (1769 tweets) random sets as training data (Multimedia Appendix 4) and held-out test data, respectively, stratified based on the distribution of the binary annotated classes. For the SVM classifier, we preprocessed the tweets by normalizing URLs, usernames, digits, and keywords related to dementia (Multimedia Appendix 1) and familial relationships (Multimedia Appendix 2), removing nonalphanumeric characters and extra spaces, and lowercasing and stemming [32] the text. We used the Weka NGram Tokenizer to extract n-grams (n=1-3) as features in a bag-of-words representation. We used the radial basis function kernel and set the *cost* at c=32. For the BERT-based classifiers, we preprocessed the tweets by normalizing URLs and usernames and lowercasing the text. For training, we used stochastic gradient descent optimization, a batch size of 8, 15 epochs, and a learning rate of 0.001. During training, we fine-tuned all layers of the transformer model with our annotated tweets. To optimize performance, the model was evaluated after each epoch on a 5% split of the training set. To assess the scalability of our approach, we then deployed automatic classification on 198,674 unlabeled tweets, posted by 119,640 users, that were continuously collected from the Twitter streaming API (Multimedia Appendix 1) between May 4, 2021, and March 9, 2022, and mentioned a select familial relationship (Multimedia Appendix 2).

XSL•FO
RenderX

Results

Table 1 presents the precision, recall, and F_1-scores of SVM and 6 deep neural network classifiers for the class of tweets indicating that the user has a family member with dementia, evaluated on a held-out test set of 1769 (20%) of the 8846 manually annotated tweets. The classifier based on a model pretrained on tweets (BERTweet-Large) achieved the highest F_1-score: 0.962 (precision=0.946 and recall=0.979). When deployed on 198,674 unlabeled tweets, posted by 119,640 users, between May 4, 2021, and March 9, 2022, the BERTweet classifier detected 128,838 tweets indicating that the user has a family member with dementia, posted by 74,290 users—that is, approximately 7500 users per month.

Table 2 presents examples of false positives and false negatives of the BERTweet classifier in the test set. Among the 68 false positives, 36 (47%) refer to people with dementia who are not or may not be select family members (Tweet 1), 8 (12%) report that a family member has a condition other than dementia (Tweet 2), and 5 (7%) merely speculate that a family member has dementia (Tweet 3). Another 8 (12%) of the 68 false positives were a result of manual annotation errors. Among the 25 false negatives, 14 (56%) use deixis or anaphora, requiring additional context in the tweet to understand that a non–first person determiner (eg, “their” in Tweet 4) actually refers to the user, or that a personal pronoun (eg, “she” in Tweet 5) refers to a select family member with dementia. Furthermore, 12 (86%) of these 14 tweets also include references to people who are not family members or do not have dementia. Another 4 (16%) of the 25 false negatives were a result of manual annotation errors.

Table 1. Precision, recall, and F_1-scores of classifiers for detecting tweets indicating that the user has a family member with dementia.

Classifier	Precision	Recall	F_1-score
SVM[a]	0.884	0.939	0.910
BERT[b]-Base-Uncased	0.924	0.954	0.938
DistilBERT-Base-Uncased	0.930	0.942	0.936
RoBERTa-Large	0.918	0.982	0.949
BioBERT-Large-Cased	0.907	0.978	0.941
Bio+ClinicalBERT	0.903	0.958	0.930
BERTweet-Large	0.946	0.979	0.962

[a]SVM: support vector machine.

[b]BERT: bidirectional encoder representations from transformers.

Table 2. Sample false positives and false negatives of a BERTweet classifier for detecting tweets indicating that the user has a select family member with dementia.

Tweet number	Tweet	Actual	Predicted
1	Evelyn has dementia, I know. But when she asked me today how my dad was doing... it still hurt.	–	+
2	We really don't have a clue about what causes Alzheimer's. We don't have a clue about Parkinson's, which is what got my dad, either.	–	+
3	I just listened to the Everywhere at The End of Time, by The Caretaker, and thought about my grandmother. The songs are about dementia, something my grandma wasn't clearly diagnosed with, but it hit hard.	–	+
4	If someone tells u their parent has Alzheimer's please don’t say your grandparent or great aunt did too. I appreciate that u can relate to the experience but it is so different. Tell me a different time.	+	–
5	I have a family member who is vulnerable and two children in their late 20s. I didn’t want to risk passing virus to her or from her to my family member. My sister made a bubble with her and her carers. She has dementia so she probably hasn’t missed me!	+	–

Discussion

Principal Findings

The benchmark performance of automatic classification demonstrates that our annotated data set has utility for accurately identifying Twitter users who have a family member with dementia, and deploying automatic classification on unlabeled tweets demonstrates that a large cohort of users can be identified. Therefore, our annotated data set enables the use of Twitter to scale up accessible, internet-based interventions directly targeted at family caregivers of people with dementia. Because our approach involves identifying tweets that mention a familial relationship, it would also enable interventions to be tailored to the care recipient.

XSL•FO
RenderX

Limitations

Our approach to identifying family caregivers assumes that having "close" relatives with dementia would likely imply the users' involvement in caregiving; however, the users identified in this study may not necessarily be caregivers or may have been caregivers but are no longer. We took this approach because we believe that limiting our identification of caregivers to users who explicitly state that they are providing ongoing care would underutilize the potential of Twitter for reaching caregivers on a large scale.

Conclusions

This paper presented an annotated data set and benchmark classification models for automatically identifying Twitter users who have a family member with dementia, enabling the use of Twitter on a large scale to not only explore family caregivers' experiences among their tweets but also directly target interventions at these users.

Acknowledgments

This work was supported by the National Library of Medicine (R01LM011176). The authors thank Ivan Flores for contributing to software applications, and Alexis Upshur and Aiden McRobbie-Johnson for contributing to annotating the Twitter data.

Authors' Contributions

AZK designed the data collection, edited the annotation guidelines, performed the support vector machine classification experiments, conducted the error analysis, and wrote the manuscript. AM performed the deep learning classification experiments, deployed the BERTweet classifier, and edited the manuscript. KO developed the annotation guidelines, annotated the Twitter data, and edited the manuscript. GGH conceptualized and guided the study and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Twitter streaming application programming interface search terms.
[TXT File , 3 KB - aging_v5i3e39547_app1.txt]

Multimedia Appendix 2
Family member keywords.
[TXT File , 0 KB - aging_v5i3e39547_app2.txt]

Multimedia Appendix 3
Annotation guidelines.
[PDF File (Adobe PDF File), 119 KB - aging_v5i3e39547_app3.pdf]

Multimedia Appendix 4
Training data.
[TXT File , 159 KB - aging_v5i3e39547_app4.txt]

References

1. Matthews KA, Xu W, Gaglioti AH, Holt JB, Croft JB, Mack D, et al. Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015-2060) in adults aged ≥65 years. Alzheimers Dement 2019 Jan 19;15(1):17-24 [FREE Full text] [doi: 10.1016/j.jalz.2018.06.3063] [Medline: 30243772]
2. Kochanek K, Xu J, Arias E. Mortality in the United States, 2019. NCHS Data Brief 2020 Dec(395):1-8. [Medline: 33395387]
3. Kasper JD, Freedman VA, Spillman BC, Wolff JL. The disproportionate impact of dementia on family and unpaid caregiving to older adults. Health Aff (Millwood) 2015 Oct;34(10):1642-1649 [FREE Full text] [doi: 10.1377/hlthaff.2015.0536] [Medline: 26438739]
4. Alzheimer's Association. 2021 Alzheimer's disease facts and figures. Alzheimers Dement 2021 Mar;17(3):327-406. [doi: 10.1002/alz.12328] [Medline: 33756057]
5. Fonareva I, Oken BS. Physiological and functional consequences of caregiving for relatives with dementia. Int Psychogeriatr 2014 May;26(5):725-747 [FREE Full text] [doi: 10.1017/S1041610214000039] [Medline: 24507463]
6. Gao C, Chapagain NY, Scullin MK. Sleep duration and sleep quality in caregivers of patients with dementia: a systematic review and meta-analysis. JAMA Netw Open 2019 Aug 02;2(8):e199891 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.9891] [Medline: 31441938]

XSL•FO
RenderX

7. Dassel K, Carr D, Vitaliano P. Does caring for a spouse with dementia accelerate cognitive decline? Findings from the health and retirement study. Gerontologist 2017 Apr 01;57(2):319-328. [doi: 10.1093/geront/gnv148] [Medline: 26582383]
8. Liu C, Fabius CD, Howard VJ, Haley WE, Roth DL. Change in social engagement among incident caregivers and controls: findings from the caregiving transitions study. J Aging Health 2021 Jan 23;33(1-2):114-124. [doi: 10.1177/0898264320961946] [Medline: 32962491]
9. Ma M, Dorstyn D, Ward L, Prentice S. Alzheimers' disease and caregiving: a meta-analytic review comparing the mental health of primary carers to controls. Aging Ment Health 2018 Nov 05;22(11):1395-1405. [doi: 10.1080/13607863.2017.1370689] [Medline: 28871796]
10. Sheehan O, Haley W, Howard V, Huang J, Rhodes J, Roth D. Stress, burden, and well-being in dementia and nondementia caregivers: insights from the caregiving transitions study. Gerontologist 2021 Jul 13;61(5):670-679 [FREE Full text] [doi: 10.1093/geront/gnaa108] [Medline: 32816014]
11. Zhu CW, Scarmeas N, Ornstein K, Albert M, Brandt J, Blacker D, et al. Health-care use and cost in dementia caregivers: Longitudinal results from the Predictors Caregiver Study. Alzheimers Dement 2015 Apr 17;11(4):444-454 [FREE Full text] [doi: 10.1016/j.jalz.2013.12.018] [Medline: 24637299]
12. Kelley AS, McGarry K, Bollens-Lund E, Rahman O, Husain M, Ferreira KB, et al. Residential setting and the cumulative financial burden of dementia in the 7 years before death. J Am Geriatr Soc 2020 Jun 18;68(6):1319-1324 [FREE Full text] [doi: 10.1111/jgs.16414] [Medline: 32187655]
13. Gaugler JE, Potter T, Pruinelli L. Partnering with caregivers. Clin Geriatr Med 2014 Aug;30(3):493-515. [doi: 10.1016/j.cger.2014.04.003] [Medline: 25037292]
14. Gitlin L, Marx K, Stanley I, Hodgson N. Translating evidence based dementia caregiving interventions into practice: state-of-the-science and next steps. Gerontologist 2015 Apr;55(2):210-226 [FREE Full text] [doi: 10.1093/geront/gnu123] [Medline: 26035597]
15. Hopwood J, Walker N, McDonagh L, Rait G, Walters K, Iliffe S, et al. Internet-based interventions aimed at supporting family caregivers of people with dementia: systematic review. J Med Internet Res 2018 Jun 12;20(6):e216 [FREE Full text] [doi: 10.2196/jmir.9548] [Medline: 29895512]
16. Leng M, Zhao Y, Xiao H, Li C, Wang Z. Internet-based supportive interventions for family caregivers of people with dementia: systematic review and meta-analysis. J Med Internet Res 2020 Sep 09;22(9):e19468 [FREE Full text] [doi: 10.2196/19468] [Medline: 32902388]
17. Yoon S. What can we learn about mental health needs from tweets mentioning dementia on World Alzheimer's Day? J Am Psychiatr Nurses Assoc 2016 Nov 01;22(6):498-503 [FREE Full text] [doi: 10.1177/1078390316663690] [Medline: 27803262]
18. Danilovich M, Tsay J, Al-Bahrani R, Choudhary A, Agrawal A. #Alzheimer's and dementia: expressions of memory loss on Twitter. Topics in Geriatric Rehabilitation 2018;34(1):48-53. [doi: 10.1097/TGR.0000000000000173]
19. Cheng TY, Liu L, Woo BK. Analyzing Twitter as a platform for Alzheimer-related dementia awareness: thematic analyses of Tweets. JMIR Aging 2018 Dec 10;1(2):e11542 [FREE Full text] [doi: 10.2196/11542] [Medline: 31518232]
20. Yoon S, Lucero R, Mittelman MS, Luchsinger JA, Bakken S. Mining Twitter to inform the design of online interventions for Hispanic Alzheimer's disease and related dementias caregivers. Hisp Health Care Int 2020 Sep 24;18(3):138-143. [doi: 10.1177/1540415319882777] [Medline: 31646904]
21. Mehta N, Zhu L, Lam K, Stall NM, Savage R, Read SH, et al. Health forums and Twitter for dementia research: opportunities and considerations. J Am Geriatr Soc 2020 Dec 07;68(12):2881-2889. [doi: 10.1111/jgs.16790] [Medline: 32894780]
22. Bacsu J, O'Connell ME, Cammer A, Azizi M, Grewal K, Poole L, et al. Using Twitter to understand the COVID-19 experiences of people with dementia: infodemiology study. J Med Internet Res 2021 Feb 03;23(2):e26254 [FREE Full text] [doi: 10.2196/26254] [Medline: 33468449]
23. Yoon S, Broadwell P, Alcantara C, Davis N, Lee H, Bristol A, et al. Analyzing topics and sentiments from Twitter to gain insights to refine interventions for family caregivers of persons with Alzheimer's disease and related dementias (ADRD) during COVID-19 pandemic. Stud Health Technol Inform 2022 Jan 14;289:170-173 [FREE Full text] [doi: 10.3233/SHTI210886] [Medline: 35062119]
24. Auxier B, Anderson M. Social media use in 2021. Pew Research Center. 2021 Apr 07. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2022-02-25]
25. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2011 Apr;2(3):1-27. [doi: 10.1145/1961189.1961199]
26. Devlin J, Cheng M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. 2019 Presented at: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers); June 2-7, 2019; Minneapolis, Minnesota, US p. 4171-4186. [doi: 10.18653/v1/N19-1423]
27. Sanh V, Debut L, Chaumond J, Wolf T. DistilBERT, a distilled version of BERTmaller, faster, cheaper and lighter. 2019 Presented at: Fifth Workshop on Energy Efficient Machine Learning and Cognitive Computing; December 13, 2019; Vancouver, Canada.

XSL•FO
RenderX

28. Liu Y, Ott M, Goyal N, Du J, Joshi M, Chen D, et al. RoBERTa: a robustly optimized BERT pretraining approach. arXiv Preprint posted online July 26, 2019. [FREE Full text]
29. Lee J, Yoon W, Kim S, Kim D, Kim S, So C, et al. BioBERT: a pre-trained biomedical language representation model for biomedical text mining. Bioinformatics 2020 Feb 15;36(4):1234-1240 [FREE Full text] [doi: 10.1093/bioinformatics/btz682] [Medline: 31501885]
30. Alsentzer E, Murphy J, Boag W, Weng W, Jindi D, Naumann T, et al. Publicly available clinical BERT embeddings. 2019 Presented at: Proceedings of the 2nd Clinical Natural Language Processing Workshop; June 7, 2019; Minneapolis, Minnesota, USA p. 72-78. [doi: 10.18653/v1/w19-1909]
31. Nguyen D, Vu T, Nguyen A. BERTweet: a pre-trained language model for English tweets. 2020 Presented at: Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations; November 16-20, 2020; Online p. 9-14. [doi: 10.18653/v1/2020.emnlp-demos.2]
32. Porter MF. An algorithm for suffix stripping. Program: electronic library and information systems 1980;14(3):130-137. [doi: 10.1108/eb046814]

Abbreviations

API: application programming interface
BERT: bidirectional encoder representations from transformers
SVM: support vector machine

Edited by J Wang, T Leung; submitted 16.05.22; peer-reviewed by K Verspoor, JW Kwon; comments to author 27.06.22; revised version received 08.07.22; accepted 08.07.22; published 16.09.22.

Please cite as:
Klein AZ, Magge A, O'Connor K, Gonzalez-Hernandez G
Automatically Identifying Twitter Users for Interventions to Support Dementia Family Caregivers: Annotated Data Set and Benchmark Classification Models
JMIR Aging 2022;5(3):e39547
URL: https://aging.jmir.org/2022/3/e39547
doi:10.2196/39547
PMID:36112408

©Ari Z Klein, Arjun Magge, Karen O'Connor, Graciela Gonzalez-Hernandez. Originally published in JMIR Aging (https://aging.jmir.org), 16.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Caregiver Availability Using Medical Notes With Rule-Based Natural Language Processing: Retrospective Cohort Study

Elham Mahmoudi[1,2], PhD; Wenbo Wu[3], PhD; Cyrus Najarian[4], MA; James Aikens[1], PhD; Julie Bynum[4], MPH, MD; V G Vinod Vydiswaran[5], PhD

[1]Department of Family Medicine, Medical School, University of Michigan, Ann Arbor, MI, United States
[2]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States
[3]Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, United States
[4]Medical School, University of Michigan, Ann Arbor, MI, United States
[5]Department of Learning Health Sciences, Medical School, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Elham Mahmoudi, PhD
Department of Family Medicine
Medical School
University of Michigan
NCRC Building 14, Room G234
2800 Plymouth Rd
Ann Arbor, MI, 48109
United States
Phone: 1 248 765 3927
Email: mahmoudi@med.umich.edu

Abstract

Background: Identifying caregiver availability, particularly for patients with dementia or those with a disability, is critical to informing the appropriate care planning by the health systems, hospitals, and providers. This information is not readily available, and there is a paucity of pragmatic approaches to automatically identifying caregiver availability and type.

Objective: Our main objective was to use medical notes to assess caregiver availability and type for hospitalized patients with dementia. Our second objective was to identify whether the patient lived at home or resided at an institution.

Methods: In this retrospective cohort study, we used 2016-2019 telephone-encounter medical notes from a single institution to develop a rule-based natural language processing (NLP) algorithm to identify the patient's caregiver availability and place of residence. Using note-level data, we compared the results of the NLP algorithm with human-conducted chart abstraction for both training (749/976, 77%) and test sets (227/976, 23%) for a total of 223 adults aged 65 years and older diagnosed with dementia. Our outcomes included determining whether the patients (1) reside at home or in an institution, (2) have a formal caregiver, and (3) have an informal caregiver.

Results: Test set results indicated that our NLP algorithm had high level of accuracy and reliability for identifying whether patients had an informal caregiver (F_1=0.94, accuracy=0.95, sensitivity=0.97, and specificity=0.93), but was relatively less able to identify whether the patient lived at an institution (F_1=0.64, accuracy=0.90, sensitivity=0.51, and specificity=0.98). The most common explanations for NLP misclassifications across all categories were (1) incomplete or misspelled facility names; (2) past, uncertain, or undecided status; (3) uncommon abbreviations; and (4) irregular use of templates.

Conclusions: This innovative work was the first to use medical notes to pragmatically determine caregiver availability. Our NLP algorithm identified whether hospitalized patients with dementia have a formal or informal caregiver and, to a lesser extent, whether they lived at home or in an institutional setting. There is merit in using NLP to identify caregivers. This study serves as a proof of concept. Future work can use other approaches and further identify caregivers and the extent of their availability.

(JMIR Aging 2022;5(3):e40241) doi:10.2196/40241

XSL•FO
RenderX

KEYWORDS

natural language processing; caregiver; medical notes; Alzheimer; dementia; pragmatic; aging; care planning; health care; elderly care; elderly population; algorithm

Introduction

Clinical practice creates a large amount of structured and unstructured data [1,2]. Although the electronic medical record (EMR) has allowed health care systems to collect clinical encounter data, the collection process and reporting are still inefficient. This inefficiency is burdensome for health care workers and providers and may negatively impact patient care [1,2]. Furthermore, a large portion of the data in health care is in a free-text format. The data are entered into the system by multiple individuals (medical students, nurses, social workers, etc) and lack a specific template, are not easily searchable by health care workers, and are not readily available for clinical decision-making. Applying natural language processing (NLP) to medical notes has shown promising results in diagnosing certain conditions [3,4], predicting adverse health events [5,6], and identifying social determinants of health [7].

Systematic collection of caregiver information in EMR is a challenging task [8]. Although caregivers play an essential role in the health and well-being of people with complex care needs, such as those with dementia or a disability [9], health care systems are not equipped to readily identify caregiver availability (or lack thereof), type of care provided, time availability, and other helpful information about caregiver support. Despite the emergence of NLP in health care [10-14], there is a paucity of work examining the pragmatic collection of caregiver information [9].

Approximately 6 million older adults in the United States live with dementia. This number is expected to double by 2050 [15]. Because of more cognitive and physical limitations, compared with other older adults, people with dementia often have complex care management needs, and their well-being depends on their caregivers [16-19]. For example, postdischarge care coordination with a patient's caregiver may reduce readmission or other adverse health events. It is critical for the health systems to quickly identify and act upon caregiver availability information for patients with complex care needs, particularly after hospital discharge.

In this work, we aimed to provide a proof of concept that NLP can reliably identify caregiver availability and type via medical notes. we examined the following three outcomes: (1) whether a patient lives at home vs in an institution, (2) whether a patient has a formal caregiver (paid), and (3) whether a patient has an informal caregiver (eg, a family member). We hypothesized that using NLP, we would be able to reliably determine each of the above outcomes.

Methods

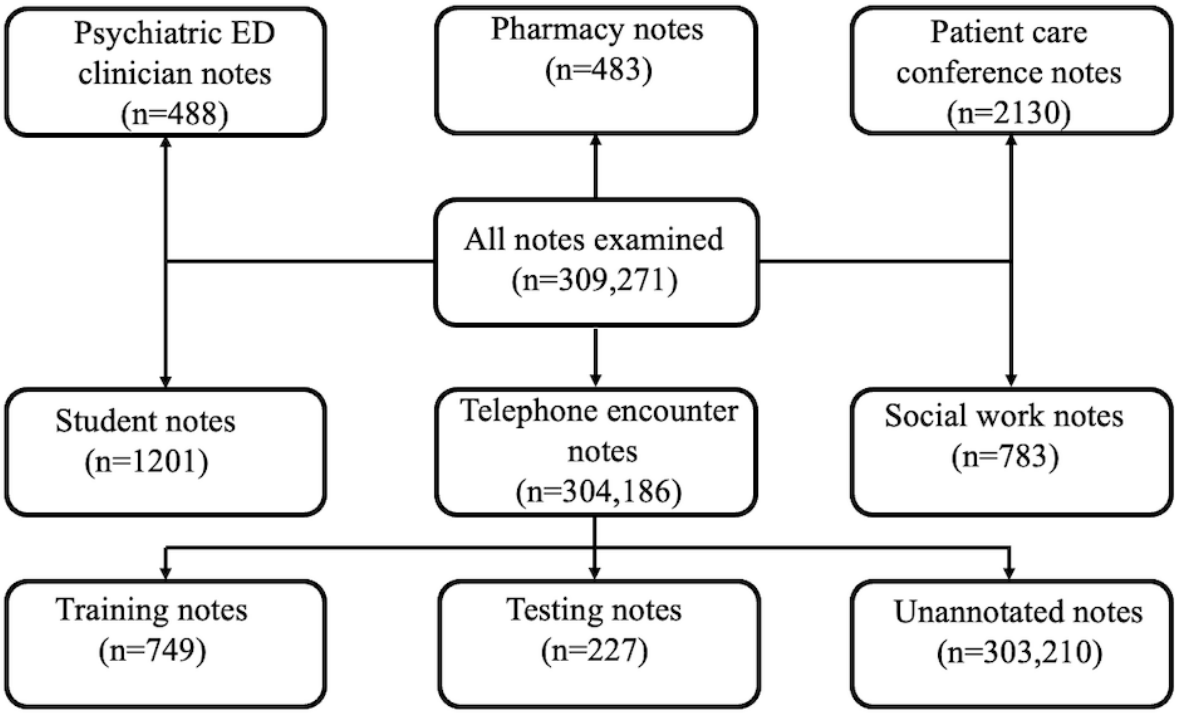
Data Source

To examine caregiver availability and type of caregivers, if any, for patients diagnosed with dementia, we used medical notes from Michigan Medicine (MM)a large academic medical center in Southeast Michigan. Our initial patient cohort was identified using the International Classification of Disease, 10th revision codes (Table S1 in Multimedia Appendix 1) from structured EMRs between October 2015 and January 2020. Using a 1-year look back period, we identified 2205 unique patients with dementia with at least one hospitalization in MM.

There are 60 different types of medical notes in MM. We randomly explored 10 notes from each category to identify the most promising type of notes for this study. Moreover, we sought expert advice from a geriatric MM nurse to identify the most promising medical notes for information about caregivers. Both approaches led us to use *telephone encounter* notes. Out of 2205 unique patients, 2017 had at least one telephone encounter note. We randomly selected and annotated a total of 976 telephone encounter notes (n=224 unique number of patients), of which 749 (77%) and 227 (23%) notes were partitioned into training (n=167 unique number of patients) and test (n=57 unique number of patients) sets, respectively. Furthermore, we ensured that all notes for each patient were kept within the same set. Figure 1 presents a schematic flow diagram of our sampling process.

XSL•FO
RenderX

Figure 1. Schematic flow diagram (source: 2016-2019 Michigan Medicine Electronic Medical Records). ED: emergency department.



Annotations

To accomplish high interrater reliability, 2 team members, a nurse experienced in reading and writing medical notes at MM and a social scientist with no medical background, independently annotated all notes. Discrepancies in annotation were resolved with all team members' participation. Our research questions included the following:

1. Does the patient reside at home or in an institution?
2. Does the patient have a formal (hired) caregiver?
3. Does the patient have an informal (a family member or a friend) caregiver?

The above research questions were chosen because the place of residence and caregiver availability are interconnected. Furthermore, our caregiver features were not mutually exclusive because it is plausible that the patient has both an informal and formal caregiver simultaneously [9]. Each criterion had the following two levels for annotation: 0 (based on lack of information or explicit negation) and 1 (based on implied or explicit narration in the note). If the note had no information about potential outcomes, we coded all features as zeros. Since a patient's circumstances (place of residence and caregiver availability) may change over time, our unit of analysis was the note (nested within individual patients). Each note was annotated independently, relying only on information found in that note. Using this method, we sought to identify the patient's place of residence and caregiver availability longitudinally at each specific time.

Model

First, we preprocessed the data based on patterns we saw in the training set and then used 2 lexicons to construct a rule-based approach to characterizing each note. We measured the model's performance in training and test sets, separately, using the F_1-score, accuracy, sensitivity, and specificity of the model against our gold standard—manual annotations of the notes. F_1-score summarizes the predictive power of an algorithm as the harmonic mean of precision and recall. Accuracy measures how many observations—positive and negative—were correctly classified [20].

Preprocessing

Through our annotation process, we discovered multiple terms that frequently led to false positives. For example, "family medicine" raised a false positive for "family," and "patient portal" raised a false positive for "patient." Additionally, some notes contain template sections and subheading phrases such as "Family History," which would list multiple familial relations that, in this context, would not be caregivers. These words and phrases were removed from medical notes before applying our algorithm (items 3 and 4 in the "Description of Rule-Based Algorithm" below).

Lexicon

Our 2 lexicons were dictionaries of terms used to identify (1) place of residence and (2) type of caregiver, if any. Specific terms (eg, "home," "atria ann arbor," and "linden square") were used to determine the current place of residence. To determine if a patient resides in an institution or at home, we used a list of nursing homes and care facilities in Washtenaw County (obtained from the University of Michigan) and a list of skilled nursing facilities in the state of Michigan (obtained from the Centers for Medicare and Medicaid Services website) [21]. Caregiver type was categorized into 2 groups—formal and informal—using general terms for caregivers (eg, "sister," "husband," or "visiting nurse"). To determine the presence and type of caregivers, we used a list of commonly used terms for friends and family members and formal caregivers based on our

XSL•FO
RenderX

consultations with practitioners (nurses, physicians, and care coordinators at MM).

Our data dictionary can be found in alphabetical order in Table S2 in Multimedia Appendix 1. Should a term from the dictionary be found in a medical note, we would use the corresponding labels associated with the term to characterize the note. For example, if a "visiting nurse" and "spouse" were both found in a note, the algorithm would rule that the patient had both formal (visiting nurse) and informal (spouse) caregivers.

Multiple rules were implemented to account for a more complex logic in determining the place of residence and caregiver presence or type. Throughout the algorithm, a 4-word window, rather than a fewer- or more-word window, was used because the 4-word window achieved the best accuracy in the training set.

Patient Verb Neighborhood

A lexicon was created that included verbs such as "agreed," "asked," or "reported," suggesting that a patient resides at home if these verbs appeared within a 4-word window of the following terms: "pt," "patient," or "patient's" (Table S3 in Multimedia Appendix 1). This was used to determine whether the patient had any relevant caregiver information, even if identified terms in the dictionary were not found in the note. Since health care workers would often discuss the "patient" in a non–caregiver related context, we could not simply search for the previous 3 words.

Institution Negation Neighborhood

In many cases, an institution's name appears in the note without any relation to patient's current place of residence or status of caregiver support. Examples include potential plans when a patient was discharged from the hospital or in discussion with family members for referral to an institution or if a patient was in an institution for subacute rehabilitation, which would not be considered a place of residence or a caregiver. We created a lexicon for institution negation words and searched for them within a 4-word window of each institution's name. If the specified negation words were found, then the institution was disregarded for that note.

If no terms were found, it was determined that the note had no information available to predict the place of residence or caregiver presence, and all fields were set to zero.

Description of the Rule-Based Algorithm

The 13 selection criteria, provided in the annotation guidelines and coded in order in the algorithm, are described below:

1. Replace "PT" (uppercase) with "physical therapy" in the original medical notes to avoid erroneous pickups of "pt" as an abbreviation of "patient."
2. Convert the original text in notes to lowercase letters.
3. Remove the following patterns in the lowercase notes to avoid false positives: "nurse navigator," "navigator nurse," "patient portal," "patient name," "relationship to patient" followed by blank spaces with no answer, "e.g., visiting nurse," "patient & caregiver," "patient or caregiver," "patient &/or caregiver," "family medicine," "family practice," "family doctor," "family physician," "alone with family," "verbalizes understanding," and "verbalized understanding."
4. To avoid falsely labeling "family" as informal caregivers, we removed each occurrence of "family history" and all words that follow until a new line character.
5. Remove all occurrences of "aid" and "aids" in a lowercase note when any of the following shows up: "sleeping aid," "sleeping aids," "sleep aid," "sleep aids," "hearing aid," "hearing aids," "hear aid," and "hear aids."
6. Substitute the following patterns with "patient" to avoid falsely picking up "want" in the proximity of "patient" or "pt": "want the patient," "wants the patient," "wanting the patient," "want the pt," "wants the pt," "wanting the pt," "want pt," "wants pt," and "wanting pt."
7. Substitute "pt or ot," "pt and ot," or "pt/ot" with "physical therapy and ot" to avoid falsely picking up "pt" as "patient."
8. Substitute "e-mail" with "email" before tokenization to avoid "e-mail" being split into "e" and "mail."
9. Substitute variants of "patient partner" (eg, "patient's partner," "patients partner," and "patient\ns partner" with "\n" being a new line character) with itself.
10. Oftentimes, the evidence of a "visiting nurse" may present itself as a variant (eg, "visit from a home care nurse"). To avoid missing such cases, for each sentence containing "nurse," we searched for variants of "visiting" (eg, "visit") before the occurrence of "nurse" within the sentence or searched for variants of "visiting" after the occurrence of "nurse" within that sentence and the sentence that follows.
11. To avoid falsely labeling informal caregiver or home when an institutional caregiver is present, in each lowercase medical note, we removed all occurrences of "patient," "pt," "care giver," "caregiver," and "guardian" in any sentence that included an institutional n-gram.
12. If any institutional term in the dictionary showed up in the note, there is evidence of institutional caregiver. To rule out false positives as potential, past, or unapproved institutional caregivers or when the service is for rehabilitation purpose only (eg, "patient discharged from Glacier Hills," "returned home from Glacier Hills," and "on the waiting list for Glacier Hills"), we looked for variants of "return from" (eg, "returning from" and "returned from"), "discharged from" (eg, "discharges from" and "discharging from"), "waiting list" (eg, "wait list" and "waitlist"), "cancel," "decline," "approve," require," "suggest," "request," and "rehabilitation" (eg, "rehab") in the sentence containing an institutional term. If any of the variants shows up in the sentence, all occurrences of the institutional term would be disregarded in the note.
13. To see whether there is evidence of home self-care (home=1) when an institutional caregiver is absent, we looked for patient verbs within a prespecified (n=4) word window of "patient," "pt," or "patient's." If there is at least one patient verb within a certain neighborhood, there is evidence that the patient is involved in the decision-making of their own health to some extent. In addition, the occurrence of "relationship to patient:" followed by "patient," "pt," or "self" in the note also constitutes evidence of home self-care.

XSL•FO
RenderX

Potential Reasons for Misclassification

The main reasons for algorithm misclassifications using annotated medical notes as our gold standard will be discussed and summarized.

Generalizability Using Other Medical Notes

To test the generalizability of our algorithm in other notes, we examined what percentage of the data dictionary features can be found in other medical notes. None of these notes were annotated. The findings provided some preliminary data for the next step, which is using other medical notes to make the algorithm more generalizable.

Ethics Approval

This research is approved by Michigan Medicine's Institutional Review Board (HUM00129193).

Results

We used R package version 3.6.3 (The R Foundation) to develop and test our NLP algorithm. Figure 1 presents the schematic flow diagram of our sampling. Out of the 304,186 available telephone encounter notes for our patient cohort, we annotated 749 notes for training and 227 notes for testing.

Table 1 displays some of the characteristics of our patient cohort (n=223). The average age was 78 (SD 10.94) years, with the majority being female (n=128, 57%). About 79% (n=176) were of White and 15% (n=33) of Black racial backgrounds. The average length of stay in the hospital was 6.78 (SD 6.54) days. Approximately 24% (n=54) of our patient population were readmitted or died within 30 days after discharge from the hospital.

Table 2 shows the results of our rule-based algorithm in our training and test sets. Among our features of interest (residing at home, residing at an institution, having a formal caregiver, and having an informal caregiver), identifying an informal caregiver was the most reliable feature. The result from our test set indicates high levels of accuracy and reliability for identifying an informal caregiver (F_1=0.942, accuracy=0.947, sensitivity=0.970, and specificity=0.928). Identifying whether the patient lives at an institution was the least reliable measure, with the algorithm being prone to false positives (F_1=0.638, accuracy=0.899, sensitivity=0.512, and specificity=0.978). The overall accuracy level for all 4 features in training and test sets were 0.858 and 0.655, respectively.

Table 3 summarizes the potential causes of misclassification, along with some examples and plausible explanations. The most common errors were related to (1) incomplete or misspelled names of the facilities; (2) past, uncertain, or undecided situations; (3) lack of specificity; (4) use of uncommon abbreviations; and (5) irregular use of templates.

To examine the generalizability of our algorithm using other medical notes, we measured the percentage of the features defined in our data dictionary in 5 different types of medical notes (patient care conference, pharmacy, psychiatric ED clinician, social work, and student) for our patient cohort (Table 4). The results indicate the highest level of generalizability for the informal caregiver. For example, 83% (n=1768) and 76% (n=595) of "patient care conference" and "social work" notes included information about informal caregivers. On the other hand, about 69% (n=333) of "pharmacy" notes had extractable information about a formal caregiver. This information can be used in future work to examine other types of medical notes.

XSL•FO
RenderX

Table 1. Descriptive characteristics of individuals included in training and test sets (N=223[a]).

Characteristics	Values
Age at the time of hospital admission, mean (SD)	77.96 (10.94)
Sex, n (%)	
Female	128 (57.4)
Male	95 (42.6)
Race or ethnicity, n (%)	
White	176 (78.9)
Black	33 (14.8)
Hispanic	3 (1.4)
Others	11 (4.9)
Length of stay in hospital, mean (SD)	6.78 (6.54)
Payor, n (%)	
Medicare+private	103 (46.2)
Medicare+Medicaid	34 (15.3)
Medicare only	53 (23.8)
Private only	6 (2.7)
Others or missing	27 (12.1)
Readmitted or died within 30 days after hospital discharge, n (%)	54 (24.2)

[a]Number of unique individuals in the sample. Each person has one or more "Telephone Encounter" medical notes.

Table 2. Model performance summary for training and test sets.

Model	Training (N=749)				Test (N=227)			
	Place of residence		Caregiver		Place of residence		Caregiver	
	Home	Institution	Formal	Informal	Home	Institution	Formal	Informal
F_1[a]	0.942	0.675	0.746	0.951	0.873	0.638	0.640	0.942
Accuracy[b]	0.923	0.964	0.923	0.964	0.837	0.899	0.841	0.947
Sensitivity[c]	0.947	0.609	0.680	0.971	0.870	0.512	0.571	0.970
Specificity[d]	0.875	0.987	0.971	0.960	0.778	0.978	0.930	0.928

[a]F_1-score: the predictive power of an algorithm as the harmonic mean of precision and recall. F_1-score ranges between 0 and 1, and the closer it is to 1, the better. F_1-score=2 * (precision*recall) / (precision + recall).

[b]Number of observations, both positive and negative, correctly classified. Accuracy = (true positive + true negative) / (true positive + false positive + true negative + false negative).

[c]Ability of the model to predict a true positive of each category.

[d]Ability of the model to predict a true negative of each category.

XSL•FO
RenderX

Table 3. Potential causes of misclassification with explanation and examples.

Cause of error	Example	Explanation
Incomplete or misspelled names	• "Pt stated that the nurse from Residential had difficulty drawing her blood." • "Medications are managed by staff at Gilbert House." • "Hartland of Ann Arbor"	• Residential is short for "Residential Home Health." If we add only "Residential" to our data dictionary, the false positive would increase. • Gilbert House is not in the dictionary. Formal name is Gilbert Residence. • Hartland is not in the dictionary. Formal name is Heartland Health Care Center.
Past, uncertain, or undecided situations	• Will also need "in Home Care" order. • "He shares that he has explored home health agencies (found them to be not suitable to what he is seeking)." • "I love that her long-term goal is already established, and Glacier Hills is her final choice." • "Will have a visit nurse in the near future (will be at sister's house)."	• "Home care" picked up by NLP as formal=1. • Falsely picked up "home health" as formal=1. • Falsely picked up institution=1 and formal=1. • Falsely picked up visiting nurse as formal=1.
Lack of specificity	• "Spoke with Donna who was caring for Mr. xxx." • "Ellen manages medications using monthly organizer." • "Calling from rehab facility and has some questions regarding wound care."	• It is not clear whether "Donna" is a formal or informal caregiver. The algorithm picked up formal=1. • Algorithm missed Ellen as a formal caregiver (formal=0). • In some cases, patient stays in the rehab facilities (institution=1 and formal=1), and in some cases, patient stays at home and goes to the rehab facility (institution=0 and formal=0). Due to this ambiguity, we did not include rehab facility in the dictionary.
Uncommon abbreviations	• "pt's dtr" • "her dau is working during the day."	• dtr and dau are short for daughter. They were not listed in the dictionary.

Table 4. Results of the natural language processing caregiver algorithm in other medical notes (the results show what percentage of the data dictionary features can be found in other medical notes).

Note type	Count, n	Overall, n (%)[a]	Resides at home, n (%)	Resides in an institution, n (%)	Formal caregiver, n (%)	Informal caregiver, n (%)
Telephone encounter	2000	1744 (87.2)	1326 (66.3)	426 (21.3)	704 (35.2)	1612 (80.6)
Patient care conference	2130	1825 (85.7)	1442 (67.7)	481 (22.6)	688 (32.3)	1768 (83.0)
Pharmacy	483	411 (85.0)	128 (26.4)	320 (66.2)	333 (68.9)	140 (29.0)
Psychiatric ED[b] clinician	488	351 (71.9)	394 (80.7)	41 (8.4)	55 (11.3)	345 (70.7)
Social work	783	621 (79.3)	612 (78.2)	147 (18.8)	212 (27.1)	593 (75.7)
Student	1201	921 (76.7)	873 (72.7)	160 (13.3)	240 (20.0)	852 (70.9)

[a]The overall percentage represents the proportion that at least one of the features in our data dictionary was used in the listed medical notes, while the feature-specific percentage indicates the proportion of notes containing information regarding the specific outcome. This was done to test the generalizability of the algorithm in other medical notes for future work.

[b]ED: emergency department.

Discussion

Principal Findings

This project was the first to assess whether medical notes can be used to identify caregiver availability and place of residence. We used a rule-based NLP algorithm on a subset of telephone encounter notes recorded between 2016 and 2019 for patients diagnosed with dementia to determine caregiver availability (formal and informal) and place of residence (home or institution). Our algorithm reliably identified the availability of an informal caregiver (F_1-score=0.94), moderately identified home as a place of residence (F_1-score=0.87), and poorly identified if the patient lives in an institution (F_1-score=0.64) or has a formal caregiver (F_1-score=0.64).

Comparison With Prior Work

Hospitals and health systems have made, and continue to make, substantial investments in their EMR systems. Although a systematic collection of salient medical and social data remains a work in progress, successful efforts using NLP algorithm have enabled efficient mining of rich free-text medical notes for various risk assessment or decision-making tools aimed at

XSL•FO
RenderX

reducing the occurrences of adverse health events and wasteful spending [22-24]. Our study aligns with this work to identify caregiver availability for patients whose well-being depends on caregivers.

For many older patients, especially people with cognitive decline or disability, information on caregiver availability has numerous applications. For instance, a recent work by Choi et al [9] reveals that among people with dementia or disability, those with a greater number of informal caregivers (ie, family members or close friends) are less likely to be institutionalized. Availability of caregiver information in medical settings may inform a better care transition (ie, discharge planning from the hospital), care use (ie, institution vs home), and care costs [25-29]. Through care coordination with caregivers, patients may experience better adherence to follow-up appointments and more effectively follow a prescribed diet and medication regimen. Moreover, patients who need a caregiver but have little or no family support can be identified by social workers or care coordinators to proactively navigate the use of formal care (ie, nursing home or paid home care) [29-32].

For this study, we used telephone encounter notes, initiated based on phone conversations with patients or their caregivers. Most of these notes are written by nurses based on their conversations with patients or family members of the patient at different time points. Perhaps because telephone encounter notes were based on direct conversations with patients or family members (or other informal caregivers), the algorithm was highly accurate in identifying informal caregivers. Furthermore, since usually informal caregivers are close family members, we had a better data dictionary to identify them in text. On the contrary, considering the vast number of places that offer a range of services from adult day care centers to independent living, perhaps we overfitted the model in training set. Hence, there was a drop in accuracy for the other 3 variables in the test set. More work is needed to detect short- and long-term residential places or paid caregiver organizations or agencies.

Furthermore, in many cases, it was hard to manually—through human interpretation—decipher the notes. Medical notes either have no standard template or the existing templates were not standardized or used irregularly. Various health care professionals (residents, physicians, nurses, social workers, etc) with limited resources and under time pressure write these notes. Hence, nonstandardized abbreviations (ie, “dau” for “daughter”), spelling errors, and incorrect and uncommon names are used regularly. Many of these issues cannot be addressed using off-the-shelf packages or programs. By contrast, although not generalizable, the rule-based NLP algorithm served as a proof of concept for addressing many institution-specific terminologies. we plan to address many of the following limitations in our future work.

Limitations, Strengths, and Future Work

Our study had a few noteworthy limitations. First, medical notes are based on unstructured text. We found large variations in the amount and type of information provided [31,32]. We used telephone encounter notes because, based on our examination of more than 60 different medical notes created within our institution, they provided the most relevant information regarding caregivers. We had, however, reasonable results detecting at least some elements of our data dictionary in other notes. In the future, we plan to make our algorithm generalizable by training and validating it using other medical notes and data from other health care centers. Second, manual annotation of the notes is resource intensive. Thus, our sample size was relatively small, which we plan to expand in the future. We will also explore the use of more sophisticated and unsupervised machine learning algorithms. Third, to make the algorithm more straightforward, we did not distinguish between a lack of objective and negative evidence. Thus, if there was no evidence about a caregiver or place of residence, we marked that outcome as zero. In our future work, we plan to make the algorithm more granular by identifying how many of the notes had (1) a positive indicator, (2) a negative indicator, and (3) no indicator. Further, to identify whether the patient lived in an institution, we used a list of skilled nursing facilities provided by the Centers for Medicare and Medicaid Services. There are, however, many unlisted independent living centers, adult day care centers, and other facilities designed to provide various services (residential and otherwise). It is challenging to include a comprehensive list of these facilities and their services. Having a reliable national directory of these facilities would help improve the model’s accuracy in determining whether a patient lives in a facility or is the recipient of paid or formal services. Finally, in this exploratory work, we only examined the binary availability of caregivers. Our future work will be focused on more critical information such as the caregiver’s proximity to the patient, the days and times of availability, the caregiver’s relation with the patient, and their capacity to help.

Conclusion

In this study, we used a rule-based approach to train, test, and develop an NLP algorithm using telephone encounter notes from our institution to identify whether a patient has a formal and informal caregiver and whether the patient resides at home or in an institution at each point in time. Our validated test results show high levels of accuracy and reliability, particularly in identifying whether a patient has an informal caregiver. This information is critical for vulnerable patient populations such as those with dementia. Our algorithm can be used as a stand-alone module or in conjunction with other tools to identify caregiver availability among high-risk patient populations. Future work will focus on making the algorithm more granular and generalizable so it can be used at other institutions.

Acknowledgments

We would like to thank Victoria Bristley, a registered nurse and clinical informaticist, for her help in annotating the medical notes.

XSL•FO
RenderX

The research reported in this publication was supported by National Institute on Aging of the National Institutes of Health (P30AG066582 and K01AG06361) and the Alzheimer's Association Research Grant (AARG-NTF-20-685960). Sponsors had no role in research design, data collection, data analysis, and research findings.

Data Availability

Since medical notes contain identifiable patient information, data sharing is not applicable. The data dictionary is available in the web-based appendix. The natural language processing algorithm is also available [33].

Authors' Contributions

EM and JB conducted the design and development of the research strategy. WW, CN, and EM were responsible for algorithm development and data validation. All authors performed the drafting and revising of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental Tables 1-3.
[DOCX File , 41 KB - aging_v5i3e40241_app1.docx]

References

1. Collier R. Electronic health records contributing to physician burnout. CMAJ 2017 Nov 13;189(45):E1405-E1406 [FREE Full text] [doi: 10.1503/cmaj.109-5522] [Medline: 29133547]
2. Moy A, Schwartz J, Chen R, Sadri S, Lucas E, Cato K, et al. Measurement of clinical documentation burden among physicians and nurses using electronic health records: a scoping review. J Am Med Inform Assoc 2021 Apr 23;28(5):998-1008 [FREE Full text] [doi: 10.1093/jamia/ocaa325] [Medline: 33434273]
3. Datta S, Bernstam EV, Roberts K. A frame semantic overview of NLP-based information extraction for cancer-related EHR notes. J Biomed Inform 2019 Dec;100:103301 [FREE Full text] [doi: 10.1016/j.jbi.2019.103301] [Medline: 31589927]
4. Castro VM, Minnier J, Murphy SN, Kohane I, Churchill SE, Gainer V, International Cohort Collection for Bipolar Disorder Consortium. Validation of electronic health record phenotyping of bipolar disorder cases and controls. Am J Psychiatry 2015 Apr;172(4):363-372 [FREE Full text] [doi: 10.1176/appi.ajp.2014.14030423] [Medline: 25827034]
5. Rumshisky A, Ghassemi M, Naumann T, Szolovits P, Castro VM, McCoy TH, et al. Predicting early psychiatric readmission with natural language processing of narrative discharge summaries. Transl Psychiatry 2016 Oct 18;6(10):e921-e921 [FREE Full text] [doi: 10.1038/tp.2015.182] [Medline: 27754482]
6. Cook BL, Progovac AM, Chen P, Mullin B, Hou S, Baca-Garcia E. Novel use of natural language processing (NLP) to predict suicidal ideation and psychiatric symptoms in a text-based mental health intervention in Madrid. Comput Math Methods Med 2016;2016:8708434-8708438 [FREE Full text] [doi: 10.1155/2016/8708434] [Medline: 27752278]
7. Patra B, Sharma M, Vekaria V, Adekkanattu P, Patterson O, Glicksberg B, et al. Extracting social determinants of health from electronic health records using natural language processing: a systematic review. J Am Med Inform Assoc 2021 Nov 25;28(12):2716-2727 [FREE Full text] [doi: 10.1093/jamia/ocab170] [Medline: 34613399]
8. Alami H, Lehoux P, Gagnon M, Fortin J, Fleet R, Ag Ahmed MA. Rethinking the electronic health record through the quadruple aim: time to align its value with the health system. BMC Med Inform Decis Mak 2020 Feb 17;20(1):32 [FREE Full text] [doi: 10.1186/s12911-020-1048-9] [Medline: 32066432]
9. Choi H, Heisler M, Norton EC, Langa KM, Cho T, Connell CM. Family care availability and implications for informal and formal care used by adults with dementia in the US. Health Aff (Millwood) 2021 Sep 01;40(9):1359-1367 [FREE Full text] [doi: 10.1377/hlthaff.2021.00280] [Medline: 34495713]
10. Hanauer DA, Mei Q, Vydiswaran VGV, Singh K, Landis-Lewis Z, Weng C. Complexities, variations, and errors of numbering within clinical notes: the potential impact on information extraction and cohort-identification. BMC Med Inform Decis Mak 2019 Apr 04;19(Suppl 3):75 [FREE Full text] [doi: 10.1186/s12911-019-0784-1] [Medline: 30944012]
11. Vydiswaran VGV, Strayhorn A, Zhao X, Robinson P, Agarwal M, Bagazinski E, et al. Hybrid bag of approaches to characterize selection criteria for cohort identification. J Am Med Inform Assoc 2019 Nov 01;26(11):1172-1180 [FREE Full text] [doi: 10.1093/jamia/ocz079] [Medline: 31197354]
12. Vydiswaran VGV, Zhang Y, Wang Y, Xu H. Special issue of BMC medical informatics and decision making on health natural language processing. BMC Med Inform Decis Mak 2019 Apr 04;19(Suppl 3):76 [FREE Full text] [doi: 10.1186/s12911-019-0777-0] [Medline: 30943961]
13. Vydiswaran V, Zhao X, Yu D. Data science and natural language processing to extract information in clinical domain. 2022 Presented at: 5th Joint International Conference on Data Science & Management of Data (9th ACM IKDD CODS and 27th COMAD); January 8-10, 2022; Bangalore, India p. 352-353. [doi: 10.1145/3493700.3493773]

XSL•FO
RenderX

14. Bynum JP, Dorr DA, Lima J, McCarthy EP, McCreedy E, Platt R, et al. Using healthcare data in embedded pragmatic clinical trials among people living with dementia and their caregivers: state of the art. J Am Geriatr Soc 2020 Jul 26;68 Suppl 2(S2):S49-S54 [FREE Full text] [doi: 10.1111/jgs.16617] [Medline: 32589274]
15. Barnes DE, Yaffe K. The projected effect of risk factor reduction on Alzheimer's disease prevalence. Lancet Neurol 2011 Sep;10(9):819-828 [FREE Full text] [doi: 10.1016/S1474-4422(11)70072-2] [Medline: 21775213]
16. Sinvani L, Warner-Cohen J, Strunk A, Halbert T, Harisingani R, Mulvany C, et al. A multicomponent model to improve hospital care of older adults with cognitive impairment: a propensity score-matched analysis. J Am Geriatr Soc 2018 Sep 11;66(9):1700-1707. [doi: 10.1111/jgs.15452] [Medline: 30098015]
17. Zheng S, Hanchate A, Shwartz M. One-year costs of medical admissions with and without a 30-day readmission and enhanced risk adjustment. BMC Health Serv Res 2019 Mar 12;19(1):155 [FREE Full text] [doi: 10.1186/s12913-019-3983-7] [Medline: 30866904]
18. Zhu CW, Cosentino S, Ornstein K, Gu Y, Andrews H, Stern Y. Use and cost of hospitalization in dementia: longitudinal results from a community-based study. Int J Geriatr Psychiatry 2014 Oct 29;30(8):833-841. [doi: 10.1002/gps.4222]
19. Mahmoudi E, Kamdar N, Kim N, Gonzales G, Singh K, Waljee AK. Use of electronic medical records in development and validation of risk prediction models of hospital readmission: systematic review. BMJ 2020 Apr 08;369:m958. [doi: 10.1136/bmj.m958] [Medline: 32269037]
20. Yacouby R, Axman D. Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models. 2020 Presented at: Proceedings of the first workshop on evaluation and comparison of NLP systems; November 20, 2020; Online p. 79-91. [doi: 10.18653/v1/2020.eval4nlp-1.9]
21. Nursing homes including rehab services datasets. Centers for Medicare and Medicaid Services. URL: https://data.cms.gov/provider-data/search?theme=Nursing%20homes%20including%20rehab%20services [accessed 2021-08-30]
22. Sung S, Chen K, Wu DP, Hung L, Su Y, Hu Y. Applying natural language processing techniques to develop a task-specific EMR interface for timely stroke thrombolysis: A feasibility study. Int J Med Inform 2018 Apr;112:149-157. [doi: 10.1016/j.ijmedinf.2018.02.005] [Medline: 29500013]
23. Ridgway JP, Uvin A, Schmitt J, Oliwa T, Almirol E, Devlin S, et al. Natural language processing of clinical notes to identify mental illness and substance use among people living with HIV: retrospective cohort study. JMIR Med Inform 2021 Mar 10;9(3):e23456 [FREE Full text] [doi: 10.2196/23456] [Medline: 33688848]
24. Wang J, Deng H, Liu B, Hu A, Liang J, Fan L, et al. Systematic evaluation of research progress on natural language processing in medicine over the past 20 years: bibliometric study on PubMed. J Med Internet Res 2020 Jan 23;22(1):e16816 [FREE Full text] [doi: 10.2196/16816] [Medline: 32012074]
25. Kocher RP, Adashi EY. Hospital readmissions and the Affordable Care Act: paying for coordinated quality care. JAMA 2011 Oct 26;306(16):1794-1795. [doi: 10.1001/jama.2011.1561] [Medline: 22028355]
26. Kripalani S, Theobald CN, Anctil B, Vasilevskis EE. Reducing hospital readmission rates: current strategies and future directions. Annu Rev Med 2014 Jan 14;65(1):471-485 [FREE Full text] [doi: 10.1146/annurev-med-022613-090415] [Medline: 24160939]
27. Joynt KE, Jha AK. Thirty-day readmissions — Truth and consequences. N Engl J Med 2012 Apr 12;366(15):1366-1369. [doi: 10.1056/nejmp1201598]
28. Kistler CE, Beeber AS, Winzelberg GS, Gabriel SL, Wretman CJ, Hanson LC. Evaluation of a training toolkit to improve clinicians' skills for dementia advance care planning. J Palliat Med 2021 Aug 01;24(8):1183-1190. [doi: 10.1089/jpm.2020.0638] [Medline: 33400605]
29. Wennberg A, Dye C, Streetman-Loy B, Pham H. Alzheimer's patient familial caregivers: a review of burden and interventions. Health Social Work 2015 Aug 25;40(4):e162-e169. [doi: 10.1093/hsw/hlv062]
30. Bressan V, Visintini C, Palese A. What do family caregivers of people with dementia need? A mixed-method systematic review. Health Soc Care Community 2020 Nov 16;28(6):1942-1960. [doi: 10.1111/hsc.13048] [Medline: 32542963]
31. Jagannatha A, Liu F, Liu W, Yu H. Overview of the first natural language processing challenge for extracting medication, indication, and adverse drug events from electronic health record notes (MADE 1.0). Drug Saf 2019 Jan 16;42(1):99-111 [FREE Full text] [doi: 10.1007/s40264-018-0762-z] [Medline: 30649735]
32. Perera S, Sheth A, Thirunarayan K, Nair S, Shah N. Challenges in understanding clinical notes: why nlp engines fall short and where background knowledge can help. 2013 Presented at: Proceedings of the 2013 international workshop on Data management & analytics for healthcare; November 1, 2013; New York, US p. 21-26. [doi: 10.1145/2512410.2512427]
33. ADRD-NLP. GitHub. URL: https://github.com/Mahmoudi-Lab/ADRD-NLP [accessed 2022-09-15]

Abbreviations

EMR: electronic medical record
MM: Michigan Medicine
NLP: natural language processing

XSL•FO
RenderX

Edited by T Leung; submitted 12.06.22; peer-reviewed by L Xia, J Kuriakose; comments to author 10.07.22; revised version received 28.07.22; accepted 16.08.22; published 22.09.22.

Please cite as:
Mahmoudi E, Wu W, Najarian C, Aikens J, Bynum J, Vydiswaran VGV
Identifying Caregiver Availability Using Medical Notes With Rule-Based Natural Language Processing: Retrospective Cohort Study
JMIR Aging 2022;5(3):e40241
URL: https://aging.jmir.org/2022/3/e40241
doi:10.2196/40241
PMID:35998328

©Elham Mahmoudi, Wenbo Wu, Cyrus Najarian, James Aikens, Julie Bynum, V G Vinod Vydiswaran. Originally published in JMIR Aging (https://aging.jmir.org), 22.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Social Media–Based Intervention for Chinese American Caregivers of Persons With Dementia: Protocol Development

Y Alicia Hong[1], PhD; Kang Shen[1], MS; Huixing Kate Lu[2], LCSW-C; Hsiaoyin Chen[2], LMSW; Yang Gong[3], MD, PhD; Van Ta Park[4], PhD; Hae-Ra Han[5], PhD

[1]Department of Health Administration and Policy, College of Public Health, George Mason University, Fairfax, VA, United States

[2]Chinese Culture and Community Service Center, Inc, Gaithersburg, MD, United States

[3]School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, TX, United States

[4]Department of Community Health Systems, School of Nursing, University of California, San Francisco, San Francisco, CA, United States

[5]School of Nursing, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
Y Alicia Hong, PhD
Department of Health Administration and Policy
College of Public Health
George Mason University
4400 University Dr. MS 1j3
Fairfax, VA, 22030
United States
Phone: 1 703 993 1929
Email: yhong22@gmu.edu

Abstract

Background: Racial/ethnic minority and immigrant caregivers of persons with dementia experience high rates of psychosocial stress and adverse health outcomes. Few culturally tailored mobile health (mHealth) programs were designed for these vulnerable populations.

Objective: This study reports the development of a culturally tailored mHealth program called Wellness Enhancement for Caregivers (WECARE) to improve caregiving skills, reduce distress, and improve the psychosocial well-being of Chinese American family caregivers of persons with dementia.

Methods: Community-based user-centered design principles were applied in the program development. First, the structure and curriculum of the WECARE program were crafted based on existing evidence-based interventions for caregivers with input from 4 experts. Second, through working closely with 8 stakeholders, we culturally adapted evidence-based programs into multimedia program components. Lastly, 5 target users tested the initial WECARE program; their experience and feedback were used to further refine the program.

Results: The resulting WECARE is a 7-week mHealth program delivered via WeChat, a social media app highly popular in Chinese Americans. By subscribing to the official WECARE account, users can receive 6 interactive multimedia articles pushed to their WeChat accounts each week for 7 weeks. The 7 major themes include (1) facts of dementia and caregiving; (2) the enhancement of caregiving skills; (3) effective communication with health care providers, care partners, and family members; (4) problem-solving skills for caregiving stress management; (5) stress reduction and depression prevention; (6) the practice of self-care and health behaviors; and (7) social support and available resources. Users also have the option of joining group chats for peer support. The WECARE program also includes a back-end database that manages intervention delivery and tracks user engagement.

Conclusions: The WECARE program represents one of the first culturally tailored social media–based interventions for Chinese American caregivers of persons with dementia. It demonstrates the use of community-based user-centered design principles in developing an mHealth intervention program in underserved communities. We call for more cultural adaptation and development of mHealth interventions for immigrant and racial/ethnic minority caregivers of persons with dementia.

(JMIR Aging 2022;5(3):e40171) doi:10.2196/40171

XSL•FO
RenderX

KEYWORDS

Alzheimer disease; dementia; caregivers; Chinese Americans; mHealth intervention; mobile health; WeChat; social media; aging

Introduction

Currently, more than 6 million Americans aged ≥65 years are living with Alzheimer disease or related dementias (ADRD) [1]. More than 11 million family caregivers of persons with dementia provide an estimated 15.3 billion hours of unpaid care valued at US $255.7 billion a year [1]. Family caregivers of people with dementia have high rates of emotional distress and negative health outcomes [2]. For example, 59% of caregivers of persons with dementia reported high rates of emotional stress [3], 40% reported depression, and 44% reported anxiety [4-6]. Family caregivers of persons with dementia also reported higher levels of physical stress [3,7], lower quality of sleep [8-10], and lower quality of life [11-13]. Some caregivers developed chronic conditions including impaired immune functions, hypertension, and coronary health diseases [14]. As the US population is aging and the number of persons with dementia is expected to reach 13 million by 2050, the burden on family caregivers and their psychosocial well-being requires more public health attention [1].

Asian Americans represent the fastest-growing racial group in the United States; they accounted for 7% of the total US population in 2020 and are projected to reach 12% in 2050. Chinese Americans represent nearly a quarter (23%) of the Asian American population [15]. The literature on Chinese American caregivers of persons with dementia is limited and mostly descriptive. Available literature indicates that cultural values of family harmony and the practice of filial piety permeate all aspects of the Chinese American caregiving process, including their appraisal of stress and coping strategies [16]. Caring for older family members is not only a sign of love and pride but also a moral obligation [17,18]. Compared to their White counterparts, Chinese Americans are more likely to live in multigenerational households [19]. Chinese caregivers are often providers of young children and older adults while also being engaged in the formal labor force themselves [19,20]. Although caregiving can strengthen attachment and emotional bonds, it also leads to tensions and feelings of being neglected [18]. These frustrations are amplified in immigrant families where cultural differences often clash with generational gaps [20].

Most of dementia care is received at home where family caregivers play a central role but often lack the knowledge and skills to perform caregiving duties to meet the needs of persons with dementia. Chinese caregivers tend to keep problems within the family and do not seek external help because of the stigma associated with dementia or cognitive impairment [17]. The isolation and challenge are exacerbated by their minority and immigrant status, and those without English proficiency are further marginalized [21]. Most Chinese caregivers have limited knowledge and use of formal care and support services; they are also disconnected from “mainstream” dementia support groups due to language and cultural barriers [22]. Therefore, they endure higher levels of stress, mental disorders, and chronic conditions [23,24].

To date, few interventions have been conducted to serve the needs of Chinese American caregivers of persons with dementia. Specifically, in a recent systematic review, only 2 interventions targeting Chinese American caregivers were identified [25]. One was a home-based in-person behavioral management program [26], and the other was a DVD-based psychoeducation skills training program [27]. Both interventions, developed by Gallagher-Thompson and colleagues in early 2000s, cannot meet needs of a large population of Chinese persons with dementia, as face-to-face intervention is resource-intensive, and DVDs are not commonly present in most American households anymore. The need for accessible and scalable mobile health (mHealth) interventions has become more salient during the COVID-19 pandemic, when in-home visits or in-person services are prohibitive, and many people adopted telehealth or mHealth programs. However, the existing mHealth interventions for caregivers of persons with dementia were mainly designed for people of higher socioeconomic status with limited programs for racial/ethnic minority or immigrant populations. Adapting evidence-based interventions for digital delivery in underserved communities remains a critical literature gap. As racial/ethnic minority and low-income populations are more likely to be smartphone-dependent for internet access (without a computer or other mobile devices) and rely on social media as a primary source of health information [28], an mHealth intervention delivered via a social media app used by the target population is a logic solution. Nevertheless, evidence or research in this regard is rather limited. For example, in a recent systematic review of mHealth interventions for dementia caregivers, none was delivered via a social media app [29].

Out of nearly 4 million Chinese Americans, more than 70% are foreign-born [30]. Over 3 million Americans speak Chinese, the third most spoken language, only after English and Spanish. As the Asian American population continues to grow, a culturally tailored and linguistically appropriate mHealth intervention has a tremendous potential to reach a large number of Chinese American caregivers of persons with dementia.

More than 95% of Chinese Americans own a smartphone [31]. Similar to other immigrants whose first language is not English, many Chinese Americans use a social media app of their native language. For example, WeChat, a highly popular social media app, has a 97% penetration rate in smartphone users in China and a 90% penetration rate in Chinese-speaking Chinese Americans [32]. As a social media app, WeChat has the functions of “moments” (sharing photos and stories with friends and receiving “likes” and feedback), texting, voice call, video call, private chat, group chat, location sharing, file transfer, and making/receiving payment. These built-in functions allow intervention developers to focus on the content, thus saving time and cost. It also enables easy adoption and long-term use, especially in low-income communities [33]. Thus, WeChat would be an efficient delivery channel for an mHealth intervention for Chinese Americans caregivers.

XSL•FO
RenderX

To address the literature gap and public health needs, our team developed a WeChat-based intervention called "Wellness Enhancement for Caregivers" (WECARE) for Chinese American family caregivers of persons with dementia. This paper reports the process of developing the WECARE program and the design features of this social media–based intervention.

Methods

Overview of Study Design

The study was conducted from September 2021 to April 2022 with the goal of developing a culturally tailored WeChat-based intervention to improve the psychosocial well-being of Chinese American family caregivers of persons with dementia. The development of the WECARE program consisted of 3 steps. First, with experts' input, we designed the structure and curriculum of the WECARE program based on existing evidence-based behavioral interventions for persons with dementia. Second, using a community-engaged user-centered design, we developed the multimedia components of the WECARE program that used the built-in functions of the WeChat app; we also developed a back-end database to manage intervention delivery and track user engagement. Finally, the complete WECARE program was refined after a beta test in target users.

Ethics Approval

The study protocol was approved by the Institutional Review Board of the George Mason University (IRB1849712). Informed consent was obtained from all participants prior to data collection. Given that this study was focused on protocol development, we did not collect participants' personal information including their demographic data.

Step 1: Structural Design Based on Experts' Input and Evidence-Based Programs

We first conducted interviews with 4 experts in the fields of ADRD, caregiving, cultural adaption, and mHealth intervention development, with 1 expert from each field. We sought input from the experts on (1) how to apply existing theories and evidence to enhance caregiving skills and reduce psychosocial stress among underserved family caregivers, (2) how to adapt evidence-based caregiving interventions for Chinese American caregivers, and (3) how to identify and prepare for potential barriers and facilitators during the intervention delivery.

The experts also suggested that the curriculum design be based on evidence-based interventions proven to be effective in minority and underserved caregivers of persons with dementia. Specifically, we used the Resources for Enhancing Alzheimer's Caregiver Health II for its major domains of the intervention [34], Building Better Care for its short courses and training materials [35], and the DVD program of Gallagher-Thompson and colleagues [27] for culturally relevant problems and solutions for Chinese American caregivers of persons with dementia [27].

Step 2: Iterative Design With Key Stakeholders

Following the principles of community-engaged user-centered design [36] while developing the program components, we worked closely with 8 stakeholders, including Chinese American family caregivers of persons with dementia (n=3), health care providers (n=3), and community leaders (n=2). Weekly meetings were organized to seek immediate input from stakeholders in terms of cultural appropriateness, ease of use, user engagement, and error reduction. The program components underwent the iterative process of being developed, reviewed, revised, tested, and refined. We also hired a software engineer to develop a back-end database to manage the WECARE delivery and user profiles. The database reflects the required functions identified in Step 1, including prescheduled automatic intervention delivery, user profile management, and user engagement tracking.

Step 3: Testing and Refinement

When the complete WECARE program was ready, we tested it among 5 target users. Of these 5 Chinese American family caregiver participants—1 man and 4 women—all had limited English proficiency and were recruited by our community partner. Participants were invited to a conference room and received the WECARE program on their WeChat accounts to test its navigation and the functions of the back-end database. During the test, participants were encouraged to "think out loud" and share their feedback while going through the program components. Participants were asked to check if all program components were delivered at prescheduled times and if all program components could be opened without problem. Meanwhile, research staff monitored all user activities on WECARE's back-end database. All interviews were conducted in Chinese or Mandarin and audio recorded, and detailed notes were taken while observing users' navigation behaviors. The research team discussed users' feedback, addressed the glitches reported in the test, and further refined the program.

Results

Program Curriculum and Components

WECARE is a 7-week program with each week focused on a theme and the final week for summary and additional resources. These themes include (1) facts of ADRD and caregiving; (2) enhancement of caregiving skills; (3) effective communication with providers, care partners, and family members; (4) problem-solving skills for caregiving stress management; (5) stress reduction and depression prevention; (6) practice of self-care and health behaviors; and (7) social support and available resources. The 7 major themes were derived from the evidence-based programs (see Step 2). The curriculum schedule is detailed in Multimedia Appendix 1. The sample screenshots of WECARE are illustrated in Figure 1. By subscribing to the official WECARE account on their own WeChat accounts, participants could receive interactive multimedia programs on their smartphone or tablet once a day, 6 times a week, for 7 weeks. Participants who did not open the program components within a week would be reminded to do.

XSL•FO
RenderX

Figure 1. Wellness Enhancement for Caregivers (WECARE) screenshots: (A) a multimedia article with audio recording on how to deal with the angry behaviors of persons with dementia; (B) a short video clip as a case study to explain angry behaviors of persons with dementia; and (C) push notifications and progress summary.



Cultural Adaptation

The content of the WECARE program was adapted for Chinese American culture based on input from our stakeholders, and the following changes were made. First, we ensured that the language translation reflected Chinese values and tradition. For example, following the tradition of respecting seniority, Chinese participants did not like the literal translation of “care recipient” or “care partner”; instead, they preferred to use “the elder” (“Lao Ren” in Chinese) to refer to the care partner, which we adopted in the WECARE program. Second, program components specific for our target users were added. For example, given that many Chinese caregivers have limited English proficiency, we added a section on how to communicate with health care providers with a practical checklist as well as a list of terminologies commonly used in medical encounters and dementia care. Additionally, as many Chinese caregivers have limited knowledge and use of formal care and support services [22], we added a section on local resources specific to the participant’s location, including health insurance, dementia care support, transportation assistance for medical care, and translation services. Third, we modified components to meet users’ demand and characteristics. For example, many stakeholders demanded “useful skill-building” content; we included videos demonstrating caregiving skills such as how to transfer, bath, feed, and clean the care partner. We also included many real-life cases to illustrate how to deal with difficult situations such as when a care partner has problematic behaviors in the public. Fourth, we adapted components to reflect cultural practice. For example, many Chinese caregivers live in multigenerational households and value “filial piety”; thus, we modified the component on how to communicate between family members, including communication on how to share responsibilities and how to discuss sensitive topics such as death. Additionally, as stress and depression are not commonly discussed in Chinese culture, we included a section to explain the importance of stress reduction and depression prevention from the perspective of family and love, citing real-life stories and demonstrating practical stress-reduction techniques.

Multimedia Features

The WECARE program consists of multimedia components to engage users and enhance understanding. Considering that many caregivers are older adults and many have lower levels of health literacy, each article is accompanied by an audio recording, so participants with vision impairment can listen to audio recordings for most of the WECARE content. We also included many short videos (3-5 minutes) adapted from other caregiver interventions with subtitles for illustration. Culturally relevant characters, storylines, and background music are embedded in all program components.

Social Networking and Social Support

As WECARE is delivered via the popular WeChat social media app, the built-in functions of social networking in WeChat were used to enhance social support among participants. For example, participants were invited to attend staff-moderated group meetings scheduled at week 3, week 5, and week 7 and welcome to share their personal experiences during the course of WECARE. Participants could also initiate their own “group chat”; in a group chat, they could “friend” any fellow caregiver participant for a private chat. Prior to the group meetings, participants agreed to the protocol that all participants in the group meetings address each other by first names only and that no personal information discussed in the meetings be shared with other people outside the group.

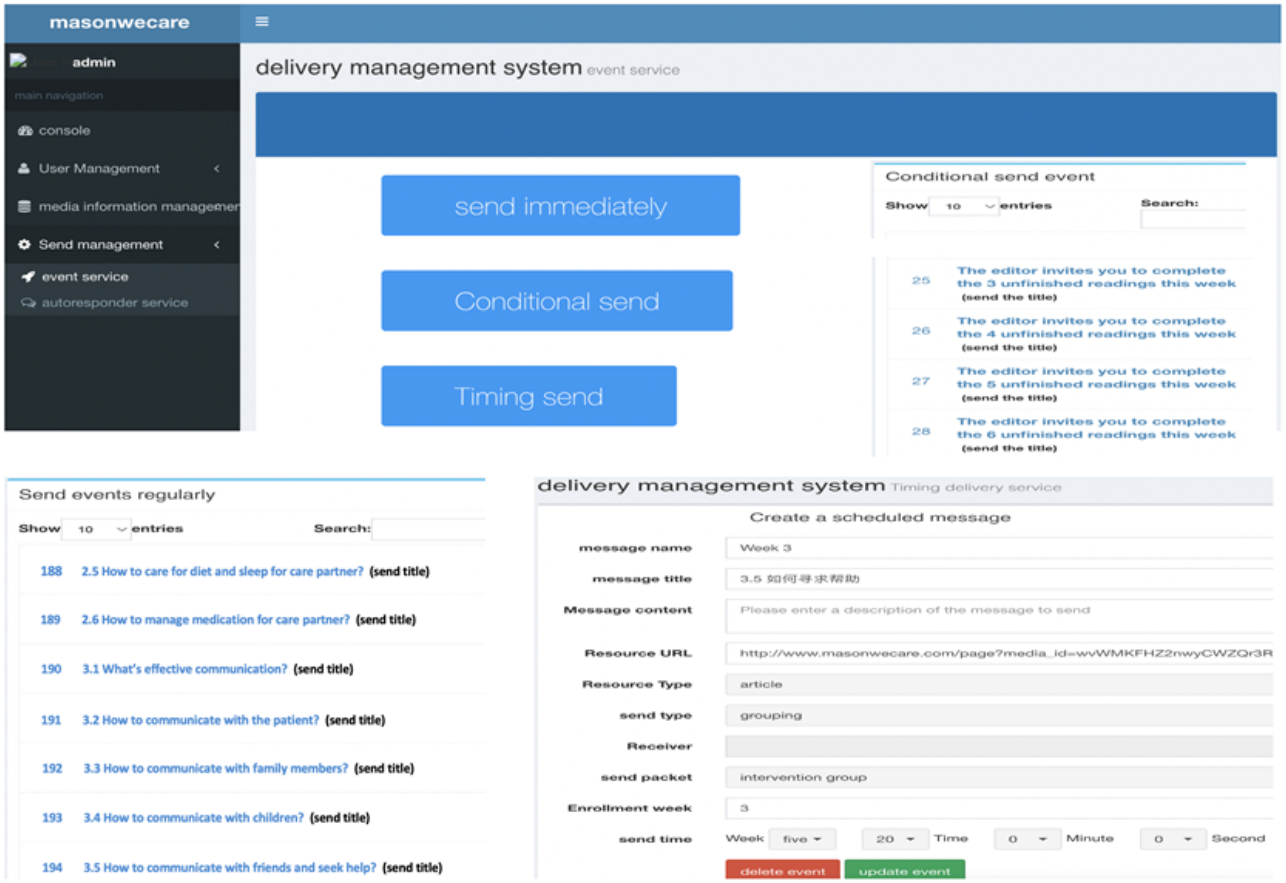
Back-end Database

Along with developing the front-end program components of WECARE, we also created an interactive web-based application

XSL•FO
RenderX

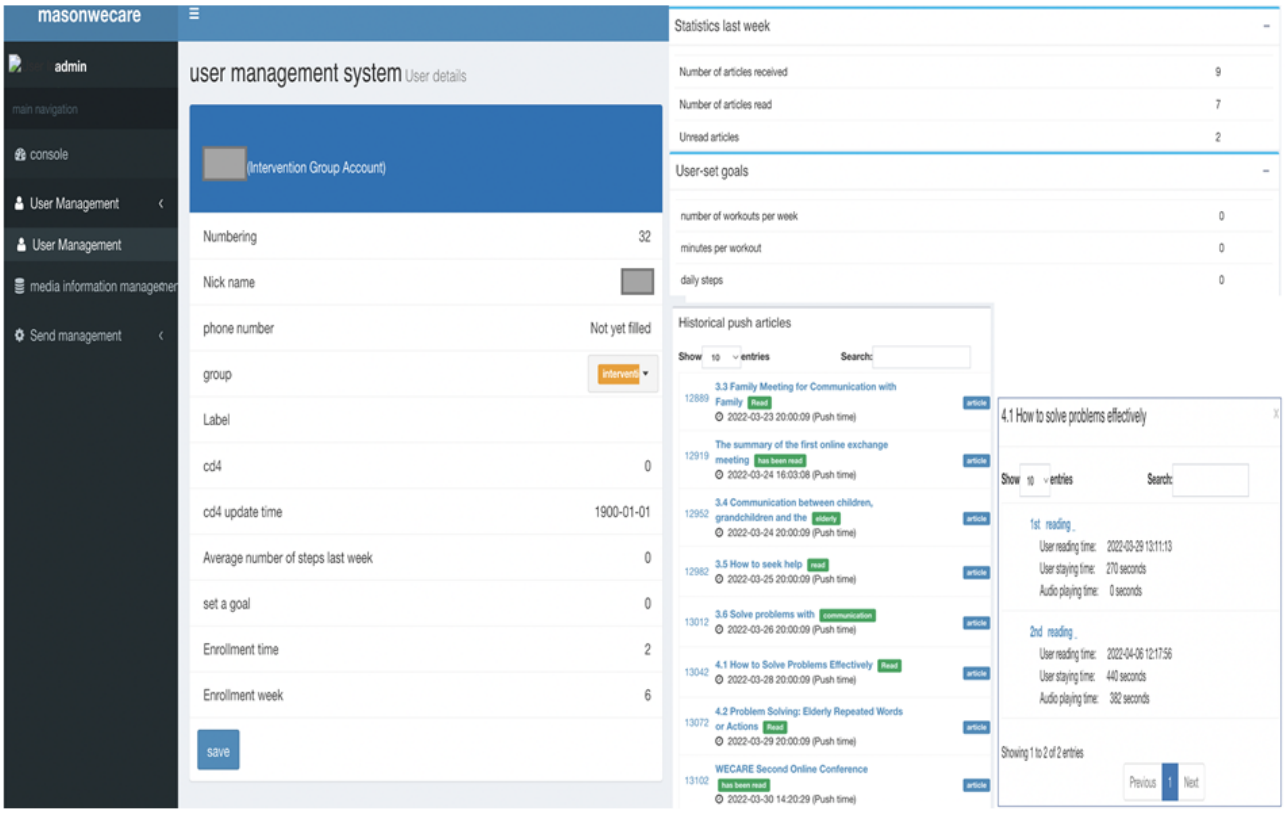
to serve as the back-end database to manage intervention delivery and monitor user engagement. This interactive website has 3 core functions: (1) storing information on every participant enrolled, including characteristic and user preferences; (2) pushing the program components based on user preferences and response (eg, if a participant does not open the WECARE program component within a week, a reminder message will be sent); and (3) tracking program receipt and responses. User engagement indicators tracked in the database include whether a program component is opened, how many times the program is opened or played, and how much time is spent on each component. Figure 2 illustrates the scheduled delivery system through which the components of the WECARE program can be sent to users at prescheduled time. Figure 3 illustrates the user management system that stores user information and tracks user activities.

Figure 2. Back-end database: scheduled delivery system.



XSL•FO
RenderX

Figure 3. Back-end database: user management system.



Discussion

Principal Findings

This study reports the process of developing WECARE, a culturally tailored social media–based intervention to enhance caregiving skills and reduce psychosocial distress in Chinese American caregivers of persons with dementia, and its key features. Community-based user-centered design principles were applied in the intervention development. We first crafted the structure of the WECARE curriculum based on existing evidence-based interventions and input from experts. Second, working closely with key stakeholders and through an iterative design-discuss-revise process, we developed culturally tailored multimedia program components at the frontend and an interactive database for intervention delivery and user profile management at the backend. Finally, we conducted a beta test of the complete WECARE program in target users and further refined it. The resulting WECARE is a 7-week mHealth program. Through subscribing to the official WECARE account, users could receive 6 multimedia articles pushed to their WeChat account each week for 7 weeks. Users could also use the built-in functions of WeChat for social networking. The back-end database automatically pushes program components with a preset schedule and tracks user activities on WECARE.

Strengths

The First mHealth Intervention for Chinese American Caregivers of Persons With Dementia

The WECARE program represents one of the first culturally tailored mHealth programs for Chinese American caregivers of persons with dementia. The health disparities experienced by minority and immigrant caregivers of persons with dementia have been exacerbated during the COVID-19 pandemic. There is an urgent need to adapt evidence-based interventions for wider dissemination in underserved populations. The cultural adaptation and digitalization of evidence-based interventions, such as developing the WECARE program as described in this paper, could be an effective approach to address the literature gap and public health needs.

Capitalizing on a Popular Social Media App for Intervention Delivery

When adapting and developing mHealth interventions, we need to consider the mobile use behaviors of the target population. The popular social media apps used by minority populations can serve as an efficient channel for intervention delivery [37]. The WECARE program is delivered via WeChat, a social media app with a high penetration rate in Chinese Americans. This program is the first time a social media app popularly used by minority populations is being used to deliver an intervention for minority caregivers of persons with dementia. Delivering mHealth interventions via such apps can save costs, increase accessibility, and enhance sustainability [38].

Cultural Adaptation for Target Users

Cultural sensitivity is critical for developing interventions for minority populations. A community-engaged user-centered design is an effective approach for the cultural adaptation of evidence-based programs. Using this approach, we made substantial changes in the WECARE program components in response to the input of our stakeholders and target users. The resulting program reflects the values, needs, and practices of Chinese American caregivers of persons with dementia.

XSL•FO
RenderX

Multimedia Features

The WECARE program features multimedia components of audio recording, short video clips, pictorial messages as well as automatic and interactive delivery based on user preferences and responses. The built-in functions in WeChat such as "private chat," "group chat," and "video chats" facilitate social networking and enhance social support. Literature suggests that mHealth programs with multimedia features and social networking functions are likely to engage target users, especially those with lower levels of health literacy [39]. The design and innovative features of WECARE will inform future designs of mHealth interventions for caregivers.

Back-end Database

The back-end database has the functions of user profile management, the automatic delivery of program components, and user activity tracking. These functions are similar to another WeChat-based intervention and enable potential scale-up and long-term follow-up [40]. The user activity tracking function also allows a future study to examine the relationship between user engagement and intervention effect [41].

Limitations

First, the beta test of the WECARE program was based on a small sample of 5 target users. The development was a rigorous process following community-engaged user-centered design principles [36] with input from experts and iterative discussion with key stakeholder. Further, according to Nielsen [42], 5 users are sufficient to elaborate usability. Second, the back-end database was designed specifically for WECARE delivery and requires a software engineer. We are seeking open-sourced solutions to lower costs and increase the potential of WECARE's adoption and scale-up. Third, the biggest limitation of the study is that we do not have data on the feasibility and effectiveness of WECARE as a complete program as well as user feedback on its innovative features and functions. A pilot study is ongoing, and the results will be available once the trial is complete.

Future Directions

In conclusion, this paper reports the development process and key features of the WECARE program—a culturally tailored, linguistically appropriate, and interactive social media–based interventions to improve caregiving skills and reduce psychosocial distress among Chinese American caregivers of persons with dementia. As the US population becomes older and more diverse, an urgent need exists for more culturally sensitive mHealth interventions for minority and immigrant caregivers of persons with dementia. We advocate for more research and practice of the cultural adaptation of evidence-based program for digital delivery, capitalizing on the widespread use of smartphones and highly popular social media apps, to meet the needs of racial/ethnic minority and immigrant patients with dementia and their caregivers.

Acknowledgments

The study was funded by the Virginia Center for Aging Alzheimer's and Related Diseases Research Award Fund.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Wellness Enhancement for Caregivers (WECARE) program schedule.
[PDF File (Adobe PDF File), 124 KB - aging_v5i3e40171_app1.pdf]

References

1. Alzheimer's Association. 2021 Alzheimer's disease facts and figures. Alzheimers Dement 2021 Mar 23;17(3):327-406. [doi: 10.1002/alz.12328] [Medline: 33756057]
2. Schulz R, Beach SR, Czaja SJ, Martire LM, Monin JK. Family caregiving for older adults. Annu Rev Psychol 2020 Jan 04;71(1):635-659 [FREE Full text] [doi: 10.1146/annurev-psych-010419-050754] [Medline: 31905111]
3. Mausbach BT, Patterson TL, Rabinowitz YG, Grant I, Schulz R. Depression and distress predict time to cardiovascular disease in dementia caregivers. Health Psychol 2007 Sep;26(5):539-544. [doi: 10.1037/0278-6133.26.5.539] [Medline: 17845105]
4. Sallim AB, Sayampanathan AA, Cuttilan A, Ho R. Prevalence of mental health disorders among caregivers of patients with Alzheimer disease. J Am Med Dir Assoc 2015 Dec;16(12):1034-1041. [doi: 10.1016/j.jamda.2015.09.007] [Medline: 26593303]
5. Vitaliano PP, Ustundag O, Borson S. Objective and subjective cognitive problems among caregivers and matched non-caregivers. Gerontologist 2017 Aug 01;57(4):637-647 [FREE Full text] [doi: 10.1093/geront/gnv690] [Medline: 26903233]
6. Dassel KB, Carr DC, Vitaliano P. Does caring for a spouse with dementia accelerate cognitive decline? findings from the health and retirement study. Gerontologist 2017 Apr 01;57(2):319-328. [doi: 10.1093/geront/gnv148] [Medline: 26582383]
7. Pinquart M, Sörensen S. Correlates of physical health of informal caregivers: a meta-analysis. J Gerontol B Psychol Sci Soc Sci 2007 Mar 01;62(2):P126-P137. [doi: 10.1093/geronb/62.2.p126] [Medline: 17379673]

XSL•FO
RenderX

8. von Känel R, Mausbach BT, Ancoli-Israel S, Mills PJ, Dimsdale JE, Patterson TL, et al. Positive affect and sleep in spousal Alzheimer caregivers: a longitudinal study. Behav Sleep Med 2014 Sep 03;12(5):358-372 [FREE Full text] [doi: 10.1080/15402002.2013.819470] [Medline: 24156281]
9. Peng HL, Chang YP. Sleep disturbance in family caregivers of individuals with dementia: a review of the literature. Perspect Psychiatr Care 2013 Apr 07;49(2):135-146. [doi: 10.1111/ppc.12005] [Medline: 23557457]
10. Gao C, Chapagain NY, Scullin MK. Sleep duration and sleep quality in caregivers of patients with dementia: a systematic review and meta-analysis. JAMA Netw Open 2019 Aug 02;2(8):e199891 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.9891] [Medline: 31441938]
11. Vitaliano PP, Zhang J, Scanlan JM. Is caregiving hazardous to one's physical health? a meta-analysis. Psychol Bull 2003 Nov;129(6):946-972. [doi: 10.1037/0033-2909.129.6.946] [Medline: 14599289]
12. Goren A, Montgomery W, Kahle-Wrobleski K, Nakamura T, Ueda K. Impact of caring for persons with Alzheimer's disease or dementia on caregivers' health outcomes: findings from a community based survey in Japan. BMC Geriatr 2016 Jun 10;16(1):122 [FREE Full text] [doi: 10.1186/s12877-016-0298-y] [Medline: 27287238]
13. Välimäki TH, Martikainen JA, Hongisto K, Väätäinen S, Sintonen H, Koivisto AM. Impact of Alzheimer's disease on the family caregiver's long-term quality of life: results from an ALSOVA follow-up study. Qual Life Res 2016 Mar 9;25(3):687-697. [doi: 10.1007/s11136-015-1100-x] [Medline: 26350541]
14. Vitaliano PP, Scanlan JM, Zhang J, Savage MV, Hirsch IB, Siegler IC. A path model of chronic stress, the metabolic syndrome, and coronary heart disease. Psychosom Med 2002;64(3):418-435. [doi: 10.1097/00006842-200205000-00006] [Medline: 12021416]
15. Budiman A, Ruiz NG. Asian Americans are the fastest-growing racial or ethnic group in the U.S. Pew Research Center. 2021 Apr 09. URL: https://www.pewresearch.org/fact-tank/2021/04/09/asian-americans-are-the-fastest-growing-racial-or-ethnic-group-in-the-u-s/ [accessed 2022-09-23]
16. Sun F, Ong R, Burnette D. The influence of ethnicity and culture on dementia caregiving: a review of empirical studies on Chinese Americans. Am J Alzheimers Dis Other Demen 2012 Feb 30;27(1):13-22 [FREE Full text] [doi: 10.1177/1533317512438224] [Medline: 22467411]
17. Zhan L. Caring for family members with Alzheimer's disease: perspectives from Chinese American caregivers. J Gerontol Nurs 2004 Aug;30(8):19-29. [doi: 10.3928/0098-9134-20040801-06] [Medline: 15359526]
18. Lai DWL. Filial piety, caregiving appraisal, and caregiving burden. Res Aging 2009 Nov 13;32(2):200-223. [doi: 10.1177/0164027509351475]
19. Weng SS, Nguyen PV. Factors affecting elder caregiving in multigenerational Asian American families. Fam Soc 2018 May 03;92(3):329-335. [doi: 10.1606/1044-3894.4133]
20. Miyawaki CE. Caregiving attitudes and needs of later-generation Chinese-American family caregivers of older adults. J Fam Issues 2020 Jun 17;41(12):2377-2399. [doi: 10.1177/0192513x20930366]
21. Jang Y, Kim MT. Limited English proficiency and health service use in Asian Americans. J Immigr Minor Health 2019 Apr 24;21(2):264-270 [FREE Full text] [doi: 10.1007/s10903-018-0763-0] [Medline: 29797103]
22. Sun F, Mutlu A, Coon D. Service barriers faced by Chinese American families with a dementia relative: perspectives from family caregivers and service professionals. Clin Gerontol 2014 Feb 20;37(2):120-138. [doi: 10.1080/07317115.2013.868848]
23. Holland JM, Thompson LW, Tzuang M, Gallagher-Thompson D. Psychosocial factors among Chinese American women dementia caregivers and their association with salivary cortisol: results of an exploratory study. Ageing Int 2010 Jul 6;35(2):109-127. [doi: 10.1007/s12126-010-9057-0]
24. Hu M, Ma C, Sadarangani T, Wu B. Social-behavioral interventions for Asian and Hispanic American dementia caregivers: an integrative review. Aging Health Res 2021 Sep;1(3):100027. [doi: 10.1016/j.ahr.2021.100027]
25. Ma KPK, Saw A. An international systematic review of dementia caregiving interventions for Chinese families. Int J Geriatr Psychiatry 2020 Nov 10;35(11):1263-1284. [doi: 10.1002/gps.5400] [Medline: 32964577]
26. Gallagher-Thompson D, Gray HL, Tang PC, Pu CY, Leung LY, Wang P, et al. Impact of in-home behavioral management versus telephone support to reduce depressive symptoms and perceived stress in Chinese caregivers: results of a pilot study. Am J Geriatr Psychiatry 2007 May;15(5):425-434. [doi: 10.1097/JGP.0b013e3180312028] [Medline: 17463192]
27. Gallagher-Thompson D, Wang P, Liu W, Cheung V, Peng R, China D, et al. Effectiveness of a psychoeducational skill training DVD program to reduce stress in Chinese American dementia caregivers: results of a preliminary study. Aging Ment Health 2010 Apr;14(3):263-273. [doi: 10.1080/13607860903420989] [Medline: 20425645]
28. Tsetsi E, Rains SA. Smartphone internet access and use: extending the digital divide and usage gap. Mobile Media & Communication 2017 Jun 13;5(3):239-255. [doi: 10.1177/2050157917708329]
29. Faieta J, Sheehan J, DiGiovine C. mHealth interventions to improve health and quality of life related outcomes for informal dementia caregivers: a scoping review. Assist Technol 2022 May 04;34(3):362-374. [doi: 10.1080/10400435.2020.1829174] [Medline: 33006526]
30. Budiman A, Ruiz NG. Key facts about Asian Americans, a diverse and growing population. Pew Research Center. 2021 Apr 29. URL: https://www.pewresearch.org/fact-tank/2021/04/29/key-facts-about-asian-americans/ [accessed 2022-09-23]
31. Nielsen. Asian-American consumers are predictive adopters of new media platforms, online shopping, and smartphone use. PR Newswire. 2018 May 08. URL: https://www.prnewswire.com/news-releases/

XSL•FO
RenderX

nielsen-asian-american-consumers-are-predictive-adopters-of-new-media-platformsonline-shopping-and-smartphone-use-300644499.html [accessed 2022-09-23]
32. Number of monthly active WeChat users from 2nd quarter 2011 to 1st quarter 2020 (in millions). Statista. 2020 May. URL: https://www.statista.com/statistics/255778/number-of-active-wechat-messenger-accounts/ [accessed 2022-09-23]
33. Pagoto S, Waring ME, May CN, Ding EY, Kunz WH, Hayes R, et al. Adapting behavioral interventions for social media delivery. J Med Internet Res 2016 Jan 29;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
34. Belle SH, Burgio L, Burns R, Coon D, Czaja SJ, Gallagher-Thompson D, Resources for Enhancing Alzheimer's Caregiver Health (REACH) II Investigators. Enhancing the quality of life of dementia caregivers from different ethnic or racial groups: a randomized, controlled trial. Ann Intern Med 2006 Nov 21;145(10):727-738 [FREE Full text] [doi: 10.7326/0003-4819-145-10-200611210-00005] [Medline: 17116917]
35. Lorig K, Laurent D, Schreiber R, Gecht-Silver M, Gallagher Thompson D, Minor M, et al. Building Better Caregivers: A Caregiver's Guide to Reducing Stress and Staying Healthy. Boulder, CO: Bull Publishing; Jun 19, 2018.
36. Chen E, Leos C, Kowitt SD, Moracco KE. Enhancing community-based participatory research through human-centered design strategies. Health Promot Pract 2020 Jan 25;21(1):37-48. [doi: 10.1177/1524839919850557] [Medline: 31131633]
37. Hong YA, Juon HS, Chou WYS. Social media apps used by immigrants in the United States: challenges and opportunities for public health research and practice. mHealth 2021 Oct;7:52-52 [FREE Full text] [doi: 10.21037/mhealth-20-133] [Medline: 34805383]
38. Pagoto S, Waring ME, Xu R. A call for a public health agenda for social media research. J Med Internet Res 2019 Dec 19;21(12):e16661 [FREE Full text] [doi: 10.2196/16661] [Medline: 31855185]
39. Henderson S, Kendall E, See L. The effectiveness of culturally appropriate interventions to manage or prevent chronic disease in culturally and linguistically diverse communities: a systematic literature review. Health Soc Care Community 2011 May;19(3):225-249. [doi: 10.1111/j.1365-2524.2010.00972.x] [Medline: 21208326]
40. Guo Y, Hong YA, Qiao J, Xu Z, Zhang H, Zeng C, et al. Run4Love, a mHealth (WeChat-based) intervention to improve mental health of people living with HIV: a randomized controlled trial protocol. BMC Public Health 2018 Jun 26;18(1):793 [FREE Full text] [doi: 10.1186/s12889-018-5693-1] [Medline: 29940921]
41. Li Y, Guo Y, Hong YA, Zeng Y, Monroe-Wise A, Zeng C, et al. Dose-response effects of patient engagement on health outcomes in an mHealth intervention: secondary analysis of a randomized controlled trial. JMIR mHealth uHealth 2022 Jan 04;10(1):e25586 [FREE Full text] [doi: 10.2196/25586] [Medline: 34982724]
42. Nielsen J. Why you only need to test with 5 users. Nielsen Norman Group. 2000 Mar 18. URL: http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2022-08-12]

Abbreviations

ADRD: Alzheimer disease or related dementias
mHealth: mobile health
WECARE: Wellness Enhancement for Caregivers

Edited by T Leung; submitted 08.06.22; peer-reviewed by Y Du, Z Wang; comments to author 10.07.22; revised version received 28.07.22; accepted 07.09.22; published 29.09.22.

Please cite as:
Hong YA, Shen K, Lu HK, Chen H, Gong Y, Ta Park V, Han HR
A Social Media–Based Intervention for Chinese American Caregivers of Persons With Dementia: Protocol Development
JMIR Aging 2022;5(3):e40171
URL: https://aging.jmir.org/2022/3/e40171
doi:10.2196/40171
PMID:36173667

©Y Alicia Hong, Kang Shen, Huixing Kate Lu, Hsiaoyin Chen, Yang Gong, Van Ta Park, Hae-Ra Han. Originally published in JMIR Aging (https://aging.jmir.org), 29.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

App-Based Evaluation of Older People's Fall Risk Using the mHealth App Lindera Mobility Analysis: Exploratory Study

Nicole Strutz[1], MSc; Hanna Brodowski[2], MSc; Joern Kiselev[3], Dr rer med; Anika Heimann-Steinert[1], Dr rer med; Ursula Müller-Werdan[1], Prof Dr

[1]Geriatrics Research Group, Charité - Universitätsmedizin Berlin (corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin), Berlin, Germany

[2]Department of Physiotherapy, Pain and Exercise Research Lübeck, Institute of Health Sciences, University of Lübeck, Lübeck, Germany

[3]Department of Anesthesiology and Operative Intensive Care Medicine (Charité Campus Virchow Clinic/Campus Charité Mitte), Charité - Universitätsmedizin Berlin (corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin), Berlin, Germany

Corresponding Author:
Nicole Strutz, MSc
Geriatrics Research Group
Charité - Universitätsmedizin Berlin (corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin)
Reinickendorfer Straße 61
Berlin, 13407
Germany
Phone: 49 030 450 553369
Email: nicole.strutz@charite.de

Abstract

Background: Falls and the risk of falling in older people pose a high risk for losing independence. As the risk of falling progresses over time, it is often not adequately diagnosed due to the long intervals between contacts with health care professionals. This leads to the risk of falling being not properly detected until the first fall. App-based software able to screen fall risks of older adults and to monitor the progress and presence of fall risk factors could detect a developing fall risk at an early stage prior to the first fall. As smartphones become more common in the elderly population, this approach is easily available and feasible.

Objective: The aim of the study is to evaluate the app Lindera Mobility Analysis (LIN). The reference standards determined the risk of falling and validated functional assessments of mobility.

Methods: The LIN app was utilized in home- and community-dwelling older adults aged 65 years or more. The Berg Balance Scale (BBS), the Tinetti Test (TIN), and the Timed Up & Go Test (TUG) were used as reference standards. In addition to descriptive statistics, data correlation and the comparison of the mean difference of analog measures (reference standards) and digital measures were tested. Spearman rank correlation analysis was performed and Bland-Altman (B-A) plots drawn.

Results: Data of 42 participants could be obtained (n=25, 59.5%, women). There was a significant correlation between the LIN app and the BBS (r=–0.587, P<.001), TUG (r=0.474, P=.002), and TIN (r=–0.464, P=.002). B-A plots showed only few data points outside the predefined limits of agreement (LOA) when combining functional tests and results of LIN.

Conclusions: The digital app LIN has the potential to detect the risk of falling in older people. Further steps in establishing the validity of the LIN app should include its clinical applicability.

Trial Registration: German Clinical Trials Register DRKS00025352; https://tinyurl.com/65awrd6a

(JMIR Aging 2022;5(3):e36872) doi:10.2196/36872

KEYWORDS

mobility; fall risk; smartphone; app; analysis; older people; accuracy; mobility restriction

Introduction

As part of the aging process, older adults are affected by an increasing risk of falling as well as accidental falls [1]. In Europe, this development leads to fall incidence rates for older adults aged 70 years or more between 7500 and nearly 20,000 falls per 100,000 inhabitants and a death rate of up to 153.2 per 100,000 inhabitants [2]. In a study by Choi et al [3], observed fall-related injury locations in older adults (≥60 years, n=1840) included lower and upper extremities (32.06% and 23.12%,

XSL•FO
RenderX

respectively) but also 15.26% of falls resulting in head injuries, while 30.9% suffered 1 or more fractures. Additionally, falls and the risk of falling have a variety of effects on older adults' attitudes and behavior. Falls and even the risk of falling can pose a high risk of losing independence [4]. The risk of falls in older people changes over time as health status [5] or medication [6], either prescribed by a doctor or self-medication, changes. Often, the risk of falls increases with age-related decline in body musculature [7] and overall decrease in functional performance [8]. The risk of falling develops over time, and it is often underdiagnosed [9]. Therefore, the risk of falling is often not properly detected until the first fall. One possible solution to this dilemma is a more frequent self-assessment that should start before the first fall. Technology-based assessments of fall risk can assist an older adult in assessing their own fall risk. In this area, analyzing gait patterns is a widely used strategy to track the progress of functional abilities and to assess the risk of falling. However, gait analysis systems, such as GAITRite or SensFloor, cannot be applied at home with minimal effort. In contrast, as mobile phones become more widespread in the elderly population [10], an app-based fall risk assessment would be easily available and feasible. Mobile applicable apps, such as FallSA (a fall risk–screening app) [11] and Lindera Mobility Analysis (LIN; Lindera GmbH, Berlin, Germany) [12], 2 commercially available apps, are location independent and applicable at home.

As scientific evidence on the validity of such apps is limited, the aim of this explorative study was to evaluate the app LIN in comparison to established and validated functional assessments of mobility as a reference standard.

Methods

Study Design and Ethical Considerations

In 2021, this explorative validation study was conducted in Germany by the Geriatrics Research Group of Charité – Universitätsmedizin Berlin. The study was approved by the Ethics Committee of Charité – Universitätsmedizin Berlin (#EA1/363/20; date of approval: April 4, 2021). A sample size calculation was not performed as the study was exploratory in nature.

Recruitment

Participants were recruited from 3 sources: (1) the Geriatrics Research Group database, comprising older people who gave their consent to be contacted for participation in research projects; (2) older people who were staying in a geriatric hospital or day-care facility; and (3) a group of nursing home residents. Contact was made by mail, telephone, or a personal interview on-site. Inclusion criteria were age 65 years or older, being able to walk, and getting up from a chair and sitting down again. Participants were allowed to use walking aids, such as a wheeled walker or crutches. Exclusion criteria were defined as any fall events in the week before recruitment, more than 3 fall events during the past 6 months, and incapability of giving consent.

Data Collection

Data collection was conducted in the laboratory of the Geriatrics Research Group as well as in a nursing home and 2 day-care facilities. In addition to sociodemographic data, such as age and gender, the care level, degree of disability, data of mobility, and fall risk of the participants were recorded. The official care level within the German health care system ranges from level 0 (no need for care) to level 5 (maximum need for care)—§61b (1) German Social Code (SGB) XII, where SGB refers to the German Social Code. The official level of disability is characterized by level 20 (low disability) to level 100 (maximum disability)—§2 SGB IX. In addition, 4 mobility tests were performed, 3 reference assessments and LIN. In all measurements, LIN was used first. For this, participants filled out the app's questionnaire independently or, if preferred, together with the researcher. A video of the patient's gait was recorded using LIN on a smartphone. In a second step, 3 reference assessments were used to test the participants' fall risk and mobility restrictions. Between assessments, the participants could rest by answering the questionnaire on sociodemographic data. All data were collected within 1 session.

Lindera Mobility Analysis

LIN version 10.3.0 was used to determine the fall risk by computing a fall risk score. Input parameters to compute the fall risk score included (1) video analysis of each participant's gait through an artificial intelligence–based algorithm [13] and (2)a standardized questionnaire on further fall risk factors.

The assessment was conducted with a mobile app using a smartphone with an integrated camera. The fall risk score is the weighted sum of 14 fall risk factors, as defined by the German National Expert Standard Fall Prevention [14], a guideline developed and published by German Network for Quality Development in Nursing (DNQP) [15]. The standardized questionnaire addresses both person-related risk factors, such as polypharmacy, diseases, or alcohol consumption ("How often do you consume alcoholic beverages during the week: not at all, 1x-2x/week, 3x-5x/week, or 6x-7x/week?") and incontinence ("How often do you feel a sudden and urgent need to visit the toilet: never, rarely, sometimes, often, or always?", as can be seen in Figure 1), and environmental risk factors, such as floor coverings or door sills, as stated in the German National Expert Standard Fall Prevention.

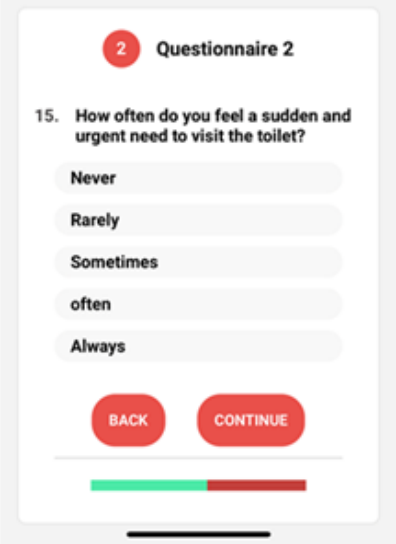
The results of the gait analysis and the questionnaire were computed into a score of 0-100 points, with a higher scoring indicating a higher fall risk.

The technical validity of LIN has been described elsewhere in several publications [12]. Thus, here, we provide a short summary.

The scientific approach underlying the app is based on a modular algorithm consisting of a video tester, a skeleton estimator (skeleton estimator 2D, skeleton estimator 3D, skeleton optimization 3D), and an analysis of mobility parameters. The skeleton estimator plays a central role. Both the validity of the mobility parameters and the validity of the analysis substantially depend on the spatial and temporal precision of the skeleton estimator [12].

XSL•FO
RenderX

Figure 1. Questionnaire: example of a person-related risk factor.



Reference Standards

Clinical guidelines recommend the evaluation of gait or balance disturbances to detect fall risk, but there is no gold standard for assessing the risk of falling in older adults measuring functional abilities [16]. However, there are several functional assessments available that have demonstrated good validity for identifying older people with a risk of falling. Three of the most widely used mobility assessments performed in therapeutic and nursing contexts are the Berg Balance Scale (BBS) [17], the Tinetti Test (TIN) [18], and the Timed Up & Go Test (TUG) [19]. In this study, these assessments were used as reference standards to evaluate functional mobility and balance.

TUG is a short-duration simple test on mobility [19], with a wide variety in clinical use. At the beginning of TUG, the participant sits on a chair, with arms placed on the armrests. On a command, the participant stands up, walks 3 m to a mark on the floor, turns around, and walks back to sit on the chair. TUG measures the time needed to complete the task in seconds. TUG is recommended as a routine screening test for falls in guidelines published by the American Geriatric Society and the British Geriatric Society [20] and has moderate-to-good sensitivity for predicting falls in older adults [16].

The BBS and TIN are scored based on a person's ability to perform specific tasks. The BBS was developed in 1989 to determine balance stability among older adults [17]. Today, it is commonly used to measure balance in people with various disabilities and health conditions. The BBS consists of 14 items assessing static and dynamic components of mobility and balance ability on multiple levels, including standing, transitional movement, and a narrowed base of support. Each item is scored on a 5-point Likert scale from 0 to 4, with 0 indicating the lowest level of function and 4 the highest. The maximum score is 56, with higher scores indicating higher levels of functional mobility and lower risk of falling [21].

A score below 45 points indicates a higher risk of falls [22]. Based on a systematic review [23], the BBS has high interrater reliability with a pooled estimate of 0.97 (95% CI 0.96-0.98) and high intrarater reliability with a pooled estimate of 0.98 (95% CI 0.97-to 0.99). The BBS can differentiate between fallers and nonfallers in community-dwelling older people [23].

TIN, also called Performance-Oriented Mobility Assessment (POMA), is a clinical balance assessment tool originally developed for use with institutionalized patients. It measures both balance and gait performance. Several versions of TIN are available, with varying numbers of items and score ranges [24]. In the version used in our study, mobility is assessed with 8 items each for balance and gait performance. The items are scored on a 2-4-point Likert scale, with a maximum score of 28 points. A score below 19 points indicates a high risk of falls [18]. TIN showed good-to-excellent interrater and intrarater reliability (intraclass correlation coefficient [ICC]>0.80) in a cohort of 30 participants with Parkinson disease [25].

Statistical Analysis

Baseline and sociodemographic data were collected, and Spearman rank correlation analysis was conducted.

Additionally limits of agreement (LOA) between LIN and TIN, the BBS, and TUG were evaluated using Bland-Altman (B-A) plots [26]. For plots with TIN and the BBS, we reversed the scale of LIN to adjust the direction of the scale to those of the reference scales. Next, we transformed results from TIN and the BBS into a ratio scale (0-100). As TUG and LIN both had the same direction of results (a higher score indicating a higher fall risk) and a transformation of TUG was not feasible, we performed all further steps with the original results obtained. Differences between each of the assessments and the results of LIN, as well as the mean of both respective measurements, and normal distributions of the differences of both observations were calculated. For this, we used the Shapiro-Wilk tests due to the relatively low number of participants, as well as visual inspection.

As the Shapiro-Wilk tests revealed mostly nonnormal distributions for the calculated differences between the measurements, we used the median and defined the upper and lower 95% of the sorted results as the threshold instead of the ±1.96 SD used for B-A plots with normal-distributed data. This approach was recommended by Bland and Altman [26] in their

XSL•FO
RenderX

original publication on drawing B-A plots for nonnormal-distributed data sets.

Baseline and sociodemographic data as well as all correlation analyses were calculated using SPSS Statistics version 28 (IBM Corporation, Armonk, NY, USA). All B-A plots were drawn using Microsoft Excel 2016.

Results

Participant Characteristics

Data of 42 participants, with a mean age of 77.6 (SD 7.3) years were analyzed. As can be seen in Table 1, there was a higher percentage of female participants (n=25, 59.5%). In addition, 25 (59.5%) of the participants did not have a care level, and 26 (65%) of 40 participants did not have a level of disability based on the grading within the German health care system.

One participant was not able to perform TUG due to difficulty in rising from the chair. Additionally, in 3 cases, data from LIN could not be interpreted and had to be discarded. Therefore, all correlation analyses were performed and B-A plots drawn with 39 and 38 data sets, respectively.

As can be seen in Table 2, low scores for TUG indicated a high degree of functional mobility, while for the BBS and TIN, high scores indicated a high degree of mobility, and low scores for LIN indicated a low level of fall risk.

In Table 3, the correlations of the analogous fall risk and mobility assessments and LIN are presented.

Low scores for TUG indicated a high degree of functional mobility, while for the BBS and TIN, high scores indicated a high degree of mobility, and low scores for LIN indicated a low level of fall risk.

The results of LIN demonstrated a high correlation with the BBS (r_s=–0.611) and a moderate-to-high correlation with TUG (r_s=0.583) and TIN (r_s=–0.563).

As can be seen in Figures 2-4, the results of the nonparametric B-A plots revealed a median of differences of –8.71 (TIN), 5.64 (BBS), and 3.3 (TUG). Most data pairs were within the predefined LOA. Only 2 data pairs (5.1%) outside the LOA could be found for the BBS, while for TIN, 3 outliers could be observed (7.7%) and only 1 for TUG (2.6%). However, a proportional bias could be observed in all 3 plots based on a significant linear regression coefficient (0.014 for TIN and <0.001 for the BBS and TUG). Visual inspection of the 3 plots revealed a tendency for higher differences between measurements for all comparisons. Additionally, as can be seen in Figure 4, the B-A plot comparing LIN and TUG showed a visible trend of a negative difference between the 2 measurements for lower means, while demonstrating positive differences for higher means.

Table 1. Baseline data.

Characteristics	Participants
Age (years; N=42), mean (SD)	77.6 (7.3)
Female gender (N=42), n (%)	25 (59.5)
Level of disability (N=40)[a], n (%)	
No level	26 (65)
<30	1 (2.5)
31-60	8 (20.0)
61-80	5 (12.5)
>80	0
Care level (N=42), n (%)	
0	25 (59.5)
1	2 (4.7)
2	7 (16.7)
3	7 (16.7)
4	1 (2.4)
5	0

[a]The official level of disability is characterized by level 20 (low disability) up to level 100 (maximum disability)—§2 German Social Code (SGB) IX.

XSL•FO
RenderX

Table 2. Mobility data.

Assessment	Mean (SD)	Minimum	Maximum
TUG[a] (N=40)	13.7 (5.8)	6.9	36
TIN[b] (N=42)	23.9 (5.3)	8	28
BBS[c] (N=42)	44.7 (13.0)	7	56
LIN[d] (N=39)	19.8 (12.4)	5	68

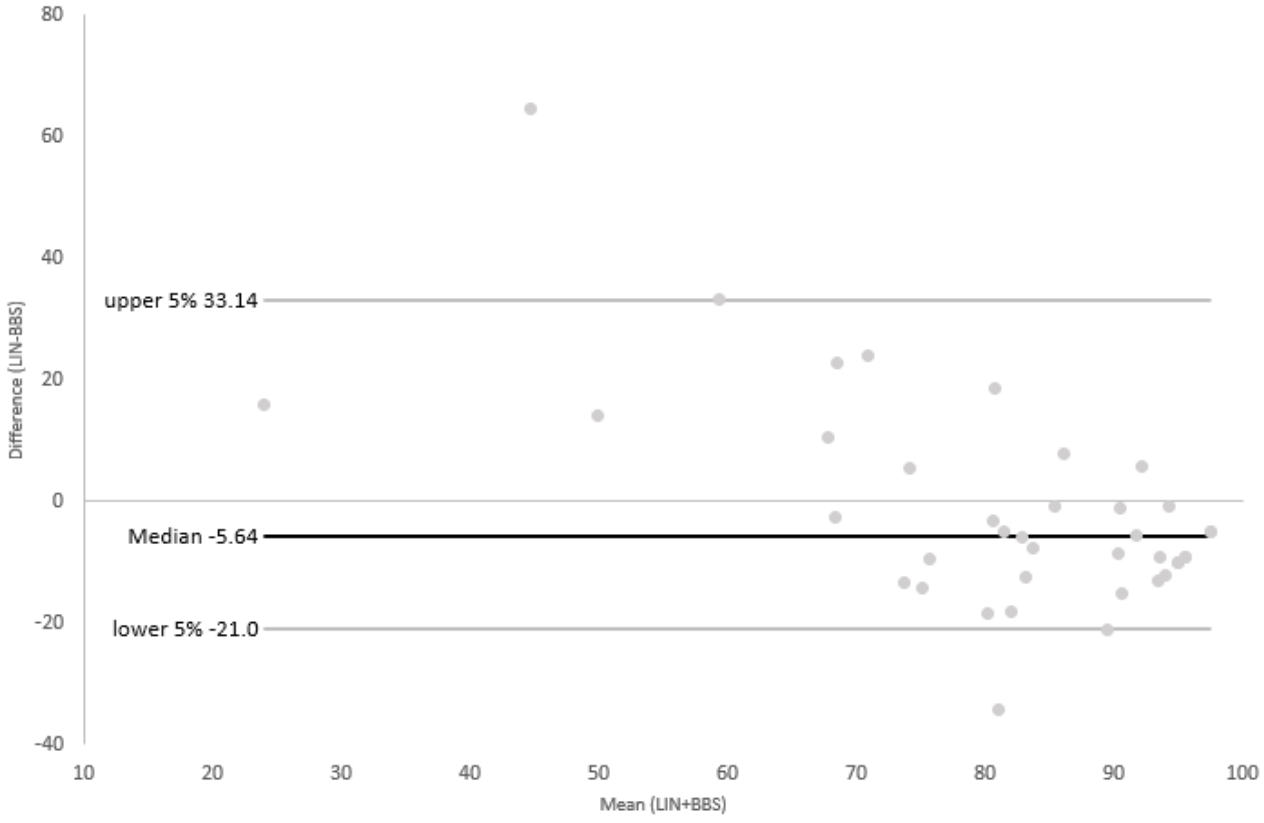
[a]TUG: Timed Up & Go Test.
[b]TIN: Tinetti Test.
[c]BBS: Berg Balance Scale.
[d]LIN: Lindera Mobility Analysis.

Table 3. Spearman rank correlation of analog and digital fall risk and mobility assessment.

Assessment	TUG[a]			TIN[b]			LIN[c]		
	r_s	*P* value	N	r_s	*P* value	N	r_s	*P* value	N
BBS[d]	–0.770[e]	.001	40	.730[e]	.001	42	–0.611[e]	.001	39
TUG	N/A[f]	N/A	N/A	–0.526[e]	.001	40	.583[e]	.001	38
TIN	N/A	N/A	N/A	N/A	N/A	N/A	–0.563[e]	.001	39

[a]TUG: Timed Up & Go Test.
[b]TIN: Tinetti Test.
[c]LIN: Lindera Mobility Analysis.
[d]BBS: Berg Balance Scale.
[e]The correlation was significant at the level of .01.
[f]N/A: not applicable.

Figure 2. B-A plot of LIN and the BBS. B-A: Bland-Altman; BBS: Berg Balance Scale; LIN: Lindera Mobility Analysis.



XSL•FO
RenderX

Figure 3. B-A plot of TIN and LIN. B-A: Bland-Altman; LIN: Lindera Mobility Analysis; TIN: Tinetti Test.

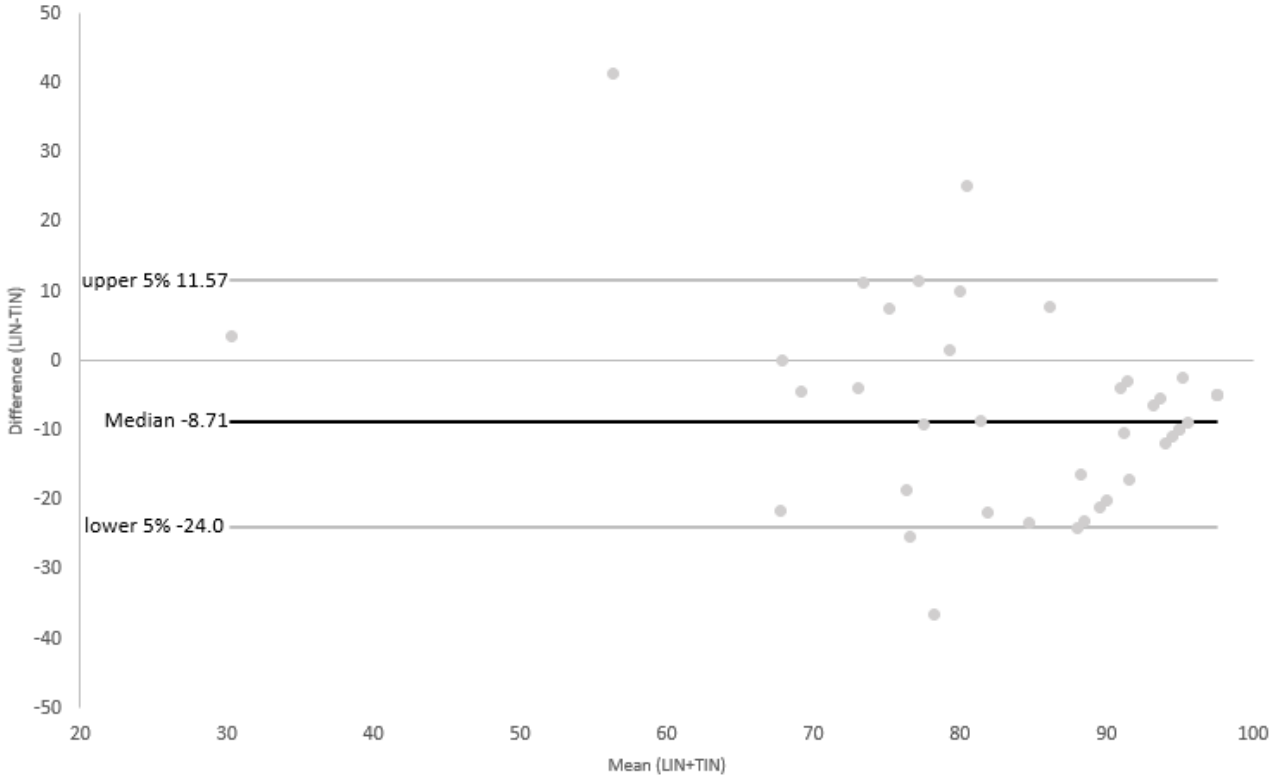
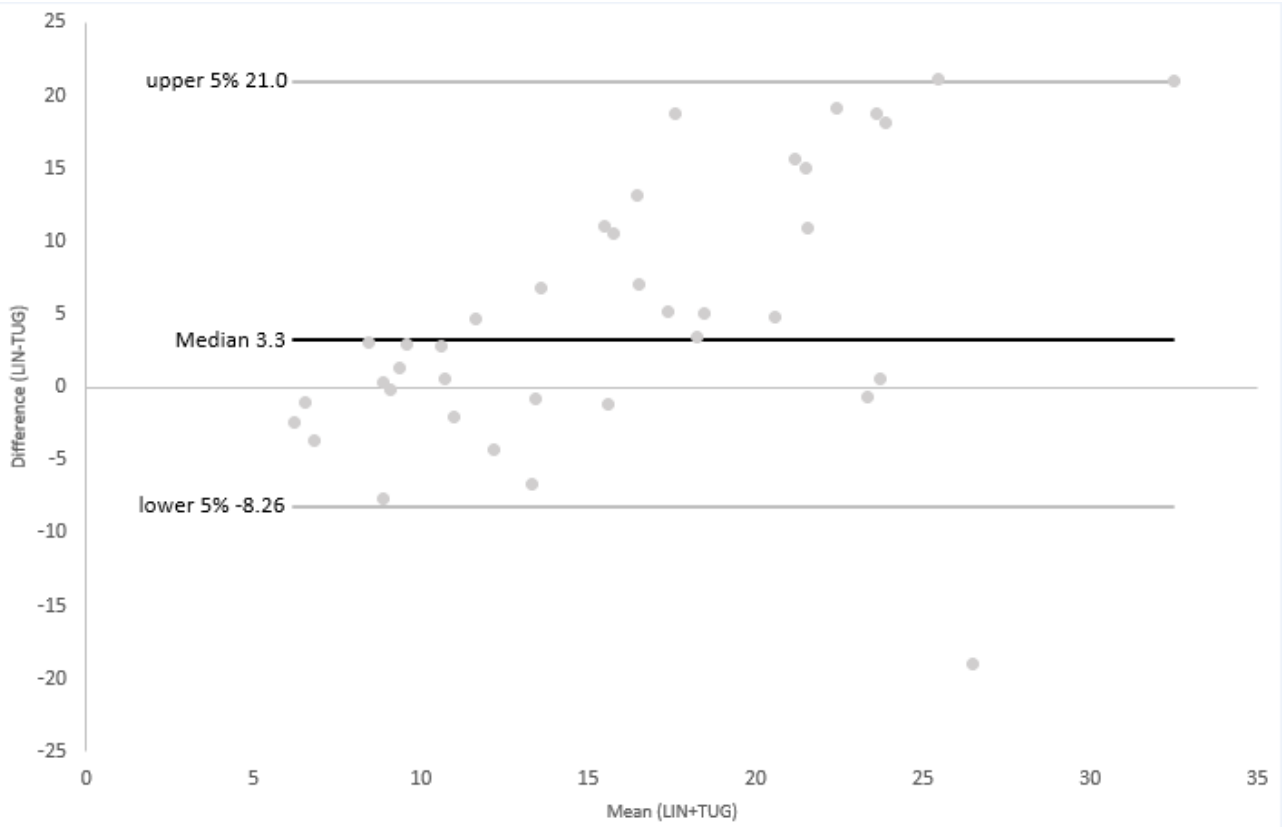


Figure 4. B-A plot of TUG and LIN. B-A: Bland-Altman; LIN: Lindera Mobility Analysis; TUG: Timed Up & Go Test.



XSL•FO
RenderX

Discussion

Principal Findings

The aim of this study was to evaluate the accuracy of LIN compared to reference standards for analog objective measures of older people's fall risk. As our study shows, a moderate-to-high correlation according to Cohen [27] was found for LIN and the BBS, TIN and TUG. In this, the lowest correlation for the 3 reference assessments could be observed for LIN and TUG.

The results of our correlation analyses were verified by the B-A plots drawn. The B-A plots showed only a minority of the data pairs outside the predefined 95% limits. However, we observed a low-to-moderate proportional bias of the differences between results of LIN and the respective reference standards, indicating that both respective measurements might not be depicting the same construct. Moreover, we observed a skew in all plots, validating the observation of the correlation analyses. Due to the range and direction of the scales indicating a higher fall risk, we needed to transform our data for 2 plots in order to be able to obtain interpretable results. Additionally, as the differences between measurements were not normally distributed, we had to draw our B-A plots based on a nonparametric version.

This might have contributed to the results of the drawn plots. However, results from both correlation analyses and B-A plots could be interpreted as a sign that LIN can actually be superior in detecting older people at risk of falling compared to the 3 reference standards.

All 3 assessments are established tools for predicting falls in older people; however, none of them can be labeled as a gold standard. Although there might be different reasons for this, all of them have known flaws that have to be considered when planning to use any of them. As mentioned before, there are several versions available for TIN, making comparison between studies difficult. Additionally, both TIN and the BBS demonstrate only good but not high sensitivity and specificity for fall prediction in older adults living in care residence facilities [28]. The authors recommended using a combination of the BBS and a gait speed test in order to obtain more dependable results in this population. For TUG, Haines et al [29] found comparable problems in a population of older adults in a geriatric ward.

This merits some consideration. In contrast to TUG, LIN, TIN, and the BBS record complex movement sequences and thus evaluate balance, postural control, and gait symmetry.

In contrast, TUG merges all these functional requirements into 1 single information piece, the time needed to complete TUG. As a consequence, a lot of technology-based research aims at increasing the information value gathered through the relative easy-to-administer TUG, where TUG performance is often used to gather not only the TUG time but also the TUG stride length, as well as the forward und lateral tilt of the trunk and gait symmetry. Although TUG's ability to predict falls in older adults has been established [19], several attempts have been made to increase the level of obtainable information while performing TUG, using video data and different sensor arrays [30-32]. All these studies have been, at least partly, successful in gathering information about gait and balance while performing TUG, but it still makes direct comparison between the original TUG and the expanded, technology-based versions difficult. In our study, LIN, in addition to information from a questionnaire and the time to complete TUG, measured other factors, such as stride length and the forward and lateral tilt of the trunk and gait symmetry. As stated before, TIN and the BBS evaluate complex movement sequences that resemble a wide variety of everyday activities and thus test a participant's balance, postural control, and gait symmetry. Therefore, the gathered data seem to be more comparable to more complex (and time-demanding) assessments, such as the BBS and TIN. This is, in our opinion, reflected in the high correlation coefficients between LIN and the BBS and TIN in contrast to the more modest correlation with TUG.

Additionally, LIN uses an additional questionnaire based on the German National Expert Standard Fall Prevention and as such provides a guideline for the prevention of falls [14]. The questionnaire encompasses items about not only intrinsic factors, such as comorbidities, incontinence, fear of falling, and prior falls, but also extrinsic factors, such as mobility aids, barriers in the living environment, shoes used at home, and several other factors that have been identified as contributors to the risk of falling. Therefore, LIN includes, in comparison to the functional assessments of gait and balance that are recommended in geriatrics and were used in this study, more dimensions of the phenomenon of falls in older adults and is, thus, in our opinion, more comprehensive that a purely functional assessment for identifying patients with fall risk. Whether this leads to any potential superiority of LIN cannot be answered based on the available data. For this, additional research is necessary that includes the prospective establishment of diagnostic criteria as well as its ability to prevent falls. We conclude therefore that for gaining deeper insight into the potential of technology-based mobility and fall risk assessments, more detailed comparators are needed.

Despite these limitations, we deem our results satisfactory. The low number of data pairs outside the LOA indicate, in our estimation, a satisfactory level of comparability of the results of LIN with our reference standards. The observable bias in all 3 plots is, in our estimation, acceptable. Due to the reason stated before insofar, a complete agreement between the measurements cannot be expected. However, we are aware of the fact that the results presented here have to be interpreted with caution and have to be verified in further studies.

Compared to other apps for fall risk analysis, such as FallSA [33], LIN showed a slightly higher significant correlation with the established BBS. FallSA was significant moderately correlated (r=0.518, P<.001) with the Physical Profile Assessment [11]. In 2021, iPhone manufacturer Apple Inc. offered a function in the current version of its iOS (iOS 15) that is supposed to prevent falls. As the manufacturer stated, "Walking Steadiness on iPhone is a first-of-its-kind health metric that can give you insight into your risk of falling. It uses custom algorithms that assess your balance, strength, and gait" [34]. Based on calculated gait stability, the software is supposed to predict the risk of falling. Both FallSA and iOS 15 measure

XSL•FO
RenderX

functional ability. In contrast, LIN is based on the measurement of functional ability and surveying intrinsic factors of its users. Furthermore, the FallSA app as well as iOS 15 are not specifically labeled as medical devices in the sense of the European Medical Device Regulation—Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on Medical Devices. In contrast, LIN is a class 1 medical device. Being a medical device allows professionals involved in care, such as nurses, physical therapists, and physicians, to use the results of the app to assist their nursing appraisals or diagnoses.

Using LIN or other medical devices with the ability to identify fall risks in older people while involving health professionals offers great potential. In 2021, Meekes et al [35] studied the level of information general practitioners (GPs) had available for any of their patients with frailty about their fall history as well as the occurrence of fear of falling. In their study, GPs had no information about fall history in 668 (48%) of the affected patients [35].

Additionally, as several studies have demonstrated that a significant portion of patients tend to underestimate their own fall risk [36,37], the LIN app offers high potential for determining one's own fall risk as a nonprofessional. This gives older people an opportunity to self-assess their own fall risk and, with repeated measurements, any changes in their fall risk status over time.

Conclusion

Using LIN has the potential to enable older people to be more independent of the initial determination of a fall risk by GPs or other health care professionals and also enables them to identify and respond to positive or negative changes in their own fall risk. This provides older adults with the ability to manage their own fall risk in an effective and adequate manner. Using LIN can help reduce fall events in people aged 65 years or more. Further study is indicated to verify validity.

Acknowledgments

We thank Lindera GmbH for providing the app Lindera Mobility Analysis (LIN) for the study. We acknowledge support from the German health insurance company Allgemeine Ortskrankenkasse (AOK).

We also acknowledge financial support from the Open Access Publication Fund of Charité – Universitätsmedizin Berlin and the German Research Foundation (DFG).

Conflicts of Interest

The authors report that there are no competing interests to declare. The study was commissioned by Lindera GmbH. Lindera GmbH did not influence the results presented in this study.

References

1. Todd CEA. World Health Organisation Global Report on Falls Prevention in Older Age/Citation Formats. URL: https://tinyurl.com/msuxm9us [accessed 2021-12-15]
2. Haagsma JA, Olij BF, Majdan M, van Beeck EF, Vos T, Castle CD, et al. Falls in older aged adults in 22 European countries: incidence, mortality and burden of disease from 1990 to 2017. Inj Prev 2020 Oct;26(Supp 1):i67-i74 [FREE Full text] [doi: 10.1136/injuryprev-2019-043347] [Medline: 32111726]
3. Choi NG, Choi BY, DiNitto DM, Marti CN, Kunik ME. Fall-related emergency department visits and hospitalizations among community-dwelling older adults: examination of health problems and injury characteristics. BMC Geriatr 2019 Nov 11;19(1):303 [FREE Full text] [doi: 10.1186/s12877-019-1329-2] [Medline: 31711437]
4. Deshpande N, Metter EJ, Lauretani F, Bandinelli S, Guralnik J, Ferrucci L. Activity restriction induced by fear of falling and objective and subjective measures of physical function: a prospective cohort study. J Am Geriatr Soc 2008 Apr;56(4):615-620 [FREE Full text] [doi: 10.1111/j.1532-5415.2007.01639.x] [Medline: 18312314]
5. Hoffman GJ, Rodriguez HP. Examining contextual influences on fall-related injuries among older adults for population health management. Popul Health Manag 2015 Dec;18(6):437-448 [FREE Full text] [doi: 10.1089/pop.2014.0156] [Medline: 25919228]
6. Lippert T, Maas R, Fromm MF, Luttenberger K, Kolominsky-Rabas P, Pendergrass A, et al. [Impact of sedating drugs on falls resulting injuries among people with dementia in a nursing home setting]. Gesundheitswesen 2020 Jan 21;82(1):14-22 [FREE Full text] [doi: 10.1055/a-1071-7911] [Medline: 31962367]
7. Yeung SS, Reijnierse EM, Pham VK, Trappenburg MC, Lim WK, Meskers CG, et al. Sarcopenia and its association with falls and fractures in older adults: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2019 Jun 16;10(3):485-500 [FREE Full text] [doi: 10.1002/jcsm.12411] [Medline: 30993881]
8. Unhjem R, van den Hoven LT, Nygård M, Hoff J, Wang E. Functional performance with age: the role of long-term strength training. J Geriatr Phys Ther 2019;42(3):115-122. [doi: 10.1519/JPT.0000000000000141] [Medline: 28786909]
9. Pérez-Ros P, Martínez-Arnau FM, Orti-Lucas RM, Tarazona-Santabalbina FJ. A predictive model of isolated and recurrent falls in functionally independent community-dwelling older adults. Braz J Phys Ther 2019 Jan;23(1):19-26 [FREE Full text] [doi: 10.1016/j.bjpt.2018.05.005] [Medline: 29914855]

XSL•FO
RenderX

10. Kakulla BN. 2020 Tech Trends of the 50+. URL: https://www.aarp.org/research/topics/technology/info-2019/2020-technology-trends-older-americans.html [accessed 2021-12-15]
11. Singh DKA, Goh JW, Shaharudin MI, Shahar S. A mobile app (FallSA) to identify fall risk among Malaysian community-dwelling older persons: development and validation study. JMIR Mhealth Uhealth 2021 Oct 12;9(10):e23663. [doi: 10.2196/23663]
12. Azhand A, Rabe S, Müller S, Sattler I, Heimann-Steinert A. Algorithm based on one monocular video delivers highly valid and reliable gait parameters. Sci Rep 2021 Jul 07;11(1):14065 [FREE Full text] [doi: 10.1038/s41598-021-93530-z] [Medline: 34234255]
13. Stamm O, Heimann-Steinert A. Accuracy of monocular two-dimensional pose estimation compared with a reference standard for kinematic multiview analysis: validation study. JMIR Mhealth Uhealth 2020 Dec 21;8(12):e19608 [FREE Full text] [doi: 10.2196/19608] [Medline: 33346739]
14. Schmidt S. Expertenstandard Sturzprophylaxe in der Pflege. In: Expertenstandards in der Pflege - eine Gebrauchsanleitung. Berlin, Heidelberg: Springer; 2016.
15. Büscher A, Möller A. Deutsches Netzwerk für Qualitätsentwicklung in der Pflege (DNQP): Aktueller Stand. Public Health Forum 2014;22(2):27-28. [doi: 10.1016/j.phf.2014.03.015]
16. Nordin E, Lindelöf N, Rosendahl E, Jensen J, Lundin-Olsson L. Prognostic validity of the Timed Up-and-Go test, a modified Get-Up-and-Go test, staff's global judgement and fall history in evaluating fall risk in residential care facilities. Age Ageing 2008 Jul;37(4):442-448. [doi: 10.1093/ageing/afn101] [Medline: 18515291]
17. Berg K. Measuring balance in the elderly: preliminary development of an instrument. Physiother Can 1989 Nov;41(6):304-311. [doi: 10.3138/ptc.41.6.304]
18. Tinetti ME. Performance-oriented assessment of mobility problems in elderly patients. J Am Geriatr Soc 1986 Feb;34(2):119-126. [doi: 10.1111/j.1532-5415.1986.tb05480.x] [Medline: 3944402]
19. Podsiadlo D, Richardson S. J Am Geriatr Soc 1991 Feb 27;39(2):142-148. [doi: 10.1111/j.1532-5415.1991.tb01616.x] [Medline: 1991946]
20. Panel on Prevention of Falls in Older Persons, American Geriatrics SocietyBritish Geriatrics Society. Summary of the Updated American Geriatrics Society/British Geriatrics Society clinical practice guideline for prevention of falls in older persons. J Am Geriatr Soc 2011 Jan;59(1):148-157. [doi: 10.1111/j.1532-5415.2010.03234.x] [Medline: 21226685]
21. Lajoie Y. Effect of computerized feedback postural training on posture and attentional demands in older adults. Aging Clin Exp Res 2013 Jul 25;16(5):363-368. [doi: 10.1007/bf03324565]
22. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83 Suppl 2:S7-S11. [Medline: 1468055]
23. Marques A, Almeida S, Carvalho J, Cruz J, Oliveira A, Jácome C. Reliability, validity, and ability to identify fall status of the balance evaluation systems test, mini-balance evaluation systems test, and brief-balance evaluation systems test in older people living in the community. Arch Phys Med Rehabil 2016 Dec;97(12):2166-2173.e1. [doi: 10.1016/j.apmr.2016.07.011] [Medline: 27497826]
24. Köpke S, Meyer G. The Tinetti test: Babylon in geriatric assessment. Z Gerontol Geriatr 2006 Aug;39(4):288-291. [doi: 10.1007/s00391-006-0398-y] [Medline: 16900448]
25. Kegelmeyer DA, Kloos AD, Thomas KM, Kostyk SK. Reliability and validity of the Tinetti Mobility Test for individuals with Parkinson disease. Phys Ther 2007 Oct;87(10):1369-1378. [doi: 10.2522/ptj.20070007] [Medline: 17684089]
26. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999 Jun 02;8(2):135-160. [doi: 10.1177/096228029900800204] [Medline: 10501650]
27. Cohen J. Set correlation as a general multivariate data-analytic method. Multivariate Behav Res 1982 Jul 01;17(3):301-341. [doi: 10.1207/s15327906mbr1703_2] [Medline: 26800754]
28. Harada N, Chiu V, Damron-Rodriguez J, Fowler E, Siu A, Reuben DB. Screening for balance and mobility impairment in elderly individuals living in residential care facilities. Phys Ther 1995 Jun;75(6):462-469. [doi: 10.1093/ptj/75.6.462] [Medline: 7770493]
29. Haines T, Kuys SS, Morrison G, Clarke J, Bew P. Balance impairment not predictive of falls in geriatric rehabilitation wards. J Gerontol A Biol Sci Med Sci 2008 May 01;63(5):523-528. [doi: 10.1093/gerona/63.5.523] [Medline: 18511758]
30. Botolfsen P, Helbostad JL, Moe-Nilssen R, Wall JC. Reliability and concurrent validity of the Expanded Timed Up-and-Go test in older people with impaired mobility. Physiother Res Int 2008 Jun;13(2):94-106. [doi: 10.1002/pri.394] [Medline: 18288773]
31. Dibble LE, Lange M. Predicting falls in individuals with Parkinson disease: a reconsideration of clinical balance measures. J Neurol Phys Ther 2006 Jun;30(2):60-67. [doi: 10.1097/01.npt.0000282569.70920.dc] [Medline: 16796770]
32. Fudickar S, Kiselev J, Frenken T, Wegel S, Dimitrowska S, Steinhagen-Thiessen E, et al. Validation of the ambient TUG chair with light barriers and force sensors in a clinical trial. Assist Technol 2020 May 17;32(1):1-8. [doi: 10.1080/10400435.2018.1446195] [Medline: 29482463]
33. Lord SR, Menz HB, Tiedemann A. A physiological profile approach to falls risk assessment and prevention. Phys Ther 2003;83(3):237-252. [doi: 10.1093/ptj/83.3.237]
34. Apple (Deutschland). iOS - Gesundheit. URL: https://www.apple.com/de/ios/health/ [accessed 2021-12-20]

XSL•FO
RenderX

35. Meekes WMA, Leemrijse CJ, Weesie YM, van de Goor IAM, Donker GA, Korevaar JC. Falls prevention at GP practices: a description of daily practice. BMC Fam Pract 2021 Sep 21;22(1):190 [FREE Full text] [doi: 10.1186/s12875-021-01540-7] [Medline: 34548022]
36. Braun BL. Knowledge and perception of fall-related risk factors and fall-reduction techniques among community-dwelling elderly individuals. Phys Ther 1998 Dec;78(12):1262-1276. [doi: 10.1093/ptj/78.12.1262] [Medline: 9859946]
37. Moreira MN, Bilton TL, Dias RC, Ferriolli E, Perracini MR. What are the main physical functioning factors associated with falls among older people with different perceived fall risk? Physiother Res Int 2017 Jul 07;22(3):e1664. [doi: 10.1002/pri.1664] [Medline: 26949232]

Abbreviations

B-A: Bland-Altman
BBS: Berg Balance Scale
GP: general practitioner
LIN: Lindera Mobility Analysis
LOA: limits of agreement
SGB: German Social Code
TIN: Tinetti Test
TUG: Timed Up & Go Test

Edited by T Leung, J Wang; submitted 28.01.22; peer-reviewed by F Fischer, H Wang; comments to author 18.02.22; revised version received 07.06.22; accepted 27.06.22; published 16.08.22.

Please cite as:
Strutz N, Brodowski H, Kiselev J, Heimann-Steinert A, Müller-Werdan U
App-Based Evaluation of Older People's Fall Risk Using the mHealth App Lindera Mobility Analysis: Exploratory Study
JMIR Aging 2022;5(3):e36872
URL: https://aging.jmir.org/2022/3/e36872
doi:10.2196/36872
PMID:35972785

©Nicole Strutz, Hanna Brodowski, Joern Kiselev, Anika Heimann-Steinert, Ursula Müller-Werdan. Originally published in JMIR Aging (https://aging.jmir.org), 16.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An Unsupervised Data-Driven Anomaly Detection Approach for Adverse Health Conditions in People Living With Dementia: Cohort Study

Nivedita Bijlani[1], BEng, MS; Ramin Nilforooshan[2,3,4], MD, MRCPsych; Samaneh Kouchaki[1,3], BSc, MSc, PhD

[1]Centre for Vision, Speech and Signal Processing, University of Surrey, Guildford, United Kingdom

[2]Surrey and Borders Partnership NHS Foundation Trust, Guildford, United Kingdom

[3]Care Research and Technology Centre, UK Dementia Research Institute, Imperial College, London, United Kingdom

[4]School of Psychology, University of Surrey, Guildford, United Kingdom

Corresponding Author:
Nivedita Bijlani, BEng, MS
Centre for Vision, Speech and Signal Processing
University of Surrey
388 Stag Hill
Guildford, GU2 7XH
United Kingdom
Phone: 44 1483 300 800
Fax: 44 1483 300 803
Email: n.bijlani@surrey.ac.uk

Abstract

Background: Sensor-based remote health monitoring can be used for the timely detection of health deterioration in people living with dementia with minimal impact on their day-to-day living. Anomaly detection approaches have been widely applied in various domains, including remote health monitoring. However, current approaches are challenged by noisy, multivariate data and low generalizability.

Objective: This study aims to develop an online, lightweight unsupervised learning–based approach to detect anomalies representing adverse health conditions using activity changes in people living with dementia. We demonstrated its effectiveness over state-of-the-art methods on a real-world data set of 9363 days collected from 15 participant households by the UK Dementia Research Institute between August 2019 and July 2021. Our approach was applied to household movement data to detect urinary tract infections (UTIs) and hospitalizations.

Methods: We propose and evaluate a solution based on Contextual Matrix Profile (CMP), an exact, ultrafast distance-based anomaly detection algorithm. Using daily aggregated household movement data collected via passive infrared sensors, we generated CMPs for location-wise sensor counts, duration, and change in hourly movement patterns for each patient. We computed a normalized anomaly score in 2 ways: by combining univariate CMPs and by developing a multidimensional CMP. The performance of our method was evaluated relative to Angle-Based Outlier Detection, Copula-Based Outlier Detection, and Lightweight Online Detector of Anomalies. We used the multidimensional CMP to discover and present the important features associated with adverse health conditions in people living with dementia.

Results: The multidimensional CMP yielded, on average, 84.3% recall with 32.1 alerts, or a 5.1% alert rate, offering the best balance of recall and relative precision compared with Copula-Based and Angle-Based Outlier Detection and Lightweight Online Detector of Anomalies when evaluated for UTI and hospitalization. Midnight to 6 AM bathroom activity was shown to be the most important cross-patient digital biomarker of anomalies indicative of UTI, contributing approximately 30% to the anomaly score. We also demonstrated how CMP-based anomaly scoring can be used for a cross-patient view of anomaly patterns.

Conclusions: To the best of our knowledge, this is the first real-world study to adapt the CMP to continuous anomaly detection in a health care scenario. The CMP inherits the speed, accuracy, and simplicity of the Matrix Profile, providing configurability, the ability to denoise and detect patterns, and explainability to clinical practitioners. We addressed the need for anomaly scoring in multivariate time series health care data by developing the multidimensional CMP. With high sensitivity, a low alert rate, better overall performance than state-of-the-art methods, and the ability to discover digital biomarkers of anomalies, the CMP is a

XSL•FO
RenderX

clinically meaningful unsupervised anomaly detection technique extensible to multimodal data for dementia and other health care scenarios.

(JMIR Aging 2022;5(3):e38211) doi:10.2196/38211

KEYWORDS

contextual matrix profile; multidimensional anomaly detection; outlier detection; sensor-based remote health monitoring; dementia; unsupervised learning

Introduction

Background

Dementia is a progressive and irreversible decline in a wide range of brain activities, including impaired memory, thinking, orientation, comprehension, calculation, learning capacity, language, and judgment, beyond what might be expected from natural biological aging. The World Health Organization estimates that approximately 55 million people have dementia worldwide, which is set to rise to 78 million in 2030 and 139 million in 2050 [1]. Managing the care of this growing population incurs significant costs. The Alzheimer's Society puts the cost of care for people with dementia in the United Kingdom at GBP 34.7 billion (US $40 billion), rising sharply to GBP 94.1 billion (US $108.6 billion) by 2040 [2]. The hospitalization of people living with dementia because of potentially preventable conditions such as fall injuries, sepsis, pneumonia, and urinary tract infection (UTI) puts huge pressure on health systems. To minimize preventable hospitalizations, there is a significant investment in artificial intelligence–driven technologies that enable the health of people living with dementia to be remotely monitored and assisted while they live in the comfort of their own homes.

The UK Dementia Research Institute Care Research and Technology Centre has made a significant effort in this direction with its vision to "use patient-centered technology to help people affected by dementia to live better and longer in their own homes" [3]. The team at the UK Dementia Research Institute Care Research and Technology Centre has developed a sensor-based remote health monitoring platform that enables clinicians to intervene early and allows researchers to improve their understanding of dementia onset and progression [4]. The cohort currently covers 102 people with dementia living with their caregivers in their own homes. Data collection commenced in 2019 and will continue until at least 2025, with more participants being onboarded each year, making it one of the largest, longest-running, and most diverse and unique dementia data collection programs worldwide. The sensors, framework, models, clinical monitoring workflows, app for participants, and monitoring dashboard together form a digital platform called Minder (please see the website of the UK Dementia Research Institute [4] for more information).

Occasionally, people with dementia present with behavioral and psychological symptoms such as agitation, aggression, sleep disturbances, urinary system disorders, dehydration, and falls. UTI is the most diagnosed infection in older adults, and early identification is key to preventing further complications [5,6]. The diagnosis of UTI remains problematic because of the presence of a range of nonspecific symptoms, a high prevalence of asymptomatic bacteriuria, and reduced help-seeking behavior [7-9].

An "anomaly" in the context of home health monitoring can be simply understood as an unexpected but significant irregularity in otherwise normal data, which is indicative of an adverse condition. Anomalies are difficult to detect within overwhelming volumes of normal data. The cost of missing or misclassifying anomalies can be high (eg, failing to detect a UTI could be catastrophic). Current methods for health care anomaly detection are challenged by one or more real-world issues: high-dimensional and multivariate data; little to no information on the distinction between normal and abnormal data; time course data and the need to make predictions with low latency; patient-to-patient variability; noise and lack of periodicity because of social visits, pets, sensor issues, and noisy labels; high false alert rate; high tuning needs; and low explainability to clinical monitoring teams and caregivers [10].

The aim of our work was to develop a clinically useful, domain-agnostic, fast, lightweight, unsupervised anomaly detection approach for real-world noisy health care data. We accounted for individual variability, generalizability across individuals and domains, and explainability to clinicians and carers in the form of digital biomarker discovery. Our work makes the following contributions: (1) it offers the first use case for the Contextual Matrix Profile (CMP) for adaptive anomaly detection in health care, specifically in a real-world remote health monitoring scenario; (2) it develops the multidimensional CMP and uses it to identify and score anomalous patient days; (3) it demonstrates the effectiveness of CMP-based anomaly scoring over state-of-the-art methods; and (4) it uses the CMP to discover biomarkers of anomalies using household movement data.

Prior Work

Overview

Approaches to anomaly detection can be broadly categorized as statistical, distance-based, reconstruction-based, domain-based or decision boundary–based, information-theoretic, and graph-based [11]. Many approaches in the literature use combinations of techniques such as visual, knowledge-based, and machine learning approaches. We highlight how some of these techniques have been applied to anomaly detection in remote health monitoring scenarios.

Statistical Methods

Statistical thresholding is a popular approach to finding point anomalies. A National Institutes of Health–funded pilot study used statistical thresholding to generate alerts for UTI and offered early interventions for 37 older adult participants, some

XSL•FO
RenderX

with Alzheimer disease, residing in apartments equipped with motion, pressure, and temperature sensors [12]. Clustering-based techniques were used in the study by Mori et al [13] to detect anomalies in the timing and duration of different activities. Statistical methods typically ignore the multivariate nature of anomalous events and can generate numerous false positives [14].

Machine Learning Approaches

Using early data from our Minder study, Enshaeifar et al [15] used a Markov chain to model activity sequences along with an entropy rate to quantify the regularity of an individual's patterns in their day-to-day life. They used a training set to construct the Markov model and a verification set to define a confidence threshold for deviations [15]. Novák et al [16] detected anomalies such as long periods of inactivity, unusual presence, and changes in daily activity patterns using a combination of self-organizing maps for activity classification followed by a Markov model for next activity prediction. The limitations of the Markov approach include the inability to address parallel activities, activities that involve the same event with different probabilities, and scalability issues [17,18].

Arifoglu and Bouchachia [5] explored convolutional neural networks to capture temporal and spatial representations of activity and detect abnormal behavior related to repeating activities, sleep disruption, and confusion. Sensor data were sliced into time windows, and activities were labeled via sequence labeling to train convolutional neural networks that could detect deviations from normal daily life sequences. Supervised learning and interpretability are some limitations of this approach.

Akl et al [19] used signal processing with machine learning algorithms to detect mild cognitive impairments in older adults. They used sensors to extract the average, probability density, and trajectory of measures over sliding windows of sensor data as input to support vector machines and random forest classifiers to assess cognitive status. This approach requires training data annotations for cognitive status and has missing data issues in time windows.

Jakkula et al [20] considered the problem of anomaly detection based on temporal relationships. They expressed relationships between temporal events based on temporal logic, such as before, after, meets, overlaps, and contains, and used these to identify frequently occurring relationships between them. Adopting a probability-based model based on prior evidence from an inhabitant's history, they reported low-probability events as anomalies. The study acknowledges that hundreds of sensors must be used to identify temporal relations at a granular level. It also requires a large training data set that must be updated to capture changing patterns.

Using data from our own remote monitoring study, Palermo et al [21] developed a supervised long short-term memory network to analyze the risk of agitation episodes in people with dementia using environmental, physiological, and sleep data. They used weak learning and label augmentation to address noise and class imbalance. In another Minder study, Li et al [22] adopted a semisupervised machine learning approach to predict the risk of UTI in people with dementia using environmental and physiological data. A convolutional autoencoder was used to learn a representation of the unlabeled sensor data. The encoder was used to extract the corresponding features from a smaller set of positively labeled data, which were then used to train a supervised classifier—a probabilistic neural network with a fully connected layer. Although this model is robust and learns continually, it approximates sensor data using Lagrangian approximation, requires interpretability, and takes a generalized versus patient-specific approach to detecting UTIs.

In the study by Paudel et al [23], the authors used unsupervised graph-based anomaly detection to identify cognitive health decline in older adult residents living in smart homes. They transformed motion sensor data from raw sensor log files into individual activity graphs and performed anomaly detection based on the normative pattern derived from the minimum data length principle [24]. This study used cohort-wide thresholds instead of the users' own thresholds.

Visual Approaches

Visualization of activity density is another intuitive way of detecting anomalies in movement data. The study by Gupta et al [25] describes how unsupervised learning can be used to discover activity patterns in unlabeled data from passive infrared (PIR) sensors. In this work, user activity data were visualized and tracked through Uniform Manifold Approximation and Projection, whereas kernel density estimation was used for automatically extracting periods of dense sensor activity. Although Uniform Manifold Approximation and Projection plots are useful in informing daily patient-carer interactions, they are not readily interpretable, and this approach does not provide an anomaly score. Heat maps have also been used in conjunction with deep learning techniques to determine the probability of agitation- or UTI-related anomalies. In the study by Li et al [26], hourly heat maps based on raw sensor data were encoded via positional encoding to extract relevant time steps that were then passed into a long short-term memory model to extract relevant data and into an attention-based model to make predictions. This method uses supervised learning and, as is common with deep learning models, is computationally expensive and requires sufficient training data for accurate risk analysis and predictions.

Matrix Profile for Anomaly Detection

Research on real-world applications of Matrix Profile (MP)–based anomaly detection is scarce. Lin et al [27] used an early version of the MP to detect discords in electrocardiogram time series. More recently, researchers have used MP for web-based anomaly detection in IT operation time series [28]. In the study by Steenwinckel et al [29], researchers used an MP with knowledge-driven algorithms to create an interpretable system for sensor monitoring in the railway domain. Nieves Avendano et al [30] used MP with clustering for web-based anomaly detection and event prediction based on acoustic emission sensors that relay information about the mechanical conditions of a cold-forming manufacturing line. This method is robust to noise, missing values, and irregular sampling.

XSL•FO
RenderX

The CMP has been shown to be more flexible and effective than the MP in 2 curated non–health care web-based data sets where the authors showed how the CMP can be used to detect more subtle anomalies in addition to those detected by the MP [31].

In Figure 1 [11,32], we summarize the effectiveness of each technique framed in the context of remote health monitoring by evaluating the pros and cons of each technique presented in the survey literature. The CMP overcomes many of the drawbacks identified for distance-based methods and is well suited to remote health monitoring scenarios.

Figure 1. Suitability of anomaly detection techniques for remote health monitoring [11,32]. CMP: Contextual Matrix Profile.

Remote monitoring needs	Statistical	Distance-based	Reconstruction-based	Domain-based	Information-theoretic	Graph-based	CMP approach
Noisy data or labels	●	●	●	●	●	●	●
Data drift	●	●	●	●	●	●	●
Streaming data	●	●	●	●	●	●	●
Multidimensional data	●	●	●	●	●	●	●
Scoring anomalies	●	●	●	●	●	●	●
Unsupervised detection	●	●	●	●	●	●	●
High recall	●	●	●	●	●	●	●
Low volume of alerts	●	●	●	●	●	●	●
Point anomalies	●	●	●	●	●	●	●
Collective anomalies	●	●	●	●	●	●	●
Contextual anomalies	●	●	●	●	●	●	●
Parameter tuning	●	●	●	●	●	●	●
Complexity	●	●	●	●	●	●	●
Interpretability or explainability	●	●	●	●	●	●	●
Scalability	●	●	●	●	●	●	●

Methods

MP and CMP Preliminaries

MP Overview

The MP, detailed in the study by Yeh et al [33], is an unsupervised, state-of-the-art time series analysis technique that can be used for pattern detection, anomaly detection, time series segmentation, and change point detection. Its fast performance stems from the use of the fast Fourier transform for the *z*-normalized Euclidean distance computation. The algorithm is useful for both static data and incremental modeling of streaming values with limited slowdown on even very large and multivariate time series. In this section, we define the MP preliminaries relevant to anomaly detection in our smart home context.

MP Description

An MP P of time series t is a vector of the *z*-normalized Euclidean distances between each subsequence in an all-subsequence set A with its corresponding nearest neighbor or closest match within A (trivial matches excluded). Trivial matches are the set of subsequences around the query subsequence, which are likely to have a very small Euclidean distance from the query subsequence. This boundary is typically set to $m/2$, where m is the length of the subsequence.

Multidimensional MP

A k-dimensional MP of a multidimensional time series t with dimensionality d is a meta–time series that stores the *z*-normalized Euclidean distance between each subsequence and its nearest neighbor (the distance is computed using the k-dimensional distance function) [34]. In simple terms, the algorithm works as follows: (1) it stores the MP for each dimension (time series channel) in the subsequent rows of a 2D matrix, (2) the k-dimensional MP is computed by taking the average of the k lowest values in the columns of the matrix, and (3) the multidimensional MP is created such that row k ($0 \leq k < m$) contains the k-dimensional MP. For implementation, we refer the reader to the STUMPY library tutorial [35].

The issues with the direct application of the MP are outlined in Textbox 1.

Textbox 1. Issues with the direct application of the Matrix Profile (MP).

Direct application issues related to the MP

- The raw MP is noisy and does not give a clear indication of which discords are true anomalies.
- It is insensitive to amplitude variations and low in localization accuracy [28].
- The MP considers every subsequence for comparison with every other, which implies that the length of subsequence equals the level of granularity at which an anomaly may be identified. The two must be decoupled.
- An anomaly could be masked when its subsequence is close to another anomalous subsequence [28].
- The MP is hardwired to compute Euclidean distance. Although this has great advantages—complexity linear to the length of the time series, easy to implement, indexable, and parameter-free—it can also be sensitive to noise and exhibit misalignments in time [36].

XSL•FO
RenderX

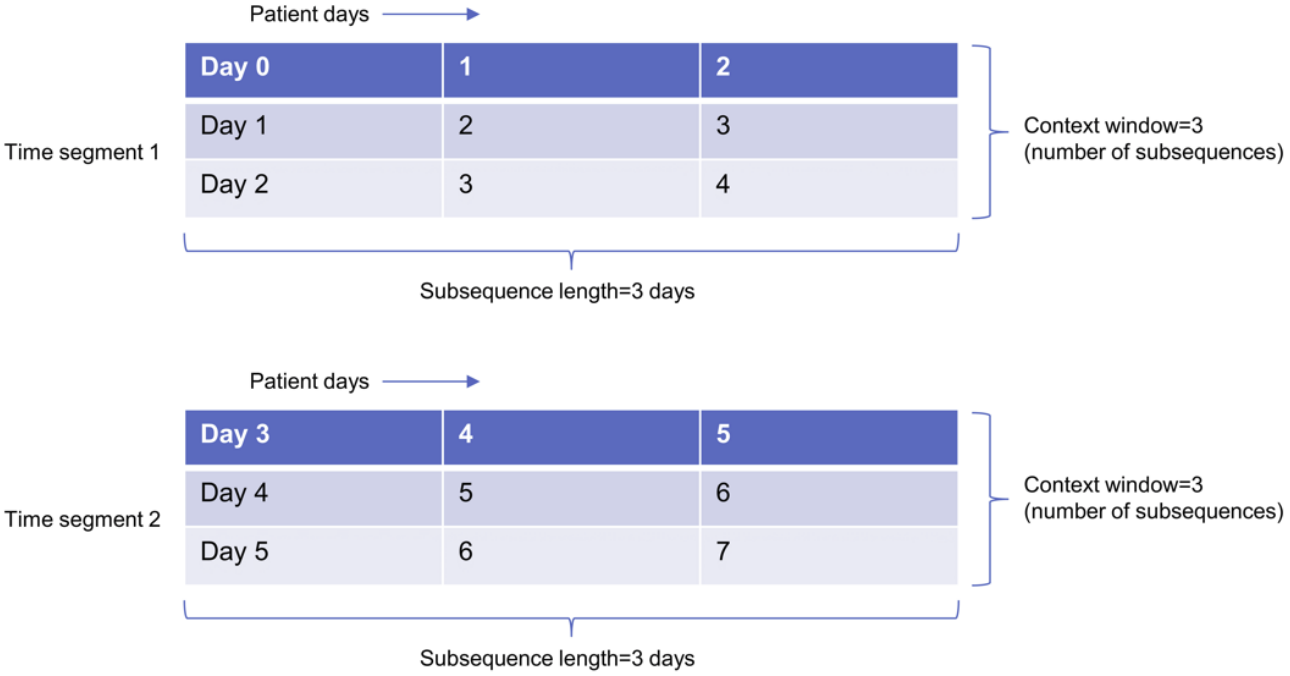
CMP Overview

The CMP is a new flexible time series analysis technique based on the MP [31]. The CMP derives its motivation from the distance matrix calculations that are used to compute the MP. This section provides details on the CMP.

Context Window

It is the number of subsequences in a single time segment or region of interest. Given a patient data set, using a context window of 3 and a subsequence length of 3 (with no subsequences omitted), the patient data will be grouped into the time segments shown in Figure 2.

Figure 2. Contextual Matrix Profile contexts or time segments are blocks of time comprising a set of subsequences. Each context shown here is made up of 3 subsequences (context window=3), each subsequence being 3 days in length. We assigned anomaly scores to contexts instead of days.



Context (or Time Segment)

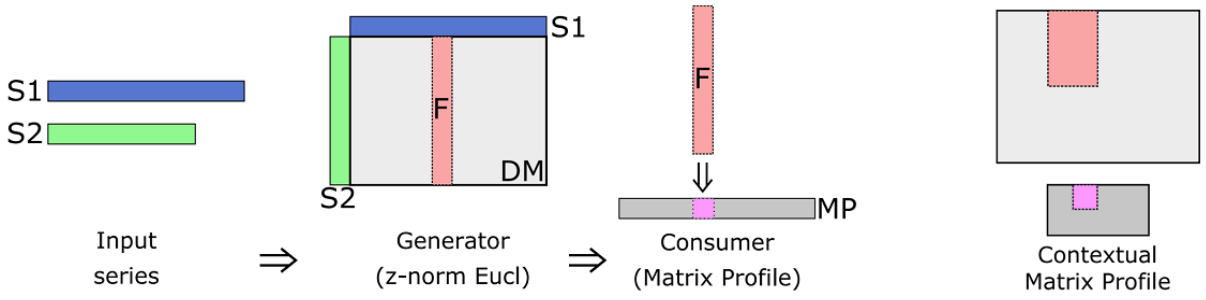
It is a single time segment with a size equal to the context window and containing subsequences of length defined by the user. One cell in the CMP represents 1 “context” or time segment.

CMP Description

It is a configurable, 2D version of the MP that tracks the minimum distance between each context of subsequences in user-defined regions of the time series. First, the user (optionally) defines regions of interest for a given time series. They then determine the subsequence length and context window size. For instance, for a subsequence length of 3 days and a context window size of 3 days, the time series is divided into contexts, as shown in Figure 2. The CMP is formed by comparing the *z*-normalized Euclidean distance between each subsequence in one context and every subsequence in another context and selecting the minimum distance, which forms 1 cell in the CMP. Figure 3 highlights the difference between the MP and CMP. The MP comprises the column-wise minimum values in the distance matrix, whereas the CMP is created by taking the minimum over rectangular areas.

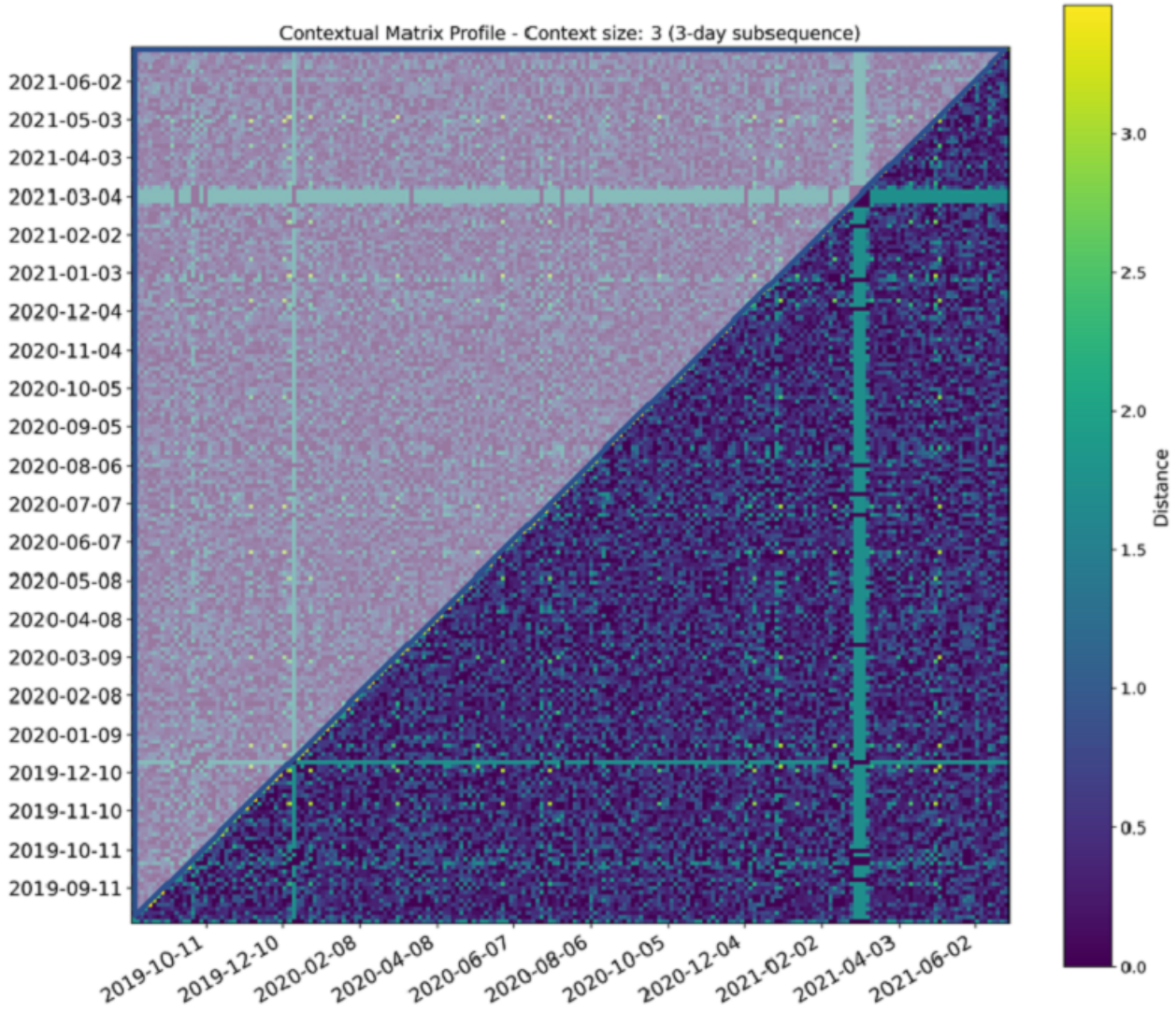
The application of the MP idea to blocks of data instead of individual subsequences serves to aggregate and denoise the distance computation and extract useful patterns. Figure 4 shows the CMP for the late-evening daily bathroom activity for one of our patients. It serves as a visual overview of the consistency of activity and any break points.

Figure 3. Matrix Profile versus Contextual Matrix Profile.



XSL•FO
RenderX

Figure 4. Each cell color codes the minimum distance between the time segments on the x- and y-axes. Green bands indicate anomalous activity or consistently large minimum distance from other time segments. The Contextual Matrix Profile is symmetric around the diagonal line.



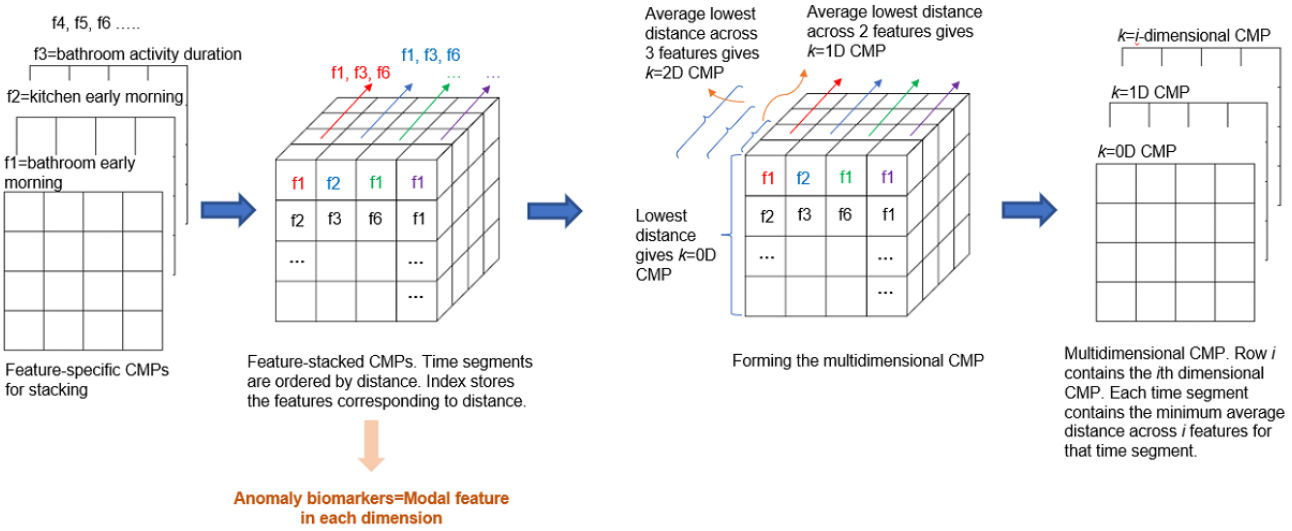
Multidimensional CMP

We developed the multidimensional CMP based on the principle of a multidimensional MP. A *k*-dimensional CMP of a multidimensional time series *t* with dimensionality *d* is a meta–time series that stores the minimum z-normalized Euclidean distance between any subsequence in one context and any subsequence in another context, with the distance computed using the *k*-dimensional distance function, which is explained in the study by Yeh et al [34]. The algorithm works as follows. First, we stack the feature-specific 2D CMPs to obtain a 3D array. We then sort the array in ascending order using this feature dimension. This gives, for each context, the minimum distance values sorted in increasing order by feature. Now, we apply the method in the study by Yeh et al [34] to obtain the *k*-dimensional CMP. For *k*=0 (ie, 1D CMP), we query the first row of the *k*-dimensional CMP. This provides the lowest nearest-neighbor distance for each context based on a single feature. The lowest-scoring feature for a context may be different from the lowest-scoring feature for a different context. Similarly, for *k*=1 (or 2D CMP), we query the second row of the *k*-dimensional CMP, which, for each context, provides the lowest average distance based on 2 features. Again, the 2 lowest-scoring features for one context may be different from those for another context. The maximum value of *k* is the number of features minus 1 (*k* is zero-based).

The multidimensional CMP (Figure 5) is key to anomaly detection in our multidimensional data. The CMP offers advantages over the original MP (Textbox 2).

XSL•FO
RenderX

Figure 5. Multidimensional Contextual Matrix Profile (CMP) formation involves stacking feature-specific CMPs and then arranging each cell (time segment) in order of distance. The final multidimensional CMP is formed so that each cell in row i contains the average of the i+1 lowest distances for the cell.



Textbox 2. Advantages of the Contextual Matrix Profile (CMP).

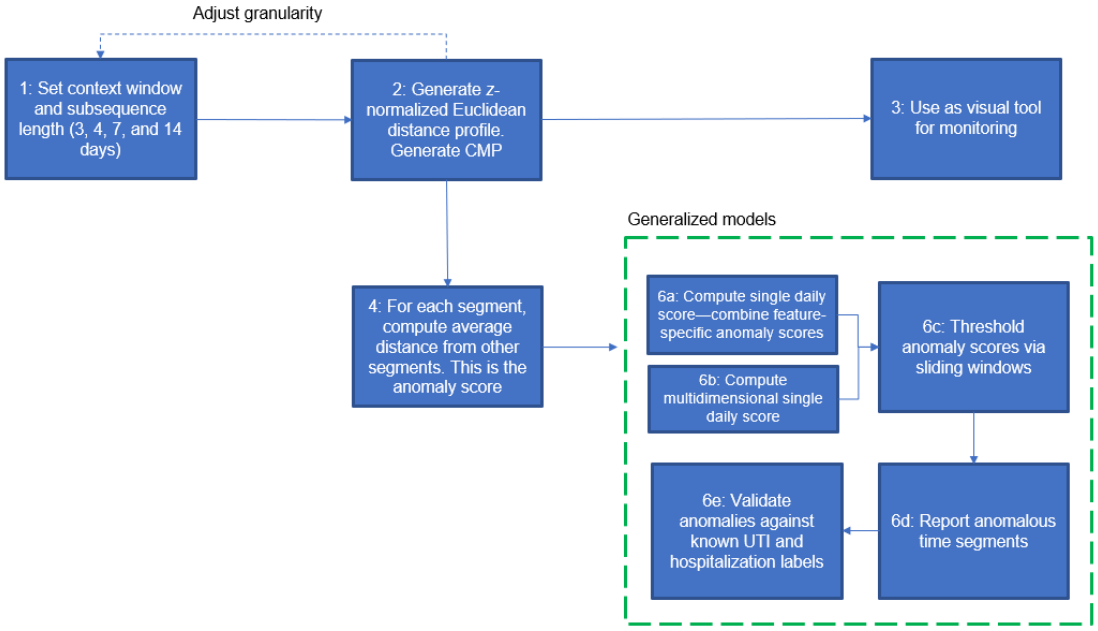
CMP advantages

- It compares distance over a context instead of at a subsequence level, which is important for denoising the time series.
- The context size is configurable. In addition, the CMP allows for specific regions to be defined to detect patterns and anomalies, and the distance matrix need not be covered in its entirety.
- Other distance measures in addition to Euclidean distance can be used.
- The CMP offers an intuitive way of visualizing time series window regions and detecting anomalies.
- Anomalies cannot be easily masked, even if another similar anomaly has occurred elsewhere in the time series.

Anomaly Scoring With the CMP

We used the anomaly detection pipeline (Figure 6) described in Textbox 3.

Figure 6. Contextual Matrix Profile (CMP)–based anomaly detection pipeline. UTI: urinary tract infection.



XSL•FO
RenderX

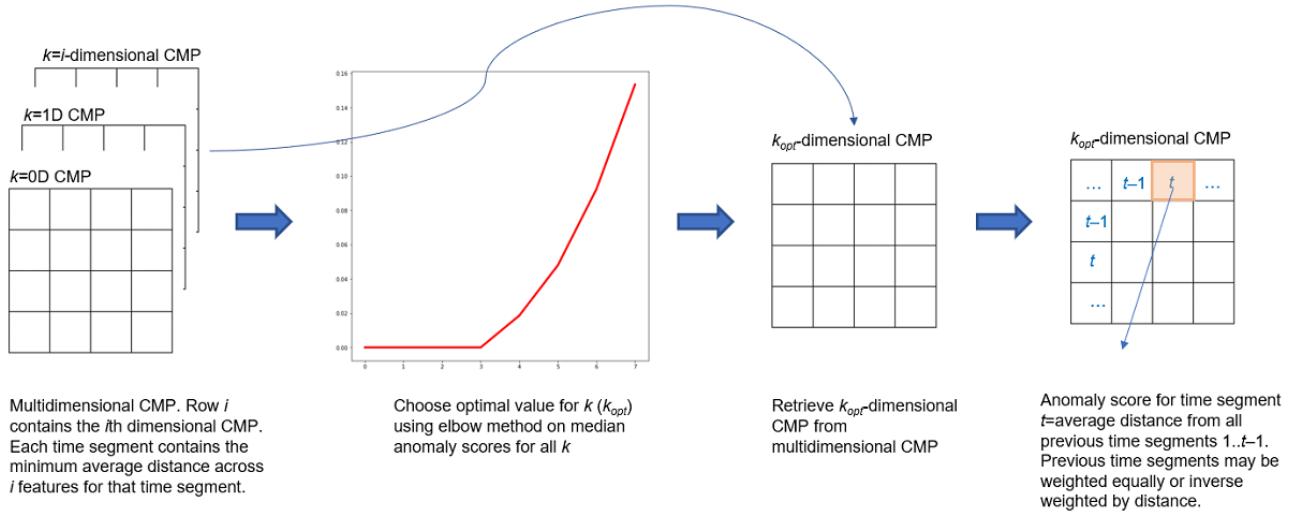
Textbox 3. Anomaly detection pipeline.

Pipeline for anomaly detection

1. We decided on a suitable context window. We used a context size of 3 and a subsequence length of 3 days taking into consideration the need for maximum granularity, denoising, explainability, and time taken for the onset of an anomaly.
2. For each patient time series, we generated the z-normed Euclidean distance matrix for a self-join and then the Contextual Matrix Profile (MP) based on our context window setting.
3. We used the CMPs directly or adapted them for visualization and monitoring purposes.
4. We computed the anomaly score for each context. This is the average distance between the current context and contexts in the past. This step was repeated for every time series to obtain feature-specific anomaly scores.
5. The anomaly scores were used as inputs in different machine learning models trained for specific health events. This will be explored in future work.
6. The models were prepared to obtain a single-valued score for each time segment. We evaluated the following methods:
    - Combining feature-specific anomaly scores based on the sum of scores, median of scores, mean of scores, maximum of scores, and entropy-based weighting (the scores were combined based on the entropy of the underlying time series using inverse weighting; greater entropy implies lower weighting of the anomaly score obtained from using that time series). Two types of entropy measures were used:
        - Approximate Entropy (ApEn): ApEn approximates the exact regularity statistic Kolmogorov-Sinai entropy and reflects the predictability of a time series by exploring repetitive patterns in the data. It is applicable to noisy data sets [37]. It relies on the Heaviside function to define the similarity between 2 patterns. ApEn generates a unitless number from 0 (perfectly periodic) to 2 (noisy) [38].
        - Fuzzy Entropy (FuzzyEn): This also uses the Heaviside function, although similarity is evaluated by a fuzzy function that computes a membership coefficient ranging from 0 to 1. Consequently, in addition to the selection of N (length over which to compute entropy), m (subsequence length), and r (tolerance in terms of the number of SDs), FuzzyEn requires a fourth parameter, n, the gradient of the boundary of the exponential function used to assess similarity [39]. FuzzyEn provides a graded similarity instead of binary similarity between parts of the time series [37].
    - Multidimensional CMP-based scoring: We used the multidimensional CMP to generate the multivariate anomaly score for each context using 2 different settings for k:
        - k=auto: Here, we considered the optimal value of k when predicting true likely anomalies for a patient (Figure 7). To do this, we used the elbow method on each patient's multidimensional CMP. Specifically, we computed the median distance in each of the k-dimensional CMPs for the patient and used the "kneedle" algorithm to automatically find the optimal value of k at which the inflection point occurred [40]. We then chose this optimal k row from the overall k-dimensional CMP to use this to extract the single-valued patient anomaly scores for each context. Once the "optimal" CMP was obtained, we scored each context in 2 ways:
            - Distance-weighted multidimensional CMP scoring: The anomaly score for a context was calculated as the inverse-weighted average of its nearest-neighbor distance from previous contexts. Thus, if a context is 3 hops in the past from the current context being scored, its distance is given one-third weight when calculating the anomaly score for the current context.
            - Equal-weighted multidimensional CMP: The anomaly score for a context was calculated as the simple average of its nearest-neighbor distance from previous contexts.
        - k=1: We took the CMP that is based on the top 2 features for each context.
    - We performed sliding window thresholding (7-, 14-, 21-, 30-, 60-, and 90-day windows) on the single context score using robust z, IQR, and quantile-based methods to predict true likely anomalies and report the best results.
    - The predicted anomalies were then "soft" validated against the anomaly labels available in the data set to compute recall.

XSL•FO
RenderX

Figure 7. Multivariate anomaly scoring. Starting with the multidimensional Contextual Matrix Profile (CMP), we chose the optimal dimension for each patient by taking the median of their scores in each dimension and then selected the dimension at the inflection point. This optimal dimensional CMP was then used for distance-based anomaly scoring.



Data Set Description and Preprocessing

Data Set

The data for our study came from an ongoing real-world remote health care monitoring study (the "Minder" study mentioned in the *Introduction* section) of 102 people living with dementia at home and supported by their carers in England, United Kingdom. This includes 51% (52/102) men (mean age 81.94, SD 6.34 years) and 49% (50/102) women (mean age 80.80, SD 15.76 years). Movement data are captured via PIR sensors installed in different parts of the home—hallway, bathroom, bedroom, lounge, and kitchen—that are triggered passively based on movement throughout the day. There are door sensors, smart plugs for appliances, light and temperature sensors, a sleep mat, and physiological data recorders as well. We considered only PIR data in this study as they are the least missing, most reliable, and available with the finest granularity across the cohort. Physiological data are currently self-reported by the person living with dementia or their carer once or twice a day and with greater missingness, which would require imputation. Sleep data are sparse for this cohort.

In our study, we focused on the 15 patients with dementia who had had at least one clinically validated incidence of UTI. This included 53% (8/15) men (mean age 85.13, SD 5.57 years) and 47% (7/15) women (mean age 82.86, SD 6.79 years). Of these 15 patients, 7 (47%) had also experienced ≥1 hospitalization event. Altogether, we had 31 UTI and 10 hospitalization labels across a total of 9363 patient days, making up approximately 0.44% (41/9363) of the overall data set. The UTI labels were manually annotated after validation by the clinical monitoring team using urine samples from patients. However, it is worth noting that older adult patients often present with atypical symptoms, making the differentiation of asymptomatic bacteriuria from symptomatic UTI challenging [6]. Moreover, the start time of UTI and the duration of symptoms are not clearly defined. The list of hospitalization events was collated based on information from general practitioners. It included the date of hospitalization and, in some but not all cases, the cause of hospitalization.

Preprocessing

Daily Aggregation

Household movement data captured via PIR motion sensors were first aggregated daily to reduce noise, as hourly counts can vary widely from one day to the next, and the high granularity and variation make anomalies less discernible. We ignored consecutive firing events from the same sensor, considering the first firing event to compute the duration at the previous location and the last firing event to compute the duration at the current location. Any consecutive sensor firings from the same sensor between the first and last firing were ignored, thus reducing redundancy and noise.

Feature Engineering

To capture different types of movement-related information, we calculated the features outlined in Textbox 4 for the daily activity data across the various locations—bathroom, bedroom, kitchen, lounge, and hallway.

XSL•FO
RenderX

Textbox 4. Movement-related features.

Feature and description

- Location count: this is the daily count of sensor firings for each location.
- Location early-morning count: this is the count of sensor firings between midnight and 6 AM on the current day.
- Location late-evening count: this is the count of sensor firings between 6 PM and midnight on the previous day.
- Location duration: this is the daily total number of minutes spent at each location.
- Location hourly movement change: this is the Wasserstein distance between the hourly sensor distribution at a location on the previous day with that on the current day; a larger Wasserstein distance implies a greater change in hourly pattern from one day to the next; this measure is robust to different motion densities across patient households. The Wasserstein distance or "earth mover" distance is a single explainable metric that measures the approximate *minimal work* required to move between 2 probability distributions, where "work" can be loosely defined as the product of how much of the distribution's mass moves and the distance by which it must be moved [41]. Unlike other measurements such as L2, Kullback-Leibler divergence, and Jensen-Shannon divergence, the Wasserstein distance is sensitive to geometry [42].

Feature Selection

Similar to the study by Skubic et al [43], we applied the methods outlined in Textbox 5 to select the features for anomaly detection.

The simplified list of features included bathroom Wasserstein distance, hallway Wasserstein distance, lounge Wasserstein distance, bathroom early-morning and late-evening activity, kitchen early-morning and late-evening activity, bedroom early-morning and late-evening activity, bedroom activity duration, and bathroom activity duration.

Textbox 5. Methods for selecting features for anomaly detection.

Anomaly detection feature selection methods

- Domain knowledge: the study by Pevný [44] showed that detectors using only features that explain anomalies had equal or better performance than detectors using all features. Erratic bathroom activity can strongly suggest urinary tract infection [45], and therefore, we captured daily changes in bathroom activity. Similarly, disturbed sleep, agitation, and wandering are common characteristics in patients with dementia [46,47]. Hence, we included daily changes in the hourly distribution of bedroom, hallway, kitchen, and lounge activity. These features help capture unusual daytime and night-time activity across locations and follow recommendations by clinical researchers in a similar study supporting the modeling of health decline with behavioral biomarkers [43].
- The significant online discords technique was used to find the common features that are associated with the highest median recall value for urinary tract infection and hospitalization using cross-validated data from our patient cohort [48].
- We eliminated redundant variables based on the correlation between the features.
- We also eliminated duration-related features for communal spaces such as hallways, living rooms, and kitchens, where distinguishing between patient and carer activity is difficult as of yet.
- The variables that were robust to differences in activity levels across households were retained.

Experiments

We conducted our experiments on the household movement data of 15 patients selected from the ongoing Minder study, which had 31 UTI and 10 hospitalization labels across a total of 9363 patient days. All experiments were run on a 64-bit Intel i7-8700K central processing unit, 3.7 GHz Windows 10 machine with 32 GB of RAM.

Our experimental settings are listed in Table 1. For each anomaly-scoring model, we experimented with every combination of window size, IQR threshold, robust *z* threshold, and quantile threshold and reported the best results obtained.

XSL•FO
RenderX

Table 1. Experimental parameters considered in this study.

Setting	Values
Context window	3
Subsequence length (days)	3
Window sizes for sliding window thresholding (days)	7, 14, 21, 30, 60, and 90
IQR threshold	1.0 and 1.2
Robust *z* threshold	1.65, 1.8, 3, and 4
Quantile threshold	0.95, 0.96, 0.97, and 0.98
Entropy-based methods	
N (data size)	500
r (SD tolerance)	0.2
m (subsequence length)	7
Soft buffer for label validation (days around actual label)	−10 to +7

Model Evaluation

We used the data from our 15 patients to evaluate multidimensional CMP-based anomaly scoring relative to univariate methods and 3 other popular modern, parameter-free, and interpretable methods in the literature: Angle-Based Outlier Detection (ABOD), Copula-Based Outlier Detection (COPOD), and Lightweight Online Detector of Anomalies (LODA). To be suitable for use in an unsupervised, streaming scenario, we used only historical data at each time point.

ABOD measures the variance of the angle (cosine) spectrum of the data points weighted by the corresponding distances. ABOD works on the principle that if the spectrum of the observed angles for a point is small, other points will be positioned only in certain directions. This means that the point is positioned outside of some sets of points that are grouped together, implying that the point is an outlier [49]. COPOD is inspired by copulas for modeling multivariate data distributions. COPOD first constructs an empirical copula and then uses it to predict the tail probabilities of each given data point to determine its level of "extremeness." The outlier scores produced by COPOD measure the likelihood of a point relative to the other points in the data set. The method outputs a "dimensional outlier graph" that provides insights into outlier subspaces or features for a given outlier point [50]. LODA comprises a collection of *k* 1D histograms, each approximating the probability density of the input data projected onto a single projection vector. Projection vectors act to diversify individual histograms, which enables the ensemble system to improve the performance of a single detector. The complexity of LODA is linear with respect to the number of training samples and the dimension of the input space [44].

We used 3 thresholding criteria for scores (Textbox 6).

To determine how competitive CMP-based anomaly scoring is in identifying anomalies in real-world remote monitoring data for patients with dementia, we report the measures shown in Textbox 7 for each model.

An anomaly is assumed to be correctly identified if the predicted date is within the soft buffer of the labeled date of anomaly. For transparency, we report both the average recall and patient-wise recall. From a clinical perspective, this measure is a direct indication of a model's effectiveness.

When choosing between models, a clinician will likely choose a model with a higher average recall, as the cost and inconvenience of false alerts in our scenario are considerably less than the cost of missing a real anomaly.

Textbox 6. Thresholding criteria for scores.

Criteria for thresholding of scores

- The robust *z* thresholding or Median Absolute Deviation method is less influenced by outliers and is used to calculate a modified *z* score that quantifies the anomaly score in terms of SD units away from the median [51].
- Tukey or IQR thresholding uses the IQR of anomaly scores in the sliding window as the basis for thresholding. Any value greater than the third quartile+ *x* times IQR is deemed anomalous, where *x* is the IQR threshold from Table 1.
- Quantile-based thresholding uses a fixed percentile of anomaly scores as the basis for thresholding taken from Table 1.

XSL•FO
RenderX

Textbox 7. Model evaluation.

Model evaluation measures

- Number of patients with >33% recall: given that the average patient had only 3 validated anomalies, we ranked the models based on how many patients had greater than one-third of their anomalies correctly identified. This makes it transparent whether the model is just effective for a small proportion of patients or across the cohort.
- Average recall: this is the average percentage recall across the 15 patients, where recall=true positives or all known anomalies in the data set.
- Average number of anomalies detected: to minimize false alerts made to the clinical monitoring team, lower is better.
- Average recall percentage versus anomalies raised: according to the study by Pimentel et al [11], effectiveness in novelty detection is based on the detection rate and the false alarm rate. The best model will demonstrate high recall together with a low number of anomalies raised.
- Precision: here, *precision* has little meaning, as outliers may result from different types of health indicators, sensor failures, visitors, pet activity, or rare unusual activities by the patient or carer, which are not labeled in our data set. Although we still report this metric, relative precision across methods is more meaningful.

Digital Biomarkers

Digital biomarkers are consumer-generated physiological and behavioral measures collected through connected digital tools that can be used to explain, influence, or predict health outcomes [52]. The Food and Drug Administration-National Institutes of Health "Biomarkers, EndpointS, and other Tools" classification for traditional biomarkers classifies their use into the following categories: susceptibility or risk determination, diagnostic use to detect and confirm the presence of a condition of interest, monitoring of the status of a condition, prognostic use to identify likelihood, recurrence or progression of a condition, predictive use, and measurement of response through exposure to a medical product or agent [53]. We envisage these biomarkers of anomalies to be used for susceptibility determination and assistance with diagnosis, prognosis, and prediction of UTI or another adverse clinical event.

The creation of the multidimensional CMP involves the intermediate step of combining feature-specific CMPs such that each context is arranged in ascending order of the feature-wise nearest-neighbor distance. This implies that if we simply keep track of the ordered set of features for each context in the ordered stacked CMP, we can discover the most common contributing feature in each of its dimensions. The modal feature in the 0th dimension will be the most important biomarker associated with the patient's anomaly score. The modal feature in the first dimension will be the second most important biomarker and so on. Subsequently, by looking across the ordered stacked CMPs for the entire cohort, we can determine the generalized top *k* important biomarkers.

Ethics Approval

This study received ethics approval from South East Coast Surrey National Health Service Research Ethics Committee (Health Research Authority); Technology Integrated Health Management Research Ethics Committee Reference: 16/LO/1802; Integrated Research Application System ID: 211318.

Results

Model Evaluation

We report the best results for each type of univariate and multivariate model (Table 2).

All models could correctly identify more than one-third of the known anomalies for two-thirds of the patients in the study. Of these, the multidimensional CMP with equal-weighted context (at window size=7 days, robust z=1.65, and k=1) yielded >33% recall for 100% (15/15) of the patients. Other CMP-based methods showed similar recall for up to 93% (14/15) of the patients. This highlights the strong support for multidimensional CMP as an anomaly detection tool for this cohort.

We also measured how many anomalies were raised by each of the models across the 624 average patient days in our study. As shown, the maximum number of alerts raised by any of the CMP-based models was only approximately 34 or 5.4% (34/624) of patient days. Our best-performing CMP model raised approximately 32 alerts, which is, on average, 5.1% (32/624) of patient days. Note that there were, on average, 3 labeled anomalies in our data set per patient; however, as emphasized previously, the annotated anomalies covered only UTI and hospitalization, and our models were designed to pick up on any anomalous activity.

The average recall, when viewed together with the total detected anomalies, provides a holistic view of performance, as it is easily possible to obtain a top-performing model by identifying an extraordinarily high number of anomalies. The overall best model is one that demonstrates high recall but a low number of raised anomalies. It is clear that the multidimensional CMP with equal-weighted contexts at window size=7 days, robust z=1.65, and k=1 offers the best-balanced performance, raising only 32 alerts over a 624-day patient journey on average. ABOD yields relatively low recall, whereas LODA and COPOD yield high recall but with a higher number of alerts raised than our best-performing model.

XSL•FO
RenderX

Table 2. Model performance (N=15).

Model	Patients with >33% recall, n (%)	Anomalies raised, mean	Recall (%), mean	Precision, %[a]
LODA[b] (w=7; IQR 1.2)	14 (93)	37.8	85.7	6.2
Sum of CMP[c] scores (w=7; quantile 0.97)	14 (93)	33.1	84.7	7.0
Mean of CMP scores (w=7; quantile 0.97)	14 (93)	33.1	84.7	7.0
Equal-weighted multidimensional CMP (w=7; *k*=1; robust *z*=1.65)	15 (100)	32.1	84.3	7.2
COPOD[d] (w=7; quantile 0.95)	13 (87)	36.8	79.1	5.9
ABOD[e] (w=21; quantile 0.95)	13 (87)	30.0	77.7	7.1
Distance-weighted multidimensional CMP (w=14; *k*=0; robust *z*=1.65)	14 (93)	33.7	76.7	6.2
ApEn[f]-weighted CMP scores (w=7; quantile 0.97)	12 (80)	29.1	69.9	6.8
Median of CMP scores (w=7; quantile 0.97)	12 (80)	30.8	68.4	6.1
Fuzzy entropy–weighted CMP scores (w=7; quantile 0.97)	10 (67)	27.7	65.5	6.5
Maximum of CMP scores (w=7; quantile 0.97)	10 (67)	24.8	57.9	6.4

[a]We have mentioned previously that it is more meaningful in this context to look at relative precision across methods and not at absolute precision.

[b]LODA: Lightweight Online Detector of Anomalies.

[c]CMP: Contextual Matrix Profile.

[d]COPOD: Copula-Based Outlier Detection.

[e]ABOD: Angle-Based Outlier Detection.

[f]ApEn: Approximate Entropy.

Digital Biomarkers

As seen previously, the multidimensional CMP for a patient can be used to discover the important digital biomarkers of anomalies. In Figure 8, we show the magnitude of the contribution of significant features toward the anomaly score across the cohort.

We discovered that early-morning (midnight to 6 AM) bathroom activity was the single largest contributor to the anomaly score by a wide margin, with a median value of approximately 30% for this cohort. This validates the findings in the literature that unusual bathroom activity is a clinically significant feature of UTIs [12,54], which comprises three-quarters of the anomalies in our labeled data set. Patient-level investigation showed this to be the top biomarker for 60% (9/15) of the patients. Late-evening (6 PM to midnight) bathroom activity also had a contribution of 12%. Both factors correlate with sleep disruption, which is commonly seen in people living with dementia. Unusual bedroom and kitchen activity in the early hours of the morning are also among the significant contributors to anomaly scores, pointing to wandering and disturbed sleep seen in dementia.

The multidimensional CMP also provides intuitive insights into patient-specific anomalies. Figure 9 shows the anomaly scores associated with 2 patients, ordered by the median anomaly score.

For patient JYN9, unusual early-morning kitchen activity was the prime biomarker of anomalous activity, where we also see the largest variance in anomaly scores. For patient SFAV, unusual bedroom activity was the largest contributor to their anomaly score. These figures indicate different anomaly patterns in the 2 patients, presumably agitation and wandering in the first patient and sleep disruption and shifting bedroom activity over time in the second patient. We can envisage an anomaly detection dashboard to provide such insight to clinicians to enable them to target interventions as needed.

We can also use the standardized anomaly scores to look at a cross-patient view (Figure 10), where we see the cross-cohort variation in multivariate anomaly scores using the patients' own optimal *k*-dimensional scores. It would be interesting to investigate patient differences in relation to their cognitive scores.

XSL•FO
RenderX

Figure 8. Top contributing digital biomarkers of anomalies. Early-morning bathroom activity had the largest median contribution of approximately 30% to the overall anomaly score.

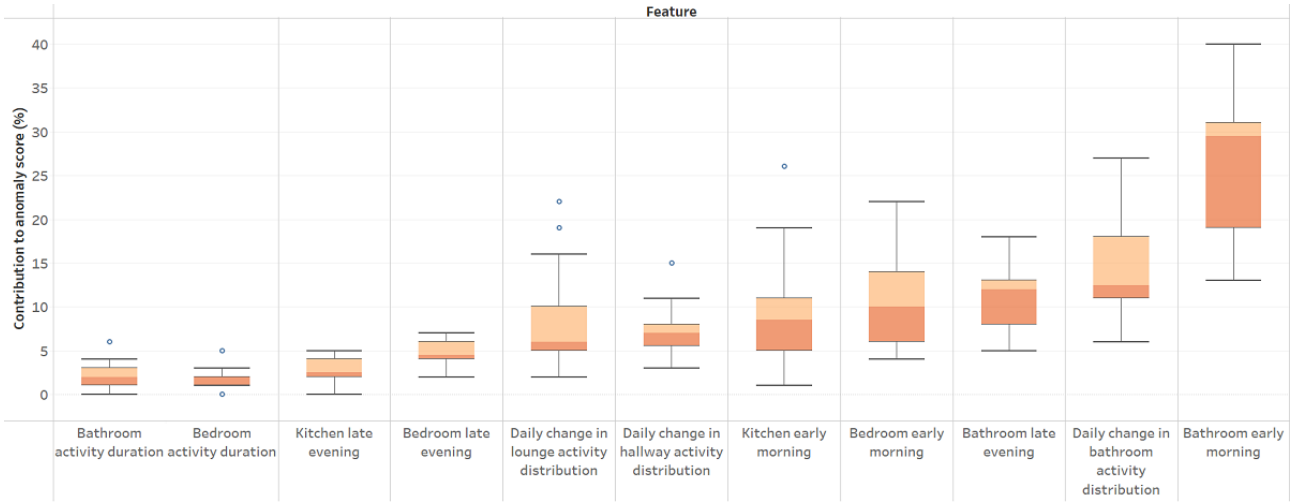
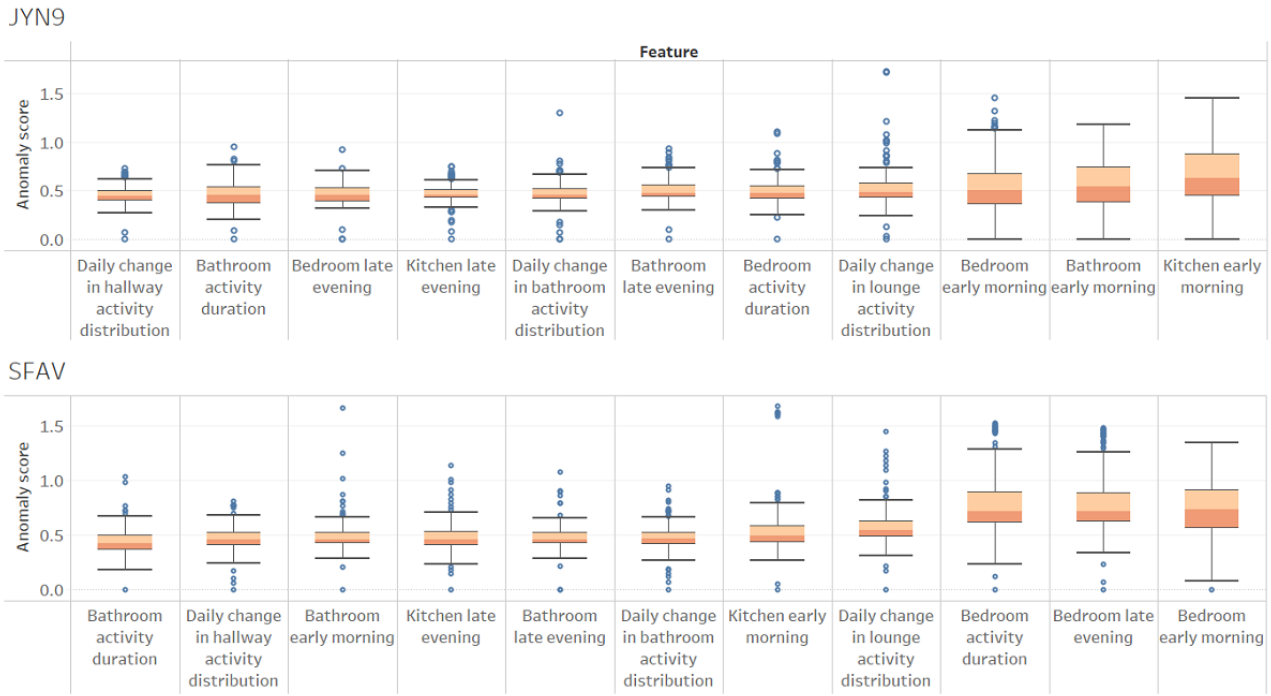
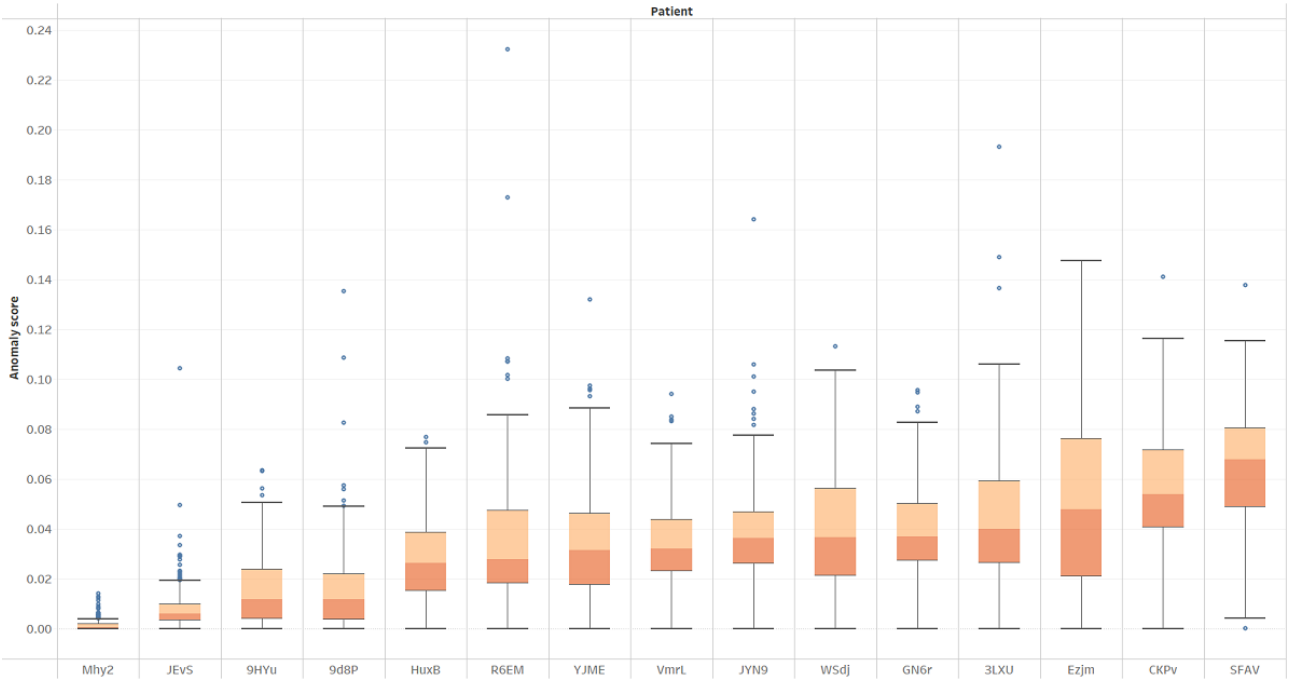


Figure 9. Univariate anomaly score distribution for 2 patients.



XSL•FO
RenderX

Figure 10. Multivariate anomaly score distribution for all patients.



Discussion

Principal Findings

Early identification of anomalies in patients living with dementia provides a window of opportunity for early intervention before a major health event occurs. This implies improved health outcomes, reduced health care costs, continued independence, and better quality of life [43]. In this study, we developed an MP-driven approach for anomaly detection and evaluated its use in a real-world study of sensor-based remote monitoring of people with dementia. We developed the multidimensional CMP to model patient household activity from sensor data and used the average Euclidean distance between activities in time segments as the basis for generating a single anomaly score. The CMP-based approach overcomes the issues with traditional distance-based anomaly detection techniques, namely, degradation because of noise, high alert rate, and identification of local novelty. Our experiments show that multidimensional CMP-based anomaly detection performs better than other comparable fast, modern, exact, and parameter-free unsupervised techniques for anomaly detection. It is well suited to real-world remote monitoring data characterized by noise and incomplete labeling and is additionally useful as a visual tool for operational monitoring, also lending itself to the discovery of personalized and cohort-wide digital biomarkers. The personalized model for each resident comes uniquely from their own sensor data patterns [43]. These aspects make CMP-based anomaly detection clinically significant, interpretable, and immediately usable, freeing up clinicians' valuable time from having to annotate patient activity. The CMP is domain agnostic and can be easily extended to different types of health care data and domains. To the best of our knowledge, our work is the first real-world use case of CMP in health care anomaly detection.

Our experimental parameters were selected to be clinically relevant. A context window of 3 and a subsequence length of 3 were selected for maximum granularity, noise resistance, and suitability for anomalies such as UTI, where a 3-day pattern is more likely to throw up anomalous behavior than a more granular daily analysis; it typically takes 3 to 8 days for a UTI to present [55]. The context window and subsequence length can be easily configured to investigate anomalies at different levels of granularity, such as 7 days, 2 weeks, or 1 month. Similarly, we used 7-, 14-, 21-, 30-, 60-, and 90-day windows to threshold anomaly scores. Time segments such as these follow typical human patterns of behavior and are easily understood by clinicians. Threshold values for IQR, robust z and quantile-based thresholding, and entropy parameters were chosen to mirror values widely used in the literature. The soft buffer for label validation (−10 days to +7 days of actual anomaly label) reflects the issue of weak labeling because of noise and inaccuracy because of manual labeling, the time it takes for a UTI to develop and be clinically diagnosed, and the need to catch anomalies early. Dau and Keogh [56] used a similar evaluation technique for weakly labeled data. We chose ABOD, COPOD, and LODA for comparison with CMP-based methods as they are similarly high-performing, parameter-free, interpretable, unsupervised anomaly detection techniques relevant to a streaming data scenario such as remote health monitoring.

Our evaluation methods were also designed to be simple, transparent, and clinically meaningful. A good method must demonstrate high cross-cohort average sensitivity but also high sensitivity for individual patients while raising minimal alerts. Therefore, we report the overall sensitivity, patient-wise sensitivity, average number of anomalies raised, and recall versus anomalies raised, which provides a rounded measure of performance. Although we report the precision for each model, it must be noted that we only considered 2 types of labels—UTI and hospitalization—whereas our models identify all types of anomalies in household movement data, many of which cannot be validated using existing labeled data. For this reason, a low

XSL•FO
RenderX

absolute precision is to be expected, and the relative precision offers a better indication of the cross-model performance in our study. Moreover, the 5% alert rate is an acceptably low rate as these alerts may have critical implications for the health of people with dementia. This was also the approach taken by Rantz et al [12], and our clinical care teams already conduct weekly check-ins with the patients.

Our results show that for our top-performing models, the optimal sliding window size for thresholding is 7 to 14 days (ie, 1-2 weeks). This makes intuitive sense as an "anomaly" regarding human behavior can be perceived as a break in their recent routine. This was also clinically validated in the study by Skubic et al [43], where clinicians recommended a 2-week moving baseline for sensor data comparison and thresholding to balance capturing sudden and gradual health changes. A short sliding window has the added advantage of being robust to variations in patient characteristics and environmental conditions. However, the ABOD technique is highly sensitive, requires sufficient data to capture true outliers, and performs best with a minimum look-back of 21 days. This behavior of increasing the sample for better performance of ABOD was also validated in the study by Domingues et al [57].

We make 3 striking observations. First, the top-performing model in terms of balancing cohort-wide sensitivity and raised anomalies was based on k=1 (ie, it considers only the top 2 contributing features for a patient). This implies that a reliable anomaly detection model based on patient activity can be simple, lightweight, easily interpretable, and generalizable. Second, univariate models derived from combining feature-specific CMPs via simple aggregation (ie, sum and mean of feature-specific anomaly scores) achieve both high recall and low volume of alerts. They are, in fact, closer in performance to the best-performing multivariate CMP model than more established high-performing models such as LODA, ABOD, and COPOD and other complex ways of combining univariate scores such as entropy-weighted scores. This shows once again that simple, interpretable models can generalize and perform competitively. Third, it is surprising that an equal-weighted time segment–scoring approach achieves considerably better recall than distance-weighted time segment scoring. We would expect that by emphasizing more recent time segments over past time segments, we might obtain an anomaly score that is reflective of a true anomaly in the current time segment. However, this appears to not be the case in this study. We aim to explore different ways of weighting previous time segments to confirm whether this behavior was because of the specific distance-weighting logic used or a more general finding.

Digital biomarkers are an incredibly useful artifact of our method. They tell us what kind of household activity was responsible for the anomaly at a specific period. Furthermore, looking across a patient's timeline, we can find the single most common activity or feature that contributed most frequently to the anomaly score in the time segments overall (ie, a digital biomarker of their anomalous behavior). We discovered that cohort-wide, early-morning (midnight to 6 AM) bathroom activity was the most common digital biomarker of anomalous behavior (9/15, 60% of the patients), followed by late-evening bathroom activity and early-morning bedroom and kitchen activity. These findings quantitatively validate observational studies of patients with dementia, where agitation, wandering at unusual hours, and unusual bathroom activity, particularly early-morning and late-night bathroom activity, were observed, especially in patients with dementia experiencing a UTI [14,20,45,46]. Finally, our method can provide a ranking of digital biomarkers for anomalies at the time segment, patient, and cohort levels. This outcome makes CMP-based anomaly scoring independently useful for clinical monitoring and for querying and validating digital biomarkers.

There are a few notable differences between this work and existing published research based on the Minder study. First, published works have used a variety of supervised and semisupervised machine learning methods to detect or predict targeted health conditions such as agitation and UTI (one study used unsupervised learning to isolate anomalous movement patterns via clustering). As such, the models were trained with data from the subset of patients clinically validated to have the specific health condition in their trajectory. In contrast, our work evaluated a lightweight, unsupervised, and parameter-free approach to detect general anomalies based on household activity data. It requires no training data but is validated on data from patients who have experienced one or more UTIs and hospitalization events. Second, existing studies incorporate patient physiological data and household appliance use in addition to household activity. We currently use only household movement data. Third, existing studies rely on either fixed training data or periodically refreshed training data, whereas our approach was designed to work in a streaming environment, implying that our daily detection and alerting algorithm uses the information in the patient timeline up to the current day. Fourth, our algorithm is patient data driven rather than cohort data driven. This means that we evaluate the average recall by considering the algorithm's performance on individual patient data. In contrast, published work takes a cohort-wide or patient-blind approach to assess algorithm performance. These factors should be collectively considered when comparing our work with other Minder-based research.

The CMP-based approach is ideally suited to anomaly detection applications where data and labels are characterized by real-world noise and annotated training data required for supervised learning may not be available because of resource constraints or in a streaming data scenario, as well as where the distinction between normal and anomalous data is not clear-cut. This includes sensor-based remote health monitoring in a variety of industrial, urban, and health care settings. The CMP-based approach excels at zooming out and focusing on temporal patterns at configurable time scales. It is also designed with personalization in mind, which makes it especially relevant for health care, where patterns of similar anomalies or the same disease can present differently in different individuals. It is ideal for situations in which explainability is key for operational monitoring teams.

The CMP-based approach may not be ideal for applications that prioritize sensitivity over interpretability. It is also not the best tool for data that have a well-defined, well-understood pattern, such as electrocardiogram data, or where noise levels are low or the distinction between normal and abnormal data is clearly

XSL•FO
RenderX

understood. Finally, the CMP pipeline would need to be augmented with feature reduction methods for it to scale to high-dimensional data.

Limitations

A limitation of the anomaly detection method presented in this paper is that cross-sensor correlations were not considered. This will be investigated in future studies using interpretable machine learning. Second, our study ignored sensor data from the front and back doors. This omission was intentional as we were interested in detecting anomalies arising from significant changes in indoor household activity instead of those arising from out-of-home situations. In addition, front door and back door opening and closing are as of yet difficult to attribute to the person with dementia. Third, a system to distinguish the patient from other household members is needed to improve the robustness of anomaly detection models based on passive sensing. Fourth, to achieve finer granularity and lower latency than 1 "context," the CMP-based anomaly detection model should be configured to ingest data hourly or at a higher resolution than 1 day. Fifth, we assume that an anomaly in a single time window can be deterministic of a complex health event. However, the presence of pets and visitors could also contribute to anomalies. To address this, we require not only a distinction between patient and carer but also a way of monitoring anomalies in subsequent time windows to correlate anomalies with health changes with high confidence. Finally, we will require a larger sample size to further validate our approach.

Conclusions and Future Work

In this study, we developed a novel lightweight unsupervised anomaly detection pipeline based on the CMP and evaluated it in sensor-based remote health monitoring of patients with dementia. We combined univariate CMP scores in novel ways, developed the multivariate CMP, and tested it for identifying anomalous patient days via thresholding in sliding windows. We demonstrated CMP-based anomaly scoring to be more effective and generalizable than other comparable methods for unsupervised anomaly detection. Specifically, the multidimensional CMP based on a 7-day sliding window and using the top 2 contributing patient-specific features exhibits 84.3% recall with only 32 alerts over the average patient timeline of 624 days. In addition, we showed how the CMP can be used to uncover and explain digital biomarkers of anomalies at the time segment, patient, and cohort levels. Our study of 9363 days collected from 15 people living with dementia who had UTI and hospitalization events in their timeline showed that unusual bathroom activity in the early and late hours of the day is a prominent biomarker of anomalies across our cohort. This helps quantitatively validate observational studies of similar behavior in patients with dementia.

Our future work will focus on the following areas: adding physiological data to the anomaly-scoring pipeline, developing the CMP as a tool for effective visual monitoring of patterns and anomalies in data and accommodating other distance metrics in addition to Euclidean distance, validating the CMP on a larger patient cohort and different kinds of anomalies, and using machine learning methods to use CMP-based scores to classify different types of anomalies. We will also investigate seasonal effects and compare our method with other relevant anomaly detection methods. We plan to integrate our model into the Minder platform to raise alerts when anomalies are detected to enable the monitoring team to investigate the underlying sensor data and offer timely intervention to patients. Alerts that are validated as true will be recorded in the patient timeline and used to monitor the operational accuracy of our model.

Acknowledgments

This research was supported by the UK Dementia Research Institute (see Multimedia Appendix 1 for a full acknowledgment list). The authors also wish to thank Dieter De Paepe for authoring and assisting with the use of the Python Series Distance Matrix library and Professor Philip Evans (Professor of Medical Radiation Imaging at the Centre of Vision, Speech, and Signal Processing at the University of Surrey) for helpful discussions.

Data Availability

Access to the complete data set requires ethics approval and is not currently available because of the sensitivity of the in-home monitoring data and privacy and security requirements. To inquire about the complete data set, a request with a statement of intent and use must be sent to Professor Payam Barnaghi, Head of the Machine Intelligence group at UK Dementia Research Institute Care Research and Technology Centre at Imperial College.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Acknowledgment list for the UK Dementia Research Institute.
[PDF File (Adobe PDF File), 51 KB - aging_v5i3e38211_app1.pdf]

References

XSL•FO
RenderX

1. Dementia. World Health Organization. 2021 Sep 2. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2021-12-01]
2. Facts for the media about dementia. Alzheimer's Society. URL: https://www.alzheimers.org.uk/about-us/news-and-media/facts-media [accessed 2021-12-01]
3. UK Dementia Research Institute Care Research & Technology Centre, Imperial College London. URL: https://www.imperial.ac.uk/uk-dri-care-research-technology [accessed 2022-09-14]
4. UK Dementia Research Institute Care Research & Technology. URL: https://ukdri.ac.uk/centres/care-research-technology [accessed 2021-12-01]
5. Arifoglu D, Bouchachia A. Detection of abnormal behaviour for dementia sufferers using Convolutional Neural Networks. Artif Intell Med 2019 Mar;94:88-95. [doi: 10.1016/j.artmed.2019.01.005] [Medline: 30871686]
6. Rowe TA, Juthani-Mehta M. Urinary tract infection in older adults. Aging health 2013 Oct;9(5):519-528 [FREE Full text] [doi: 10.2217/ahe.13.38] [Medline: 24391677]
7. Gavazzi G, Delerce E, Cambau E, François P, Corroyer B, de Wazières B, et al. Diagnostic criteria for urinary tract infection in hospitalized elderly patients over 75 years of age: a multicenter cross-sectional study. Med Mal Infect 2013 May;43(5):189-194. [doi: 10.1016/j.medmal.2013.02.006] [Medline: 23622951]
8. Marques LP, Flores JT, Barros Junior OD, Rodrigues GB, Mourão C, Moreira RM. Epidemiological and clinical aspects of urinary tract infection in community-dwelling elderly women. Braz J Infect Dis 2012 Sep;16(5):436-441 [FREE Full text] [doi: 10.1016/j.bjid.2012.06.025] [Medline: 22975174]
9. Nicolle LE. Asymptomatic bacteriuria. Infect Disease Clin North Am 2003 Jun;17(2):367-394. [doi: 10.1016/s0891-5520(03)00008-4]
10. Ahmad S, Lavin A, Purdy S, Agha Z. Unsupervised real-time anomaly detection for streaming data. Neurocomputing 2017 Nov;262:134-147. [doi: 10.1016/j.neucom.2017.04.070]
11. Pimentel MA, Clifton DA, Clifton L, Tarassenko L. A review of novelty detection. Signal Processing 2014 Jun;99:215-249. [doi: 10.1016/j.sigpro.2013.12.026]
12. Rantz M, Skubic M, Koopman R, Phillips L, Alexander G, Miller S, et al. Using sensor networks to detect urinary tract infections in older adults. In: Proceedings of the 2011 IEEE 13th International Conference on e-Health Networking, Applications and Services. 2011 Presented at: 2011 IEEE 13th International Conference on e-Health Networking, Applications and Services; Jun 13-15, 2011; Columbia, MO, USA. [doi: 10.1109/health.2011.6026731]
13. Mori T, Fujii A, Shimosaka M, Noguchi H, Sato T. Typical behavior patterns extraction and anomaly detection algorithm based on accumulated home sensor data. In: Proceedings of the Future Generation Communication and Networking (FGCN 2007). 2007 Presented at: Future Generation Communication and Networking (FGCN 2007); Dec 06-08, 2007; Jeju, Korea (South). [doi: 10.1109/FGCN.2007.226]
14. Hoque E, Dickerson R, Preum S, Hanson M, Barth A, Stankovic J. Holmes: a comprehensive anomaly detection system for daily in-home activities. In: Proceedings of the 2015 International Conference on Distributed Computing in Sensor Systems. 2015 Presented at: 2015 International Conference on Distributed Computing in Sensor Systems; Jun 10-12, 2015; Fortaleza, Brazil. [doi: 10.1109/dcoss.2015.20]
15. Enshaeifar S, Zoha A, Markides A, Skillman S, Acton ST, Elsaleh T, et al. Health management and pattern analysis of daily living activities of people with dementia using in-home sensors and machine learning techniques. PLoS One 2018;13(5):e0195605 [FREE Full text] [doi: 10.1371/journal.pone.0195605] [Medline: 29723236]
16. Novák M, Bi?as M, Jakab F. Unobtrusive anomaly detection in presence of elderly in a smart-home environment. In: Proceedings of the 2012 ELEKTRO. 2012 Presented at: 2012 ELEKTRO; May 21-22, 2012; Rajecke Teplice, Slovakia. [doi: 10.1109/elektro.2012.6225617]
17. Li H, Zhang Q, Duan P. A novel one-pass neural network approach for activities recognition in intelligent environments. In: Proceedings of the 2008 7th World Congress on Intelligent Control and Automation. 2008 Presented at: 2008 7th World Congress on Intelligent Control and Automation; Jun 25-27, 2008; Chongqing, China. [doi: 10.1109/wcica.2008.4592901]
18. Singla G, Cook DJ, Schmitter-Edgecombe M. Recognizing independent and joint activities among multiple residents in smart environments. J Ambient Intell Humaniz Comput 2010 Mar 01;1(1):57-63 [FREE Full text] [doi: 10.1007/s12652-009-0007-1] [Medline: 20975986]
19. Akl A, Taati B, Mihailidis A. Autonomous unobtrusive detection of mild cognitive impairment in older adults. IEEE Trans Biomed Eng 2015 May;62(5):1383-1394. [doi: 10.1109/tbme.2015.2389149]
20. Jakkula V, Cook D, Crandall A. Temporal pattern discovery for anomaly detection in a smart home. In: Proceedings of the 2007 3rd IET International Conference on Intelligent Environments. 2007 Presented at: 2007 3rd IET International Conference on Intelligent Environments; Sep 24-25, 2007; Ulm. [doi: 10.1049/CP:20070390]
21. Palermo F, Li H, Capstick A, Fletcher-Lloyd N, Zhao Y, Kouchaki S, et al. Designing a clinically applicable deep recurrent model to identify neuropsychiatric symptoms in people living with dementia using in-home monitoring data. arXiv. Preprint posted online October 19, 2021 [FREE Full text]
22. Li H, Enshaeifar S, Skillman S, Markides A, Kenny M, Sharp D, et al. A semi-supervised machine learning model for risk analysis of urinary tract infections in people with dementia. SSRN J 2020. [doi: 10.2139/ssrn.3682493]

XSL•FO
RenderX

23. Paudel R, Eberle W, Holder L. Anomaly Detection of Elderly Patient Activities in Smart Homes using a Graph-Based Approach. London: CSREA Press; 2018.
24. Eberle W, Holder L. Anomaly detection in data represented as graphs. Intell Data Analysis 2007 Nov 09;11(6):663-689. [doi: 10.3233/ida-2007-11606]
25. Gupta P, McClatchey R, Caleb-Solly P. Tracking changes in user activity from unlabelled smart home sensor data using unsupervised learning methods. Neural Comput Applic 2020 Jan 25;32(16):12351-12362. [doi: 10.1007/s00521-020-04737-6]
26. Li H, Rezvani R, Kolanko M, Sharp D, Wairagkar M, Vaidyanathan R, et al. An attention model to analyse the risk of agitation and urinary tract infections in people with dementia. arXiv. Preprint posted online January 18, 2021 [FREE Full text] [doi: 10.48550/arXiv.2101.07007]
27. Lin J, Keogh E, Fu A, Van HH. Approximations to magic: finding unusual medical time series. In: Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05). 2005 Presented at: 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05); Jun 23-24, 2005; Dublin, Ireland. [doi: 10.1109/cbms.2005.34]
28. Lan S, Chen R, Zhao W. Anomaly detection on IT operation series via online matrix profile. arXiv. Preprint posted online August 27, 2021 [FREE Full text] [doi: 10.48550/arXiv.2108.12093]
29. Steenwinckel B, De Paepe D, Vanden Hautte S, Heyvaert P, Bentefrit M, Moens P, et al. FLAGS: a methodology for adaptive anomaly detection and root cause analysis on sensor data streams by fusing expert knowledge with machine learning. Future Generation Comput Syst 2021 Mar;116:30-48. [doi: 10.1016/j.future.2020.10.015]
30. Nieves Avendano D, Caljouw D, Deschrijver D, Van Hoecke S. Anomaly detection and event mining in cold forming manufacturing processes. Int J Adv Manuf Technol 2020 Oct 10;115(3):837-852. [doi: 10.1007/s00170-020-06156-2]
31. De Paepe D, Vanden Hautte S, Steenwinckel B, De Turck F, Ongenae F, Janssens O, et al. A generalized matrix profile framework with support for contextual series analysis. Eng Appl Artif Intell 2020 Apr;90:103487. [doi: 10.1016/j.engappai.2020.103487]
32. Chandola V, Banerjee A, Kumar V. Anomaly detection. ACM Comput Surv 2009 Jul;41(3):1-58 [FREE Full text] [doi: 10.1145/1541880.1541882]
33. Yeh C, Zhu Y, Ulanova L, Begum N, Ding Y, Dau H, et al. Matrix profile I: all pairs similarity joins for time series: a unifying view that includes motifs, discords and shapelets. In: Proceedings of the 2016 IEEE 16th International Conference on Data Mining (ICDM). 2016 Presented at: 2016 IEEE 16th International Conference on Data Mining (ICDM); Dec 12-15, 2016; Barcelona, Spain. [doi: 10.1109/icdm.2016.0179]
34. Yeh C, Kavantzas N, Keogh E. Matrix profile VI: Meaningful multidimensional motif discovery. In: Proceedings of the 2017 IEEE International Conference on Data Mining (ICDM). 2017 Presented at: 2017 IEEE International Conference on Data Mining (ICDM); Nov 18-21, 2017; New Orleans, LA, USA. [doi: 10.1109/ICDM.2017.66]
35. Multidimensional motif discovery. Stumpy. URL: https://stumpy.readthedocs.io/en/latest/Tutorial_Multidimensional_Motif_Discovery.html#Multi-dimensional-Matrix-Profiles [accessed 2021-12-01]
36. Wang X, Mueen A, Ding H, Trajcevski G, Scheuermann P, Keogh E. Experimental comparison of representation methods and distance measures for time series data. Data Min Knowl Disc 2012 Feb 10;26(2):275-309. [doi: 10.1007/s10618-012-0250-5]
37. González C, Jensen E, Gambús P, Vallverdú M. Entropy measures as descriptors to identify apneas in rheoencephalographic signals. Entropy (Basel) 2019 Jun 18;21(6):605 [FREE Full text] [doi: 10.3390/e21060605] [Medline: 33267319]
38. Pincus SM, Gladstone IM, Ehrenkranz RA. A regularity statistic for medical data analysis. J Clin Monitor Comput 1991 Oct;7(4):335-345. [doi: 10.1007/bf01619355]
39. Montesinos L, Castaldo R, Pecchia L. Selection of entropy-measure parameters for force plate-based human balance evaluation. In: World Congress on Medical Physics and Biomedical Engineering 2018. Singapore: Springer; 2018.
40. Satopaa V, Albrecht J, Irwin D, Raghavan B. Finding a "Kneedle" in a haystack: detecting knee points in system behavior. In: Proceedings of the 2011 31st International Conference on Distributed Computing Systems Workshops. 2011 Presented at: 2011 31st International Conference on Distributed Computing Systems Workshops; Jun 20-24, 2011; Minneapolis, MN, USA. [doi: 10.1109/icdcsw.2011.20]
41. Wilde H, Knight V, Gillard J, Smith K. Segmentation analysis and the recovery of queuing parameters via the Wasserstein distance: a study of administrative data for patients with chronic obstructive pulmonary disease. arXiv. Preprint posted online August 10, 2020 [FREE Full text] [doi: 10.48550/arXiv.2008.04295]
42. Yan Y, Duffner S, Phutane P, Berthelier A, Blanc C, Garcia C, et al. 2D Wasserstein loss for robust facial landmark detection. Pattern Recognit 2021 Aug;116:107945. [doi: 10.1016/j.patcog.2021.107945]
43. Skubic M, Guevara RD, Rantz M. Automated health alerts using in-home sensor data for embedded health assessment. IEEE J Transl Eng Health Med 2015;3:1-11. [doi: 10.1109/jtehm.2015.2421499]
44. Pevný T. Loda: lightweight on-line detector of anomalies. Mach Learn 2015 Jul 21;102(2):275-304. [doi: 10.1007/s10994-015-5521-0]
45. Rantz MJ, Skubic M, Koopman RJ, Alexander GL, Phillips L, Musterman K, et al. Automated technology to speed recognition of signs of illness in older adults. J Gerontol Nurs 2012 Apr;38(4):18-23 [FREE Full text] [doi: 10.3928/00989134-20120307-01] [Medline: 22420519]

XSL•FO
RenderX

46. D'Agata E, Loeb MB, Mitchell SL. Challenges in assessing nursing home residents with advanced dementia for suspected urinary tract infections. J Am Geriatr Soc 2013 Jan 11;61(1):62-66 [FREE Full text] [doi: 10.1111/jgs.12070] [Medline: 23311553]
47. Ijaopo EO. Dementia-related agitation: a review of non-pharmacological interventions and analysis of risks and benefits of pharmacotherapy. Transl Psychiatry 2017 Oct 31;7(10):e1250 [FREE Full text] [doi: 10.1038/tp.2017.199] [Medline: 29087372]
48. Avogadro P, Palonca L, Dominoni MA. Online anomaly search in time series: significant online discords. Knowl Inf Syst 2020 Mar 09;62(8):3083-3106. [doi: 10.1007/s10115-020-01453-4]
49. Kriegel H, Schubert M, Zimek A. Angle-based outlier detection in high-dimensional data. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. 2008 Presented at: KDD08: The 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 24 - 27, 2008; Las Vegas Nevada USA. [doi: 10.1145/1401890.1401946]
50. Li Z, Zhao Y, Botta N, Ionescu C, Hu X. COPOD: copula-based outlier detection. In: Proceedings of the 2020 IEEE International Conference on Data Mining (ICDM). 2020 Presented at: 2020 IEEE International Conference on Data Mining (ICDM); Nov 17-20, 2020; Sorrento, Italy. [doi: 10.1109/icdm50108.2020.00135]
51. Zach. Statology. 2021 Apr 5. URL: https://www.statology.org/modified-z-score/ [accessed 2021-12-01]
52. Wang T, Azad T, Ranan R. The emerging influence of digital biomarkers on healthcare. RockHealth.org. 2016 Jan 31. URL: https://rockhealth.com/reports/the-emerging-influence-of-digital-biomarkers-on-healthcare/ [accessed 2021-12-01]
53. Cagney D, Sul J, Huang RY, Ligon KL, Wen PY, Alexander BM. The FDA NIH Biomarkers, EndpointS, and other Tools (BEST) resource in neuro-oncology. Neuro Oncol 2018 Aug 02;20(9):1162-1172 [FREE Full text] [doi: 10.1093/neuonc/nox242] [Medline: 29294069]
54. Foxman B. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am J Med 2002 Jul;113(1):5-13. [doi: 10.1016/s0002-9343(02)01054-9]
55. Davis CP, Shiel Jr WC. Is a urinary tract infection (UTI) contagious? RxList. 2022 Jan 8. URL: https://www.rxlist.com/is_a_urinary_tract_infection_uti_contagious/article.htm [accessed 2021-12-01]
56. Dau H, Keogh E. Matrix profile V: a generic technique to incorporate domain knowledge into motif discovery. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2017 Presented at: KDD '17: The 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; Aug 13 - 17, 2017; Halifax NS Canada. [doi: 10.1145/3097983.3097993]
57. Domingues R, Filippone M, Michiardi P, Zouaoui J. A comparative evaluation of outlier detection algorithms: experiments and analyses. Pattern Recognit 2018 Feb;74:406-421. [doi: 10.1016/j.patcog.2017.09.037]

Abbreviations

ABOD: Angle-Based Outlier Detection
CMP: Contextual Matrix Profile
COPOD: Copula-Based Outlier Detection
LODA: Lightweight Online Detector of Anomalies
MP: Matrix Profile
PIR: passive infrared
UTI: urinary tract infection

Edited by J Wang, T Leung; submitted 23.03.22; peer-reviewed by G Abebe, R Zhao; comments to author 21.06.22; revised version received 04.07.22; accepted 30.07.22; published 19.09.22.

Please cite as:
Bijlani N, Nilforooshan R, Kouchaki S
An Unsupervised Data-Driven Anomaly Detection Approach for Adverse Health Conditions in People Living With Dementia: Cohort Study
JMIR Aging 2022;5(3):e38211
URL: https://aging.jmir.org/2022/3/e38211
doi:10.2196/38211
PMID:36121687

©Nivedita Bijlani, Ramin Nilforooshan, Samaneh Kouchaki. Originally published in JMIR Aging (https://aging.jmir.org), 19.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding the Experience of Geriatric Care Professionals in Using Telemedicine to Care for Older Patients in Response to the COVID-19 Pandemic: Mixed Methods Study

Wenwen Chen[1], BHA, BSc; Ashley Flanagan[2], PhD; Pria MD Nippak[1], PhD; Michael Nicin[2], MPP; Samir K Sinha[2,3], MD, DPhil

[1]School of Health Services Management, Ted Rogers School of Management, Toronto Metropolitan University, Toronto, ON, Canada

[2]National Institute on Ageing, Ted Rogers School of Management, Toronto Metropolitan University, Toronto, ON, Canada

[3]Division of Geriatric Medicine, Department of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Samir K Sinha, MD, DPhil
National Institute on Ageing
Ted Rogers School of Management
Toronto Metropolitan University
350 Victoria Street
Toronto, ON, M5B 2K3
Canada
Phone: 1 416 586 4800 ext 7859
Email: Samir.Sinha@sinaihealth.ca

Abstract

Background: Geriatric care professionals were forced to rapidly adopt the use of telemedicine technologies to ensure the continuity of care for their older patients in response to the COVID-19 pandemic. However, there is little current literature that describes how telemedicine technologies can best be used to meet the needs of geriatric care professionals in providing care to frail older patients, their caregivers, and their families.

Objective: This study aims to identify the benefits and challenges geriatric care professionals face when using telemedicine technologies with frail older patients, their caregivers, and their families and how to maximize the benefits of this method of providing care.

Methods: This was a mixed methods study that recruited geriatric care professionals to complete an online survey regarding their personal demographics and experiences with using telemedicine technologies and participate in a semistructured interview. Interview responses were analyzed using the Consolidated Framework for Implementation Research (CFIR).

Results: Quantitative and qualitative data were obtained from 30 practicing geriatric care professionals (22, 73%, geriatricians, 5, 17%, geriatric psychiatrists, and 3, 10%, geriatric nurse practitioners) recruited from across the Greater Toronto Area. Analysis of interview data identified 5 CFIR contextual barriers (complexity, design quality and packaging, patient needs and resources, readiness for implementation, and culture) and 13 CFIR contextual facilitators (relative advantage, adaptability, tension for change, available resources, access to knowledge, networks and communications, compatibility, knowledge and beliefs, self-efficacy, champions, external agents, executing, and reflecting and evaluating). The CFIR concept of external policy and incentives was found to be a neutral construct.

Conclusions: This is the first known study to use the CFIR to develop a comprehensive narrative to characterize the experiences of Ontario geriatric care professionals using telemedicine technologies in providing care. Overall, telemedicine can significantly enable most of the geriatric care that is traditionally provided in person but is less useful in providing specific aspects of geriatric care to frail older patients, their caregivers, and their families.

(JMIR Aging 2022;5(3):e34952) doi:10.2196/34952

KEYWORDS

telemedicine; virtual care visit; geriatric care professionals; aging population; Consolidated Framework for Implementation Research; geriatric care; older adults; elderly care; telehealth; digital health; COVID-19; pandemic; technology usability

XSL•FO
RenderX

Introduction

Canada's older population remains at the greatest risk of dying from COVID-19, caused by the novel SARS-COV-2 virus [1]. Yet, the same public health measures being imposed to protect this population have also posed an ongoing challenge for older persons in accessing in-person care in a timely manner since the beginning of the COVID-19 pandemic. In response, the Government of Ontario's health care system, like several others, rapidly supported the early and widespread adoption and use of its existing and other telemedicine technologies, including the telephone or popular videoconferencing platforms, such as Zoom and Skype, for health care professionals to deliver safe and effective remote or virtual care throughout the COVID-19 pandemic [2].

For older persons, previously noted beneficial outcomes of Ontario's telemedicine services have included a decrease in wait times for access to specialists [3] and a significant reduction in emergency department (ED) admissions [4]. However, until the beginning of the COVID-19 pandemic, there remained major challenges that hindered the widespread adoption of telemedicine technologies by health care professionals across Ontario. For example, a main barrier was that prior to the COVID-19 pandemic, Ontario physicians could only be reimbursed for providing telemedicine services if they and their patients were able to use the Government of Ontario's Ontario Telemedicine Network (OTN) secure videoconferencing technologies to conduct patient visits [5]. Telephone-based consultations or follow-up were not reimbursable in general for Ontario physicians, except for geriatricians when providing caregiver advice and support to one of their established patients. Furthermore, another main barrier was that acquiring the required communication technologies to enable secure videoconferencing via telemedicine could be expensive for both patients, their caregivers and families, health care professionals, and their organizations, although this was becoming less of an issue in recent years with the greater availability of secure web-based communication technologies using standard computer equipment. Indeed, many older persons, their caregivers, and their families might also not be able to access the technology needed to use telemedicine services [6]. In addition, many older persons with cognitive and sensory impairments need to rely on health care professionals or unpaid caregivers and family members to assist with or manage the technology [6]. This is a particular concern for those who are homebound or live in isolation, as they may not readily have access to the level of health care professionals or caregiver support necessary for accessing telemedicine technology-enabled supports [7]. Nevertheless, evaluations of Ontario telemedicine programs, prior to the COVID-19 pandemic, have demonstrated high satisfaction among older patient users [3,6,8-11].

Despite positive reported patient outcomes, there remain gaps in the current literature on whether the use of telemedicine technologies can adequately meet the needs of Ontario geriatric care professionals to facilitate the delivery of the range of care they provide. Many observed findings from previous evaluations of Ontario telemedicine programs have reflected the needs of patients and program stakeholders specific to individual conditions, such as telehomecare for older patients with chronic obstructive pulmonary disease (COPD) and heart failure [6], or more specific geographical needs, such as the provision of geriatric medicine and psychiatry outreach clinics to rural and remote communities [9]. Meanwhile, the complexities that geriatric care professionals in Ontario typically manage using a holistic in-person approach that assesses the complex and often interrelated health and social issues experienced by their older patients have not been addressed through prior studies examining the use of telemedicine technologies.

Another issue is that many of the existing Ontario telemedicine programs have tended to target older persons who were more able-bodied, and have often excluded the more vulnerable, older persons with complex conditions [6,12,13]. In particular, health care professionals did not see the benefit of using telemedicine technologies with older persons with physical and cognitive impairments, as they were concerned that this older subgroup of patients could not keep up with the unique demands a remote consultation requires [6,14,15]. However, the exclusion of this subgroup in prior studies has only served to pose a greater challenge for geriatric care professionals in assessing their ability to transition to using telemedicine, particularly with their older patients with complex conditions [16].

At the outset of the COVID-19 pandemic, Ontario geriatric care professionals could not be as selective with their older patients regarding how best to care for them under existing public health measures, including when lockdowns and other restricted visiting orders were enacted and there was a general fear of the possible consequences that could befall older patients with complex conditions seeking nonemergent health care services. Ironically, all these factors could further promote physical and social isolation that increase the risk of worsening functional decline and mental health issues that could actually result in more ED visits and acute hospital admissions [15,17].

Existing telemedicine research also has not addressed how best to facilitate and support the level of care that Ontario geriatric care professionals aim to provide. The current literature indicates that virtual care visits require various ongoing background coordination supports, such as patient data management, patient care monitoring, and facilitation of communications between health care staff involved in care planning [6,15,18]. In addition, telemedicine technologies have often been associated with a limited ability to perform a physical exam [19,20] and a difficulty in observing verbal and nonverbal cues that could impact establishing reliable diagnoses [19,21]. Therefore, the consulting health care professional has often needed to rely on a health care facilitator, whenever possible, who would be on the premises with the patient (eg, a local physician or nurse), as their support, which has been seen as crucial for an effective remote visit [15]. However, in the current COVID-19 pandemic paradigm, consulting health care professionals have also had to serve as facilitators, managing all aspects of the telemedicine visit unless there was a family member, caregiver, or health care professional present to assist with a visit.

Therefore, this study aims to determine the benefits and challenges Ontario geriatric care professionals have faced in using telemedicine technologies to provide routine consultations

XSL•FO
RenderX

and follow-up care to their older patients with complex needs, their families, and their caregivers, as well as the conditions under which this method proves to be or not be an effective way to provide care. Moreover, identifying frontline benefits and challenges would also provide new learning opportunities for geriatric care professionals across different health care settings and regions in the use of telemedicine technologies [22], which is still a novel approach in the practice of geriatric medicine [13].

Methods

Study Design

This was a mixed methods study that included the following: (1) a survey that inquired about the demographic information of participating geriatric care professionals, their experience, and current satisfaction with the use of telemedicine technologies to provide care to their older patients and their caregivers and (2) a semistructured interview that reflected the objective of the study but also allowed participants to freely express their additional perspectives about the use of telemedicine technologies to provide care. Please see the Multimedia Appendix 1 for the semistructured interview guide.

The Consolidated Framework for Implementation Research (CFIR) was used to assess the barriers and facilitators toward providing care with telemedicine technologies identified in the semistructured interviews. The CFIR was identified as an appropriate methodological framework for providing a comprehensive evaluation of the barriers and facilitators in the implementation of health care technologies across multiple contexts [16,23,24]. The CFIR consists of 5 domains: intervention characteristics, outer setting, inner setting, individual characteristics, and the process of implementation [24]. Within each domain, there are various constructs that guide users to identify barriers and facilitators that impact implementation [24].

Participants

This study targeted geriatric care professionals who use telemedicine technologies with older patients, their caregivers, and their family members in an outpatient setting to provide routine consultations and follow-up care. The geriatric care professionals were recruited through the Divisions of Geriatric Medicine and Geriatric Psychiatry at the University of Toronto, Canada, and the local Regional Geriatric Program of Toronto between January and April 2021.

A total of 30 geriatric care professionals representing the fields of geriatric medicine, geriatric psychiatry, and geriatric nursing were invited to the survey and participated in a semistructured interview. These geriatric care professionals work in geriatric outpatient clinics that often do not use allied health care workers.

Ethical Considerations

Participants provided either written informed consent or audio-recorded oral consent. The study protocol was approved by the Toronto Metropolitan University Research Ethics Board (#2020-513-1).

Data Analysis

The Statistical Package for the Social Sciences (SPSS; IBM Corporation) was used to analyze the survey responses and determine participant characteristics. A deductive thematic approach to analysis was used to analyze the semistructured interviews. Authors WC and AF independently coded the transcripts using the codebook based on the CFIR constructs [24], which was modified to reflect the local geriatric care professionals' practice. NVivo (March 2020 release; QSR International) was used to facilitate the coding process. Following a codebook helped to minimize coding differences, and weekly discussions were held to resolve coding differences.

Results

Participant Characteristics

Of the 30 participants, 22 (73%) were geriatric medicine specialists or geriatricians, 5 (17%) were geriatric psychiatrists, and 3 (10%) were geriatric nurse practitioners. In addition, 28 (93%) participants completed both the survey and the semistructured interview. The survey results of 2 (7%) participants were not collected due to personal choice or technical difficulties. Tables 1 and 2 provide a detailed overview of the participant characteristics and satisfaction with telemedicine use, respectively, from the 28 completed surveys.

XSL•FO
RenderX

Table 1. Participant characteristics (N=28).

Participant characteristics	Participants[a]
Age (years), mean (range)	44 (30-74)
Gender[b], n (%)	
Woman	15 (54)
Man	12 (43)
Clinical profession, n (%)	
Doctor of medicine (MD)	26 (93)
Nurse practitioner (NP)[c]	2 (7)
Medical discipline, n (%)	
Geriatrics	23 (82)
Geriatric psychiatry	5 (18)
Years of professional experience, n (%)	
Less than 3 years	6 (21)
4-10 years	10 (36)
More than 10 years	12 (43)
Patient setting, n (%)	
Outpatient	25 (89)
Other[d]	3 (11)
Used telemedicine prior to COVID-19, n (%)	
Yes	13 (46)
No	15 (54)
Experience with telemedicine, n (%)	
3-6 months	2 (7)
6 months-1 year	15 (54)
More than 1 year	11 (39)
Frequency of telemedicine use, n (%)	
Rarely	1 (4)
Sometimes	5 (18)
Often	20 (71)
Always	2 (7)
Types of telemedicine platforms, n (%)	
OTN[e] videoconferencing	5 (18)
Zoom/Skype/Google Hangouts/Facetime	4 (14)
Combination of telemedicine platforms[f]	19 (68)

[a]The survey results of 2 participants were not collected, 1 participant declined to complete the survey, and 1 participant's survey was not collected due to technical difficulties.

[b]One participant did not provide gender information.

[c]Both nurse practitioners practiced in a geriatric medicine setting.

[d]The "Other" setting included a combination of an outpatient setting, long-term care homes, and supportive housing.

[e]OTN: Ontario Telemedicine Network.

[f]The combination of telemedicine platforms included Zoom, the OTN, email, telephone WebEx, Facebook Messenger, Microsoft Teams, electronic medical record (EMR)-based applications, WhatsApp, and Facetime.

XSL•FO
RenderX

Table 2. Geriatric care professional telemedicine satisfaction survey (N=28).

Questions	Strongly disagree, n (%)	Disagree, n (%)	Neutral, n (%)	Agree, n (%)	Strongly agree, n (%)
1. Telemedicine can increase my productivity in delivering patient care.	0	4 (14)	5 (18)	10 (36)	9 (32)
2. My patients provide me with sufficient information about their comorbidities using telemedicine.	1 (4)	3 (11)	4 (14)	16 (57)	4 (14)
3. I can conduct a comprehensive geriatric assessment using telemedicine[a].	0	8 (29)	1 (4)	16 (57)	3 (11)
4. Telemedicine services do not require a lot of training to use[a].	1 (4)	10 (36)	4 (14)	12 (43)	1 (4)
5. Telemedicine services are compatible with the existing clinical workflow.	1 (4)	4 (14)	2 (7)	17 (61)	4 (14)
6. Teleconsultation is as effective as an in-person consultation[a].	3 (11)	12 (43)	5 (18)	8 (29)	0
7. My older patients can easily communicate with me using telemedicine.	2 (7)	10 (36)	10 (36)	5 (21)	0
8. I can engage with my patients, their families, and their caregivers about treatment plans using telemedicine.	0	1 (4)	5 (18)	18 (64)	4 (14)
9. I would continue to use telemedicine to care for my older patients beyond the pandemic.	0	2 (7)	2 (7)	14 (50)	9 (36)
10. Overall, I am satisfied with using telemedicine with older patients.	0	2 (7)	6 (21)	17 (61)	3 (11)

[a]The percentages do not add up to 100% due to rounding.

Barriers to and Facilitators of Telemedicine Use

Table 3 details the barriers and facilitators associated with the implementation of telemedicine in geriatric care practice. Only relevant key constructs identified within the 5 CFIR domains are discussed herein.

XSL•FO
RenderX

Table 3. Adapted CFIR[a] operational codes.

Domains and constructs	Operational definition[b]	Facilitator/barrier
I. Intervention characteristics		
Relative advantage	Perception of geriatric care professionals seeing virtual care visits as an advantage versus in-person consultations	Facilitator
Adaptability	The degree to which the virtual care visit was tailored to meet the needs of geriatric care professionals	Facilitator
Complexity	Perceived complexity of how virtual care assessments compared to in-person assessments	Barrier
Design quality and packaging	Perceived quality of telemedicine platforms and how the innovation is bundled and presented	Barrier
II. Outer setting		
External policy and incentives	Broad constructs on government policies, such as confidentiality issues/consent with older patients, as well as discussions about how to bill for virtual care visits (consults via telephone, text messages, or videoconferencing)	Neutral
Patient needs and resources	The degree to which the needs of older patients with complex needs, their caregivers, and their families are accurately known and prioritized during virtual care visits	Barrier
III. Inner setting		
Networks and communications	The quality of information derived from fellow colleagues, caregivers, families, and local EMR[c] systems to develop collateral history regarding older patients with complex needs	Facilitator
Culture	Norms, values, and basic assumptions of geriatric care professionals toward telemedicine use prior to the COVID-19 pandemic	Barrier
Implementation climate: tension for change	The degree of willingness to transition to telemedicine use	Facilitator
Implementation climate: compatibility	The degree of tangible fit between meaning and values attached to virtual care visits, how those align with the geriatric care professionals' own norms, values, and perceived risks and needs, and how virtual care visits fit into the existing workflow and systems	Facilitator
Readiness for implementation	Geriatric care professionals' readiness to implement virtual care visits	Barrier
Readiness for implementation: access to knowledge and information	Ease of access to training and support provided on how to conduct virtual visits	Facilitator
Readiness for implementation: available resources	The level of resources provided for telemedicine use, including technological infrastructure, dedicated clinic space to conduct virtual care visits, and educational guidance	Facilitator
IV. Individual characteristics		
Knowledge and beliefs about the intervention	Geriatric care professionals' attitudes toward the values placed on virtual care, as well as familiarity with facts, truths, and principles related to telemedicine technologies	Facilitator
Self-efficacy	Geriatric care professionals' beliefs in their own capabilities in using telemedicine technologies with older patients, their caregivers, and their families	Facilitator
V. Implementation process		
Engaging: champions	Individuals who drove the implementation of virtual care visits forward	Facilitator
Engaging: external change agents	Outside individuals who formally influenced or facilitated virtual care visit decisions in a desirable direction	Facilitator
Executing	Carrying out and accomplishing tasks during care visits	Facilitator
Reflecting and evaluating	Quantitative and qualitative feedback on progress and quality to enhance virtual care visits	Facilitator

[a]CFIR: Consolidated Framework for Implementation Research.

[b]The operational definitions of the constructs are adapted to reflect the geriatric care professionals' experiences.

[c]EMR: electronic medical record.

XSL•FO
RenderX

Domain I: Intervention Characteristics

Relative Advantage

Using telemedicine technologies increased access to older patients who were homebound, were reluctant to come in due to COVID-19, or lived in remote areas. Many geriatric care professionals perceived that virtual care visits allowed older patients the convenience to receive care in their own homes without the hassle of traveling. Older patients who were previously "no shows" for their appointments were also more likely to be reached. Virtual care visits also provided greater schedule flexibility for geriatric care professionals to accommodate the schedules of their older patients, their caregivers, and their families more easily.

> *I started with this idea that there are certain patients, maybe like homebound people, who would be very difficult for them to come into clinic. And so, these kinds of people, I can provide service to who I wouldn't have been able to otherwise. Patients who are [reticent] to be in person because of the pandemic would feel comfortable that way. You know…there are some people who I would not have been able to evaluate if it were not for virtual meet means.* [Geriatric care professional 1]

> *I think one thing is because it works so well for their schedule, for their lives. Now that they can call, they can maybe work in the morning, and then they have a break from like, say, 10 to 11…So, we've had a lot more people just working the same day, but they had a break, or they took an hour off work, and they were able to just do the Zoom or the virtual, and then it worked out quite well. So yeah, I think for families and for caregivers, it was definitely a convenience.* [Geriatric care professional 26]

Adaptability

Many statements revealed that geriatric care professionals found ways to adapt their assessments virtually. Adapted methods included having the family members or caregivers assist patients in conducting tests, developing different backup communication plans, using modified clinical assessment tools, or collecting more information that was presently available.

> *Often what I was doing with the family members' assistance was just asking them to test strength. So, asking the [patient to] put their hands up, and then having the family member just press down and tell me is there resistance there, or do they just collapse? And oftentimes, I can see that over the video if they just collapse. So those are kind of the things that we would collect over video.* [Geriatric care professional 15]

> *I surprised myself that you can actually do geriatric psychiatry for the most part, on a video, and/or a telephone, which is about half my patients who do not have a computer. And maybe another 10% who don't know how to use it when they own it, and so, I conduct at least half of my interviews by telephone rather than video.* [Geriatric psychiatrist 17]

> *And then when we come to the physical exam, then…I would mute the OTN, so there's no feedback, and I would talk to [the patient] on the phone and watch them on the video. So, there's a lot of creativity that needs to happen in order to make these things go smoothly when, and not all the elderly people have a younger person, like a family member, who can physically go and help them get online.* [Geriatric care professional 20]

> *The cognitive piece, we have several tools, which can be administered virtually, like [the] MoCA [Montreal Cognitive Assessment], sharing the screen on Zoom, and you can guide people through the virtual exam. Even on the telephone, there are there are ways of doing certain [parts] of the mental status exams…you can certainly assess for depression because you ask people questions about that.* [Geriatric care professional 13]

> *So, what I've been doing is even more detailed functional history, particularly focusing on what can you do, but how has it changed over the last 6 months and things like that to see if there's a progressive nature, which I think is sort of a…you know, that's what we're worried about with cognitive decline is that there's going to be functional decline. So, I have substituted in the virtual platform [a] more detailed, functional history instead of doing [a] detailed cognitive history.* [Geriatric care professional 5]

Complexity

The complexities of using telemedicine technologies with older patients, their caregivers, and their families was an evident barrier. Geriatric care professionals discussed the following challenges: (1) navigating the transition from in-person consultations to virtual care; (2) establishing interpersonal connections with their patients for new consults; (3) difficulty conducting comprehensive geriatric assessments (CGAs), especially the physical exam component; (4) engaging with older patients with sensory or cognitive impairments or behavioral issues; and (5) gathering sufficient collateral information from caregivers, families, or information systems.

> *So initially, our nurses and even our admin, in terms of booking, there have been concerns about, well, who are we supposed to be emailing? Who are we calling? And trying to make sure we have all the right players in place because we can't see Power of Attorney documents. We need to be making decisions regarding the [institutional] hierarchy. We don't know all of that stuff up front.* [Geriatric care professional 1]

> *Part of medicine is the patient interaction, the physical exam, seeing how the person walks and moves and talks. That is one element that is missing is you cannot examine the patient beyond the very, very basic exam. Even the cognitive paper exams are limited. So, all of that needs to be taken very much into account. Certain assessments are just not going to be possible...are not going to be as accurate*

XSL•FO
RenderX

because of that absence. [Geriatric care professional 4]

The problem of course is not all doctors' files are available there. And none of the family doctors that I know of have any of their notes available on Connecting Ontario. So, you miss that real background importance that family doctors have. [Geriatric care professional 5]

The other kind of major area or barrier is that sensory impairments like hearing impairment or visual impairments certainly were barriers. And patients with cognitive impairment…it's already, like, it's an unfamiliar person, potentially, especially for a new consult. And it's a sort of disembodied head, like over the screen, and they sort of don't necessarily know what's going on. And so that was a little bit more difficult to establish rapport. [Geriatric care professional 8]

Again, if patients are very cognitively impaired, they tend not to understand what's happening. I had one lady who, like, was literally running away from the staff member with the camera because she was quite paranoid, and she thought she was being filmed and she was kind of covering her face. [Geriatric psychiatrist 25]

The biggest challenge is actually building a human connection with the patient. So, that's been very difficult; especially we have a lot of new consults, and we're trying to bring as few people into the hospital as possible. So, they hear my voice. They will sometimes see me on video, and then same with our nurse practitioners and our occupational therapist in [the] clinic. But, it's not the same. So, they don't get that same kind of connection. You don't build those same bonds. [Geriatric care professional 19]

Design Quality and Packaging

Geriatric care professionals used a variety of videoconferencing platforms and the telephone to conduct virtual care. However, many statements revealed that connectivity issues were often still a barrier for both geriatric care professionals and their older patients in using videoconferencing platforms.

Regarding the types of telemedicine technologies, many statements indicated a preference for videoconferencing over telephone communications as the geriatric care professionals could see their older patients and their living environments. Many statements also revealed that geriatric care professionals would like videoconferencing platforms to have the ability to facilitate more interactions with their older patients. However, some indicated their older patients preferred telephone communications. In addition, some found that their older patients with hearing impairments could hear better since there was the ability to adjust volume. Few geriatric care professionals indicated they used email communications or text messages with their older patients.

So, I think the connectivity has been a major issue. So oftentimes, you'll lose audio or things are so delayed, that it's really a barrier to communicating with the individual on the other side. [Geriatric care professional 1]

Like it would be actually, really nice to have some sort of digital interface where I could write something on the screen, and they could like circle it, or I can show them the cognitive testing on the screen, and they could like draw it on the screen. So, there's a bit more interaction. [Geriatric care professional 15]

The only thing we can do with them is telephone, okay with this, which is quite suboptimal because you can't actually see how they're doing. And they will just tell you whatever. And the patients don't know to report certain symptoms that may be concerning on the telephone, whereas, like, if you have a visual of the patient, you're more likely to not miss something, for example. [Geriatric care professional 20]

So, to meet somebody new and to be spilling out your guts when you don't see the person, I think it is very hard. I think that is why in those cases, the video is really, really important. [Geriatric psychiatrist 17]

Domain II: Outer Setting

External Policy and Incentives

Confidentiality and the security of networks were not a major concern for geriatric care professionals. Several geriatric care professionals questioned the future payment model and discussed the need for more guidance on the billing process for virtual visits.

So, there is a statement that we use from OMA [Ontario Medical Association]. So, we read that out to [the patient], or I usually send it to [the patient] by email. And I then document that in my consult note that your consent was reviewed and accepted. [Geriatric care professional 7]

And unless the government is prepared to pay a hell of a lot more for the geriatricians' time, or, and can continue with these billing codes that they have suddenly sprung up. [Geriatric care professional 13]

I do think that if you're going to have a certain service, and meet a certain standard, then be reimbursed at a certain level because there are medical legal implications as well as providing appropriate service. We have to be very clear on what we can and cannot provide, and what we should [and] should not be reimbursed for. [Geriatric care professional 28]

Patient Needs and Resources

Many statements revealed that telemedicine visits are more effective if the patient has their own monitoring devices that could provide clinical health information, such as vital signs, and a caregiver, family member, or health professional to assist the patient during the visit. However, many expressed concerns about the support and available resources for their older patients to use telemedicine. Several also expressed concerns about the

XSL•FO
RenderX

health and technology literacy of their older patients and caregivers.

> *So, blood pressure, usually I can get by because most families will have a blood pressure machine at home, which I recommend they bring with them so that they're able to do that for me virtually.* [Geriatric care professional 3]

> *I think a huge portion of the people I saw had caregivers who were helping navigate this, and, and the few that didn't, like some of them were able to, but those were the less impaired. The more impaired who weren't able to, if we didn't have nurses who were able to go there to support them to get onto the video call, there'd be no way.* [Geriatric care professional 15]

> *The big thing is that you are missing a lot of people that you don't even know, like all those who don't have technology, or all the people that may live in public housing, that are poor…those are individuals who probably would have come to the hospital, but who don't have the web; who don't have the iPhones, or iPads or, or, high technology, and we're missing them.* [Geriatric care professional 13]

> *People who are more health literate, and more technology literate, and have reasonable education, of course, those are moderately correlated with each other, are probably the best ones to be able to do the televisits with ideally…if they don't have a caregiver and they live alone and they are cognitively impaired or they don't speak English fluently…they are [the] ones that I think would be less well served by telemedicine.* [Geriatric care professional 28]

Domain III: Inner Setting

Networks and Communications

Many statements revealed geriatric care professionals relied on quality collateral information (eg, patient medical history) derived from their older patients' caregivers, families, referring physicians, local team members, and electronic medical record (EMR) systems.

> *I think that in the nature of how geriatrics sort of works in general, you can get a lot of information just from, sort of, descriptive, you know, scenarios and, sort of, gaining that information, enough to make significant changes and significant improvements. And even though it would not be perfect, I find that that there is still a lot of good work that can be done.* [Geriatric care professional 10]

> *Then, you mentioned the medication compliance. So that's where collateral is really important. So, we really rely on family and caregivers to tell us that, you know, the blister pack…they're pretty reliable with it…or I came by the other day, and there's three days of missing meds. So, we can't physically see the blister pack but except for video. Actually, if they have video, sometimes we do have them show it to us. But usually, the collateral there can help verify that, I think, almost just as well as if we were there in person.* [Geriatric care professional 19]

Culture

Prior to the COVID-19 pandemic, using telemedicine technologies was not the norm for geriatric care professionals. Geriatric care professionals preferred to be on-site with their patients and fellow colleagues.

> *I had historically rejected participating in the Ontario Telemedicine Network. I just wasn't interested. It was a variation on home visits from my viewpoint. And I prefer to sit and do what I did, which was see the person with their family, have that direct interaction, and proceed from there.* [Geriatric care professional 11]

Implementation Climate

This construct was broken into two subconstructs: tension for change and compatibility.

Tension for Change

The rapid implementation of telemedicine technologies by geriatric care professionals across Ontario was due to the COVID-19 pandemic.

> *Now when COVID hit, I had no choice, and we all went virtually.* [Geriatric care professional 10]

Compatibility

Many geriatric care professionals perceived that telemedicine technologies are highly compatible for addressing polypharmacy issues, effectively conducting follow-up care, and inquiring about patient medical histories. Geriatric psychiatrists found telemedicine technologies to be compatible with their clinical practice.

> *And so often, I think it's good for maybe follow-ups where, especially if it's a complex case, good to see the person in person. But then, if you just want to follow up and see how the pain is, see how they're doing cognitively, then you can do that very comfortably virtually.* [Geriatric care professional 4]

> *So, one of the big issues that older adults often face is polypharmacy, and with, like, video chatting, audio chatting, we were still able to review their medications, review the indications, side effects, what issues they were having, whether it was timing, being in bottles, it's not blister-packed. That was a big piece of a lot of the assessments or is still a big piece of a lot of the assessments. So, that was very, very helpful to still be able to do that part of the assessment.* [Geriatric care professional 15]

> *And I think it might be very different if you're geriatric physician, who [deals] with a physical problem, as opposed to a psychiatrist, where most of the [problems] we deal with are mental or psychological and, therefore, can be assessed by questions.* [Geriatric psychiatrist 14]

XSL•FO
RenderX

Readiness for Implementation

Many geriatric care professionals found the transition to virtual care visits was unexpected and sudden. However, a few statements also revealed that those who had previous experience with telemedicine use found the transition to be smooth.

> *I've already started using that in my training at [Medical Institution] before transitioning into practice. So, it wasn't a big transition for me, and I find it pretty easy to use.* [Geriatric care professional 3]

> *The training and my comfort in terms of making assessments in person that had been honed, both consciously and subconsciously, over 20 years, [have] now been abruptly changed to filter through a screen.* [Geriatric care professional 12]

This construct was also broken down into two subconstructs: available resources and access to knowledge.

Access to Knowledge

Several geriatric care professionals indicated they could adopt the training on telemedicine into their clinical practice, but there was still a learning curve.

> *Most of the session was talking about how challenging it is, which we all knew, it was very challenging…Some of the cognitive assessment ideas we got from that part of the workshop and incorporated them and just kind of adopted it from there.* [Geriatric care professional 6]

> *I think there were opportunities by OMA. There were webinars. And so yes, if somebody really needed to learn it or had questions, I think there were opportunities available for them. But of course, you had to do your work. There was a learning curve. You need to get used to it.* [Geriatric care professional 7]

Available Resources

Many geriatric care professionals had available resources and support for using communication infrastructure, standardized clinical assessment tools, and training on virtual care. A few did not initially have infrastructure available to them.

> *We did have the appropriate support in our hospital for OTN connections. We did have appropriate support in the hospital to provide the technical support to be able to do all of this, telemedicine from home actually, from my office, which currently is at home. And this was very helpful.* [Geriatric care professional 2]

Domain IV: Individual Characteristics

Knowledge and Beliefs About the Intervention

Many geriatric care professionals perceived that virtual care visits will be incorporated into their clinical practice in the future due to their benefits in reaching their older patients.

> *I think that there's definitely some of the benefits that I think are helpful…is that we have more options now. I think that it'll sort of carry over. I think, post–COVID, of having sort of the options to have different pathways to see our patients. If for whatever reason people can't come in, then our options were home-visiting teams, or things like that, that may have [been much] more limited in the sense of [a] longer waitlist. So, I think that that benefited that accessibility.* [Geriatric care professional 10]

Self-Efficacy

Many geriatric care professionals were confident in using telemedicine technologies to meet the care needs of their older patients with complex conditions, their caregivers, and their families. However, some were still apprehensive about their ability to conduct care virtually.

> *But like, I'd say, like 90% of the encounters, I was pretty satisfied with that I had achieved kind of the same level of assessment that I would have otherwise.[Geriatric care professional 8]*

> *And so aside from accuracy of diagnosis, I wonder if my therapeutic presence, which can be hard to quantify, is lost over a virtual platform? Or does the individual feel the same degree of therapeutic presence with an office virtual assessment?* [Geriatric care professional 12]

Domain V: Implementation Process

Engaging

This construct was broken down into two subconstructs: champions and external change agents.

Champions and External Change Agents

Several statements indicated that having a champion in the team or an external role model helped facilitate the implementation of telemedicine technologies.

> *We, as a clinic, were very lucky to have a clinician, which was just on top of all of these new changes, and [were] able to switch from seeing patients in person to telemedicine.* [Geriatric care professional 2]

Executing

Several geriatric care professionals found that they were more efficient with time during the virtual care visit, but it did not increase their patient capacity. Some statements revealed that there was additional follow-up work required with telemedicine use, especially if the older patient needed to be followed up in person.

> *When you're in [an in-person] clinic situation, the nurse or someone is going pop their head in and say, "[Doctor], the next patient is waiting." In a virtual, I've no one managing my time aside from me. So, I am much more efficient. If I have a 30-minute telephone assessment, I'm out at 30…because historically, I would gauge my time, and it was just about it.* [Geriatric care professional 11]

> *But, I guess the biggest increase to workload is that if I determine that we need to, I need to, see them,*

XSL•FO
RenderX

> *then it's another visit soon afterwards. And oftentimes, it's a home visit. And yeah, that adds time.* [Geriatric care professional 6]

Reflecting and Evaluating

Several statements revealed that geriatric care professionals would frequently reconvene with their clinical teams or peers to evaluate their experiences with telemedicine use.

> *And we would also meet with the RGP, which is the Regional Geriatric Program. And it was weekly meetings to kind of discuss what's working, what's not working. How are you guys using your referral forms? How are you generating email addresses? So, it was a lot of communication within the city, interestingly enough, to actually get these programs up and going. And it was cool because it was great to get those perspectives from interdisciplinary teams. I think that was the most important piece, that it was all members, so physician, administrative assistant, nurse practitioner, OT, PT, everybody was feeling it. So, we all had to sort of pitch in and collaborate.* [Geriatric care professional 23]

Discussion

Principal Findings

In this study, the CFIR was used to develop a comprehensive narrative of the current experiences of geriatric care professionals in routinely using telemedicine technologies in Ontario in the light of the COVID-19 pandemic. Although 5 barriers and 1 neutral construct were identified, so too were 13 facilitators.

This mixed methods study adds to the growing literature on the use of telemedicine to provide geriatric care before and during the COVID-19 pandemic, and we found similar findings to other recent studies that also explored geriatric care professional experiences with the use of telemedicine during the COVID-19 pandemic [22,25-29]. The ubiquitous transition to using telemedicine in the provision of geriatric care was uniformly driven by the COVID-19 pandemic. Geriatric care professionals faced an initial learning curve as they learned to incorporate telemedicine technologies into their routine clinical practices. However, for geriatric psychiatrists specifically, there was a more seamless transition to virtual care due to the nature of their clinical practice that has always been more amenable to the use of telemedicine technologies, which differs from geriatricians and geriatric nurse practitioners. For example, the physical examination is usually not necessary to complete a psychiatric assessment [30,31]. Meanwhile, the main challenge for geriatric psychiatrists was using telemedicine technologies with older patients with severe cognitive impairment. In a systematic review on telemedicine and dementia, Sekhon et al [15] had identified that in-person consultations are more appropriate for this subset of older patients with complex conditions.

Our study findings also raised additional unique insights being experienced by Ontario geriatric care professionals. Notably, our study was able to explore the range of strategies adopted by Ontario geriatric care professionals to complete their clinical assessments virtually, whereas other recent studies have largely focused on navigating the technological aspect of telemedicine use to overcome barriers [22,26]. As noted in *Adaptability*, Ontario geriatric care professionals quickly adopted the use of validated clinical tools that could enable them to virtually conduct their assessments or better prioritize assessment components when certain aspects were hindered by the challenges in using telemedicine technologies. Thus, with regard to improving telemedicine technologies to facilitate the more effective provision of geriatric care, Ontario geriatric care professionals would like videoconferencing platforms to have the ability to facilitate more interactions with their older patients, such as the capability to see how their older patients complete the actual written exercises in the validated clinical tools they use. Observing how their older patients actually complete these exercises, such as drawing a picture or connecting dots, provides important insights for the geriatric care professionals regarding the physical and cognitive abilities of their older patients. Another notable finding to support overcoming identified barriers was around the role of collateral information derived from caregivers, friends, family members, referring colleagues, and EMR systems. The responses of our study participants illuminated the importance of collateral information, as discussed in the *Networks and Communications* construct which played a crucial role across the whole implementation process in the virtual delivery of care. However, gathering sufficient collateral information was a complexity for our study participants, while Watt et al [28], in a recent study, had found that the persistent need to collect collateral information is a complexity for virtual care. Nevertheless, collateral information helps provide a comprehensive overview of the patient's medical and social history for geriatric care professionals without needing to see the patient in person to derive this. Furthermore, linkable EMR data were associated with more opportunities for understanding the patient journey through the care continuum [32]. Hence, geriatric care professionals were often still able to make effective clinical decisions virtually for their older patients when given sufficient collateral information that helped compensate for the other factors that can limit the usefulness of telemedicine. Watt et al [28] had also indicated that geriatric care professionals found collateral history to be particularly useful for telephone assessments in which visual assessment was not possible.

Due to the inherent challenges that exist in using telemedicine technologies, the *Compatibility* construct revealed a consensus from the responses of study participants that the role of telemedicine technology for Ontario geriatric care professionals was more appropriate for follow-up visits. Participants reasoned that follow-up visits do not require as comprehensive assessments as an initial consultation that would likely have components that benefit from an in-person assessment, such as a physical examination or certain cognitive tests. This aligns with findings by Watt et al [28] and studies that have evaluated physicians' experiences in telehealth visits with older patients in the context of US practice [26,29].

Regarding the observed perceptions of patient needs and resources, the responses of this study's participants echoed concerns around the "digital divide," which continues to be a

XSL•FO
RenderX

major barrier for older persons to using telemedicine technologies in Ontario [8,25,26,33]. In particular, their responses revealed important necessary aspects for an effective virtual care visit to take place without the assistance of a health care facilitator present with their older patients. Many of our study participants had indicated that their older patients often relied on the presence of a caregiver or a family member to access the communication technology and to manage the virtual care visit. This is also supported in recent studies in which primary care physicians found the assistance of family members and caregivers to be helpful in the facilitation of the telehealth visit for their older patients [22,29]. Additionally, our study found that the visits were even more effective for older patients who had monitoring devices that could provide basic health information, but not all older patients owned these devices. However, the varied level of health and technology literacy of their older patients and their caregivers or family members posed challenges for the ability of geriatric care professionals to collect information for their clinical assessments virtually.

The responses of our study participants suggest that there exist three necessary aspects to achieve an effective virtual care visit for both geriatric care professionals and their older patients: (1) access to the telemedicine visit–enabling technology (smartphones, tablets, computers, or telephone); (2) access to health-monitoring equipment to provide basic health information, such as blood pressure monitors; and (3) appropriate health and technology literacy amongst older patients and their caregivers or family members. Essentially, older patients or their caregiver or family member would need to assume the traditional role of the on-site health care facilitator. The reality, however, is that only a certain portion of older Ontarians have the means and ability to support the effective use of telemedicine services. It is less likely for older persons to use telemedicine technologies if they lack confidence with using related technologies [11,22,26] or to receive virtual care visits via videoconferencing if their caregivers could not be present [34]. Recent studies also indicate that older persons at the lower end of the socioeconomic spectrum are often overlooked as they lack equitable access to the appropriate resources and support to facilitate virtual care visits [25,26,35], and geriatric care professionals have observed worsening of this disparity during the COVID-19 pandemic [25]. This is important to note since a main purpose of telemedicine technologies is to bridge the gap in care accessibility for older persons who live in low-resource settings [36].

Lastly, an important consideration is how evolving policies and incentives could fundamentally change the landscape for providing virtual care visits in Ontario. As discussed in the External Policy and Incentives section, our study's geriatric care professionals raised concerns, including around the ambiguity about future billing processes for the provision of virtual care visits. Although the Government of Ontario quickly implemented temporary billing codes and guidelines to facilitate the provision of virtual care visits at the start of the COVID-19 pandemic [37], the future funding model for virtual care visits in Ontario will largely reflect the recent impact on the use of telemedicine technologies due to the COVID-19 pandemic [2,26]. If Ontario geriatric care professionals continue to embrace the use of virtual care visits, future funding policies will need to determine how to broadly support the appropriate use of telemedicine to provide high-quality geriatric care, while recognizing there still exist socioeconomic barriers to accessing it and trade-offs related to its use [28].

Strengths and Limitations

This is the largest known study pertinent to the real-world experience of geriatric care professionals using a wide range of telemedicine technologies in the light of and during the first year and a half of the COVID-19 pandemic. This study further used the CFIR to provide a comprehensive overview of the various strategies geriatric care professionals have used to overcome the complexities surrounding the provision of outpatient virtual care with older persons, their caregivers, and their families. Another strength is that this study included a wide-ranging age group of geriatric care professionals. In addition, this study primarily focused on the experiences of geriatric care professionals and did not evaluate older patients' and their caregivers' perspectives. Yet, despite the lack of these perspectives, responses across all the constructs were effective in revealing the various practice changes and strategies used to address the diverse needs of older patients with complex conditions in the virtual care setting.

There are several limitations to the study. First, the findings are limited to the experiences of geriatric care professionals in Canada's Greater Toronto Area. In addition, the experience of nonmedical or nursing allied health professionals, who are also integral members of some geriatric care teams, was not included. Nevertheless, the majority of Ontario's geriatric care professionals work in the Greater Toronto Area, with the vast majority being geriatricians, geriatric psychiatrists, and nursing care professionals who were included in this study. In addition, this study's CFIR construct *Reflecting and Evaluating* indicated that geriatric care professionals are continuously evaluating their own experiences that shape their knowledge and beliefs about the use of telemedicine in their practices. Hence, the findings presented herein should represent a snapshot of the current needs of geriatric care professionals that will likely evolve as we continue to navigate the opportunities to using telemedicine technologies to deliver geriatric care.

Recommendations

Based on the findings of this study, we offer the following four recommendations to support the continued and enhanced use of telemedicine technologies by geriatric care professionals in providing care to older patients, their caregivers, and their families:

1. Continuing training and education for geriatric care professionals in the use of telemedicine technologies is needed: Prior training on telemedicine use had helped facilitate a smoother transition for geriatric care professionals during the rapid transition to predominantly providing virtual care visits at the outset of the COVID-19 pandemic. In addition, the use of telehealth interventions relies on the experiences of clinicians in using the technology as intended [16]. Hence, continuing education can provide new learning opportunities for the best use of

XSL•FO
RenderX

telemedicine technologies for geriatric care professionals [22].

2. Training in the use of telemedicine technologies is needed for older patients, their caregivers, and their families, as well as on how to collect basic health information that may facilitate a telemedicine assessment. This could help alleviate some of the challenges in obtaining clinical information and further enhance the feasibility of virtual care visits without the presence of a health care facilitator with the patient. As McLean et al [11] noted, providing basic training for older patients, their caregivers, and their families could help them better navigate and feel more comfortable in using various telemedicine technologies.
3. Health care systems should maintain virtual care visits as an option available to older patients, their caregivers, and their families, with geriatric care professionals when this option may represent an equally or better way to facilitate care. This recognizes that virtual care visits give older patients, their caregivers, and their families and geriatric care professionals more flexibility to both provide and receive care. It should represent a mechanism through which to provide older patients, their families, and caregivers with the appropriate community infrastructure supports that could help reduce barriers for older patients, their caregivers, and their families in accessing telemedicine technologies.
4. Ensure that future reimbursement models to enable telemedicine or virtual care visits are financially sustainable for geriatric care professionals. Virtual care visits will likely be incorporated into the future provision of geriatric care in Ontario. Hence, temporary payment models will likely transition into long-term ones, and policy makers will need to ensure that the long-term methods of funding the provision of telemedicine-based care are financially sustainable, while ensuring the needs of geriatric care professionals and older persons can also be met.

Conclusion

The sudden need to find alternative ways to provide care in safe and effective ways at the outset of the COVID-19 pandemic forced health care systems worldwide to enable the rapid and widespread adoption and use of telemedicine technologies by geriatric care and other health care professionals. Overall, this study found that Ontario geriatric care professionals could adapt the use of telemedicine technologies to provide virtual care to meet the complex needs of their older patients, but there also exists a threshold in their ability to effectively provide geriatric care using telemedicine technologies. Indeed, geriatric care professionals have been found to perceive telemedicine technologies or virtual care methods to be more appropriate in the provision of follow-up visits that do not usually require specific assessments that are better done in-person. However, this study also noted that there are also various additional issues that will prohibit the greater widespread and permanent adoption of telemedicine technologies in Ontario, especially in the provision of geriatric care, unless specifically addressed. Further research on addressing older patient equity and inclusion, medical information infrastructure, and economic policies will be beneficial for understanding the best practices for supporting the use of telemedicine technologies to provide both more effective and equitable geriatric care.

Acknowledgments

The study was supported by Toronto Metropolitan University ’s School of Health Services Management and the National Institute on Ageing. The authors would like to thank the geriatric care professionals who participated in this study.

Authors' Contributions

WC led the overall study conception and design, data collection, and data analysis and drafted the manuscript. AF contributed to data analysis and provided qualitative data guidance and manuscript revisions. MN, PN, and SS contributed to the study design and provided manuscript revisions. All authors provided their final approval for the permission to publish.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Geriatric Care Professional Semistructured Interview Guide.
[PDF File (Adobe PDF File), 164 KB - aging_v5i3e34952_app1.pdf]

References

1. Dunning J, Dhuper M, Sinha S. The NIA's "Iron Ring" Guidance for Protecting Older Canadians in Long-Term Care and Congregate Living Settings. 2020. URL: https://static1.squarespace.com/static/5c2fa7b03917eed9b5a436d8/t/5f6ca4e9a0cbe5120880c889/1600955625522/NIA%2BIron%2BRing%2BGuidance%2BDocument%2BSeptember2020.pdf [accessed 2022-07-28]
2. Glazier RH, Green ME, Wu FC, Frymire E, Kopp A, Kiran T. Shifts in office and virtual primary care during the early COVID-19 pandemic in Ontario, Canada. CMAJ 2021 Feb 08;193(6):E200-E210 [FREE Full text] [doi: 10.1503/cmaj.202303] [Medline: 33558406]

XSL•FO
RenderX

3. Azad N, Amos S, Milne K, Power B. Telemedicine in a rural memory disorder clinic-remote management of patients with dementia. Can Geriatr J 2012 Dec 20;15(4):96-100 [FREE Full text] [doi: 10.5770/cgj.15.28] [Medline: 23259023]
4. Mierdel S, Owen K. Telehomecare reduces ER use and hospitalizations at William Osler Health System. Stud Health Technol Inform 2015;209:102-108. [doi: 10.3233/978-1-61499-505-0-102]
5. Chu C, Cram P, Pang A, Stamenova V, Tadrous M, Bhatia RS. Rural telemedicine use before and during the COVID-19 pandemic: repeated cross-sectional study. J Med Internet Res 2021 Apr 05;23(4):e26960 [FREE Full text] [doi: 10.2196/26960] [Medline: 33769942]
6. Hunting G, Shahid N, Sahakyan Y, Fan I, Moneypenny CR, Stanimirovic A, et al. A multi-level qualitative analysis of telehomecare in Ontario: challenges and opportunities. BMC Health Serv Res 2015 Dec 09;15:544 [FREE Full text] [doi: 10.1186/s12913-015-1196-2] [Medline: 26645639]
7. National Institute on Ageing. National Seniors Strategy 2020, Third Edition. 2020. URL: http://nationalseniorsstrategy.ca/wp-content/uploads/2020/09/NSS_2020_Third_Edition.pdf [accessed 2022-07-28]
8. Appireddy R, Khan S, Leaver C, Martin C, Jin A, Durafourt BA, et al. Home virtual visits for outpatient follow-up stroke care: cross-sectional study. J Med Internet Res 2019 Oct 07;21(10):e13734 [FREE Full text] [doi: 10.2196/13734] [Medline: 31593536]
9. Conn DK, Madan R, Lam J, Patterson T, Skirten S. Program evaluation of a telepsychiatry service for older adults connecting a university-affiliated geriatric center to a rural psychogeriatric outreach service in Northwest Ontario, Canada. Int Psychogeriatr 2013 Jul 22;25(11):1795-1800. [doi: 10.1017/s104161021300118x]
10. Lee JW, Mah G, Kalia S, Miyasaki J. Emergency room visit prevention for older patients with parkinsonism in a geriatric clinic. Can J Neurol Sci 2021 Sep 13;48(5):666-675. [doi: 10.1017/cjn.2020.253] [Medline: 33183363]
11. McLean B, Hossain N, Donison V, Gray M, Durbano S, Haase K, et al. Providing medical information to older adults in a web-based environment: systematic review. JMIR Aging 2021 Feb 09;4(1):e24092 [FREE Full text] [doi: 10.2196/24092] [Medline: 33560228]
12. Batsis JA, DiMilia PR, Seo LM, Fortuna KL, Kennedy MA, Blunt HB, et al. Effectiveness of ambulatory telemedicine care in older adults: a systematic review. J Am Geriatr Soc 2019 Aug 08;67(8):1737-1749 [FREE Full text] [doi: 10.1111/jgs.15959] [Medline: 31066916]
13. Murphy R, Dennehy K, Costello M, Murphy E, Judge C, O'Donnell MJ, et al. Virtual geriatric clinics and the COVID-19 catalyst: a rapid review. Age Ageing 2020 Oct 23;49(6):907-914 [FREE Full text] [doi: 10.1093/ageing/afaa191] [Medline: 32821909]
14. Razmak J, Bélanger CH. Comparing Canadian physicians and patients on their use of e-health tools. Technol Soc 2017 Nov;51:102-112. [doi: 10.1016/j.techsoc.2017.08.007]
15. Sekhon H, Sekhon K, Launay C, Afililo M, Innocente N, Vahia I, et al. Telemedicine and the rural dementia population: A systematic reviewa. Maturitas 2021 Jan;143:105-114. [doi: 10.1016/j.maturitas.2020.09.001] [Medline: 33308615]
16. Ware P, Ross HJ, Cafazzo JA, Laporte A, Gordon K, Seto E. Evaluating the implementation of a mobile phone-based telemonitoring program: longitudinal study guided by the consolidated framework for implementation research. JMIR Mhealth Uhealth 2018 Jul 31;6(7):e10768 [FREE Full text] [doi: 10.2196/10768] [Medline: 30064970]
17. Courtin E, Knapp M. Social isolation, loneliness and health in old age: a scoping review. Health Soc Care Community 2017 May;25(3):799-812. [doi: 10.1111/hsc.12311] [Medline: 26712585]
18. Gillespie SM, Moser AL, Gokula M, Edmondson T, Rees J, Nelson D, et al. Standards for the use of telemedicine for evaluation and management of resident change of condition in the nursing home. J Am Med Dir Assoc 2019 Feb;20(2):115-122. [doi: 10.1016/j.jamda.2018.11.022] [Medline: 30691620]
19. Kendzerska T, Zhu DT, Gershon AS, Edwards JD, Peixoto C, Robillard R, et al. The effects of the health system response to the COVID-19 pandemic on chronic disease management: a narrative review. RMHP 2021 Feb;14:575-584. [doi: 10.2147/rmhp.s293471]
20. Lehoux P, Sicotte C, Denis JL, Berg M, Lacroix A. The theory of use behind telemedicine: how compatible with physicians' clinical routines? Soc Sci Med 2002 Mar;54(6):889-904. [doi: 10.1016/s0277-9536(01)00063-6] [Medline: 11996023]
21. Batsis JA, Pletcher SN, Stahl JE. Telemedicine and primary care obesity management in rural areas - innovative approach for older adults? BMC Geriatr 2017 Jan 05;17(1):6 [FREE Full text] [doi: 10.1186/s12877-016-0396-x] [Medline: 28056832]
22. Chen K, Davoodi N, Strauss D, Li M, Jiménez FN, Guthrie K, et al. Strategies to ensure continuity of care using telemedicine with older adults during COVID-19: a qualitative study of physicians in primary care and geriatrics. J Appl Gerontol 2022 Jun 16:7334648221109728 [FREE Full text] [doi: 10.1177/07334648221109728] [Medline: 35711106]
23. Varsi C, Ekstedt M, Gammon D, Ruland CM. Using the Consolidated Framework for Implementation Research to identify barriers and facilitators for the implementation of an internet-based patient-provider communication service in five settings: a qualitative study. J Med Internet Res 2015 Nov 18;17(11):e262 [FREE Full text] [doi: 10.2196/jmir.5091] [Medline: 26582138]
24. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 07;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]

XSL•FO
RenderX

25. Phimphasone-Brady P, Chiao J, Karamsetti L, Sieja A, Johnson R, Macke L, et al. Clinician and staff perspectives on potential disparities introduced by the rapid implementation of telehealth services during COVID-19: a mixed-methods analysis. Transl Behav Med 2021 Jul 29;11(7):1339-1347 [FREE Full text] [doi: 10.1093/tbm/ibab060] [Medline: 34132810]
26. Goldberg EM, Jiménez FN, Chen K, Davoodi NM, Li M, Strauss DH, et al. Telehealth was beneficial during COVID-19 for older Americans: a qualitative study with physicians. J Am Geriatr Soc 2021 Nov 20;69(11):3034-3043 [FREE Full text] [doi: 10.1111/jgs.17370] [Medline: 34245165]
27. Iyer S, Mehta P, Weith J, Hoang-Gia D, Moore J, Carlson C, et al. Converting a geriatrics clinic to virtual visits during COVID-19: a case study. J Prim Care Community Health 2021;12:21501327211000235 [FREE Full text] [doi: 10.1177/21501327211000235] [Medline: 33729044]
28. Watt J, Fahim C, Straus S, Goodarzi Z. Barriers and facilitators to virtual care in a geriatric medicine clinic: a semi-structured interview study of patient, caregiver and healthcare provider perspectives. Age Ageing 2022 Jan 06;51(1):afab218. [doi: 10.1093/ageing/afab218] [Medline: 34794177]
29. Aliberti GM, Bhatia R, Desrochers LB, Gilliam EA, Schonberg MA. Perspectives of primary care clinicians in Massachusetts on use of telemedicine with adults aged 65 and older during the COVID-19 pandemic. Prev Med Rep 2022 Apr;26:101729 [FREE Full text] [doi: 10.1016/j.pmedr.2022.101729] [Medline: 35155085]
30. Öngür D, Perlis R, Goff D. Psychiatry and COVID-19. JAMA 2020 Sep 22;324(12):1149-1150. [doi: 10.1001/jama.2020.14294] [Medline: 32960247]
31. Stamenova V, Chu C, Pang A, Fang J, Shakeri A, Cram P, et al. Virtual care use during the COVID-19 pandemic and its impact on healthcare utilization in patients with chronic disease: a population-based repeated cross-sectional study. PLoS One 2022;17(4):e0267218 [FREE Full text] [doi: 10.1371/journal.pone.0267218] [Medline: 35468168]
32. Rayner J, Khan T, Chan C, Wu C. Illustrating the patient journey through the care continuum: leveraging structured primary care electronic medical record (EMR) data in Ontario, Canada using chronic obstructive pulmonary disease as a case study. Int J Med Inform 2020 Aug;140:104159. [doi: 10.1016/j.ijmedinf.2020.104159] [Medline: 32473567]
33. Zulman D, Wong E, Slightam C, Gregory A, Jacobs J, Kimerling R, et al. Making connections: nationwide implementation of video telehealth tablets to address access barriers in veterans. JAMIA Open 2019 Oct;2(3):323-329 [FREE Full text] [doi: 10.1093/jamiaopen/ooz024] [Medline: 32766533]
34. Liu L, Goodarzi Z, Jones A, Posno R, Straus SE, Watt JA. Factors associated with virtual care access in older adults: a cross-sectional study. Age Ageing 2021 Jun 28;50(4):1412-1415 [FREE Full text] [doi: 10.1093/ageing/afab021] [Medline: 33625475]
35. Li W, Ornstein KA, Li Y, Liu B. Barriers to learning a new technology to go online among older adults during the COVID-19 pandemic. J Am Geriatr Soc 2021 Nov 29;69(11):3051-3057 [FREE Full text] [doi: 10.1111/jgs.17433] [Medline: 34409589]
36. Krishnaswami A, Beavers C, Dorsch MP, Dodson JA, Masterson Creber R, Kitsiou S, Innovations, Cardiovascular Team and the Geriatric Cardiology Councils, American College of Cardiology. Gerotechnology for older adults with cardiovascular diseases: JACC state-of-the-art review. J Am Coll Cardiol 2020 Dec 01;76(22):2650-2670 [FREE Full text] [doi: 10.1016/j.jacc.2020.09.606] [Medline: 33243384]
37. Ministry of Health of Ontario. Virtual Care Billing Information Manual. 2021. URL: https://dropbox.otn.ca/files/virtual-care-billing-information-manual.pdf [accessed 2022-07-28]

Abbreviations

CGA: Comprehensive Geriatric Assessment
CFIR: Consolidated Framework for Implementation Research
COPD: Chronic obstructive pulmonary disease
ED: Emergency department
EMR: Electronic medical record
MoCA: Montreal Cognitive Assessment
NP: Nurse Practitioner
OT: Occupational Health Therapist
OMA: Ontario Medical Association
OTN: Ontario Telemedicine Network
PT: Physical Therapist
SPSS: Statistical Package for the Social Sciences
REB: Research Ethics Board
RGP: Regional Geriatrics Program

XSL•FO
RenderX

Edited by T Leung, J Wang; submitted 11.01.22; peer-reviewed by M Stein, W Tang; comments to author 30.05.22; revised version received 04.07.22; accepted 11.07.22; published 10.08.22.

Please cite as:
Chen W, Flanagan A, Nippak PMD, Nicin M, Sinha SK
Understanding the Experience of Geriatric Care Professionals in Using Telemedicine to Care for Older Patients in Response to the COVID-19 Pandemic: Mixed Methods Study
JMIR Aging 2022;5(3):e34952
URL: https://aging.jmir.org/2022/3/e34952
doi:10.2196/34952
PMID:35830331

©Wenwen Chen, Ashley Flanagan, Pria MD Nippak, Michael Nicin, Samir K Sinha. Originally published in JMIR Aging (https://aging.jmir.org), 10.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementing Technology Literacy Programs in Retirement Homes and Residential Care Facilities: Conceptual Framework

Karen S Li[1,2], BSc; Nathan Nagallo[2,3], BSc; Erica McDonald[2,4], BSc, MSc; Colin Whaley[2,5,6], BSc, MSc; Kelly Grindrod[5], BSc, PharmD, MSc; Karla Boluk[7], BA, PhD

[1]Department of Biology, Faculty of Science, University of Waterloo, Waterloo, ON, Canada

[2]enTECH Computer Club, University of Waterloo, Waterloo, ON, Canada

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

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

[5]School of Pharmacy, Faculty of Science, University of Waterloo, Waterloo, ON, Canada

[6]Michael G. DeGroote School of Medicine, McMaster University, Hamilton, ON, Canada

[7]Department of Recreation and Leisure Studies, Faculty of Health, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Colin Whaley, BSc, MSc
Michael G. DeGroote School of Medicine
McMaster University
1280 Main Street West
Hamilton, ON, L8S 4K1
Canada
Phone: 1 905 525 9140 ext 22235
Fax: 1 905 546 0349
Email: colin.whaley@mcmaster.ca

Abstract

Background: The COVID-19 pandemic caused widespread societal disruption, with governmental stay-at-home orders resulting in people connecting more via technology rather than in person. This shift had major impacts on older adult residents staying in retirement homes and residential care facilities, where they may lack the technology literacy needed to stay connected. The enTECH Computer Club from the University of Waterloo in Ontario, Canada created a knowledge translation toolkit to support organizations interested in starting technology literacy programs (TLPs) by providing guidance and practical tips.

Objective: This paper aimed to present a framework for implementing TLPs in retirement homes and residential care facilities through expanding on the knowledge translation toolkit and the framework for person-centered care.

Methods: Major concepts relating to the creation of a TLP in retirement homes and residential care facilities were extracted from the enTECH knowledge translation toolkit. The domains from the framework for person-centered care were modified to fit a TLP context. The concepts identified from the toolkit were sorted into the three framework categories: “structure,” “process,” and “outcome.” Information from the knowledge translation toolkit were extracted into the three categories and synthesized to form foundational principles and potential actions.

Results: All 13 domains from the framework for person-centered care were redefined to shift the focus on TLP implementation, with 7 domains under “structure,” 4 domains under “process,” and 2 domains under “outcome.” Domains in the “structure” category focus on developing an organizational infrastructure to deliver a successful TLP; 10 foundational principles and 25 potential actions were identified for this category. Domains in the “process” category focus on outlining procedures taken by stakeholders involved to ensure a smooth transition from conceptualization into action; 12 foundational principles and 9 potential actions were identified for this category. Domains in the “outcome” category focus on evaluating the TLP to consider making any improvements to better serve the needs of older adults and staff; 6 foundational principles and 6 potential actions were identified for this category.

Conclusions: Several domains and their foundational principles and potential actions from the TLP framework were found to be consistent with existing literatures that encourage taking active steps to increase technology literacy in older adults. Although there may be some limitations to the components of the framework with the current state of the pandemic, starting TLPs in the

XSL•FO
RenderX

community can yield positive outcomes that will be beneficial to both older adult participants and the organization in the long term.

(JMIR Aging 2022;5(3):e34997) doi:10.2196/34997

KEYWORDS

older adult; technology; retirement home; long-term care; social connections; technology literacy program; retirement; elderly; literacy; implementation; concept; framework; knowledge translation

Introduction

Older adults' ability to stay socially connected has become more challenging as a result of the COVID-19 global pandemic. The World Health Organization and public health agencies in North America have advised all individuals to follow physical distancing measures and limit gatherings to minimize transmission of COVID-19 [1-3]. Older adults (individuals who are 65 years of age or older) are identified as a group at greater risk for severe illness from COVID-19, compared to the general public [4]. Throughout the course of the pandemic, long-term care (LTC) visitor restrictions have been implemented to protect the health of residents and staff members in these settings. Although it is important to limit the spread of COVID-19, these measures have disrupted the normal social routines of older adults, which may increase the risk of anxiety, depression, cognitive dysfunction, heart disease, and overall mortality [5,6]. Throughout the pandemic, the World Health Organization and public health agencies have encouraged individuals to maintain social connections through digital alternatives, using phones, computers, tablets, and other electronic devices, but little support is available to assist those who need help using technology for these purposes.

Previous research has identified that older adults are often interested in using technology for recreation, but feelings of apprehension and difficulty in accessing and using technology limit its uptake by older adults [7-9]. Existing literature suggests that active steps should be taken to educate and support older adults in their engagement with technology to help them build and maintain social connections with their family, friends, and the wider community [9-12]. Implementing teaching and education opportunities by volunteer organizations, for instance, that involve peer-to-peer learning and intergenerational relationships to help older adults adopt technology are recommended [5,9,13].

Although technology adoption in LTC settings increased during the pandemic, ensuring continued access to technology and technology education can maintain and increase its uptake by older adults [8]. In many regions, community organizations provide technology education and assistance to older adults at a low cost or for free, such as the enTECH Computer Club, based out of the University of Waterloo in Waterloo, Ontario, Canada. Before the pandemic, enTECH volunteers worked with residents of local LTC homes and supported older adults reaching their technology goals [13,14]. To facilitate the expansion of club activities to other locations, enTECH club members have also developed a knowledge translation toolkit in consultation with the University of Waterloo faculty [15]. The toolkit serves as a starting point for LTC homes to implement technology programming, consolidating much of the club's work.

The purpose of this paper is to describe a framework to guide the implementation of technology programming in residential care facilities and retirement home settings, using the framework for person-centered care by Santana et al [16] as a template. To complement the toolkit by Nagallo et al [15], we aimed to develop a comprehensive framework with which a technology program can be started. Additionally, we wish to frame the development and rollout of a technology literacy program (TLP) as being fundamentally person focused, leading to the selection of the framework for person-centered care. Through consultation with the enTECH team, the TLP framework maps the knowledge translation toolkit onto the framework for person-centered care, in which it provides foundational principles and potential actions to guide the framework's implementation. The implementation of technology literacy programming through this framework aims to encourage older adults to use technology and to create opportunities for fostering social connections and maintaining healthy aging.

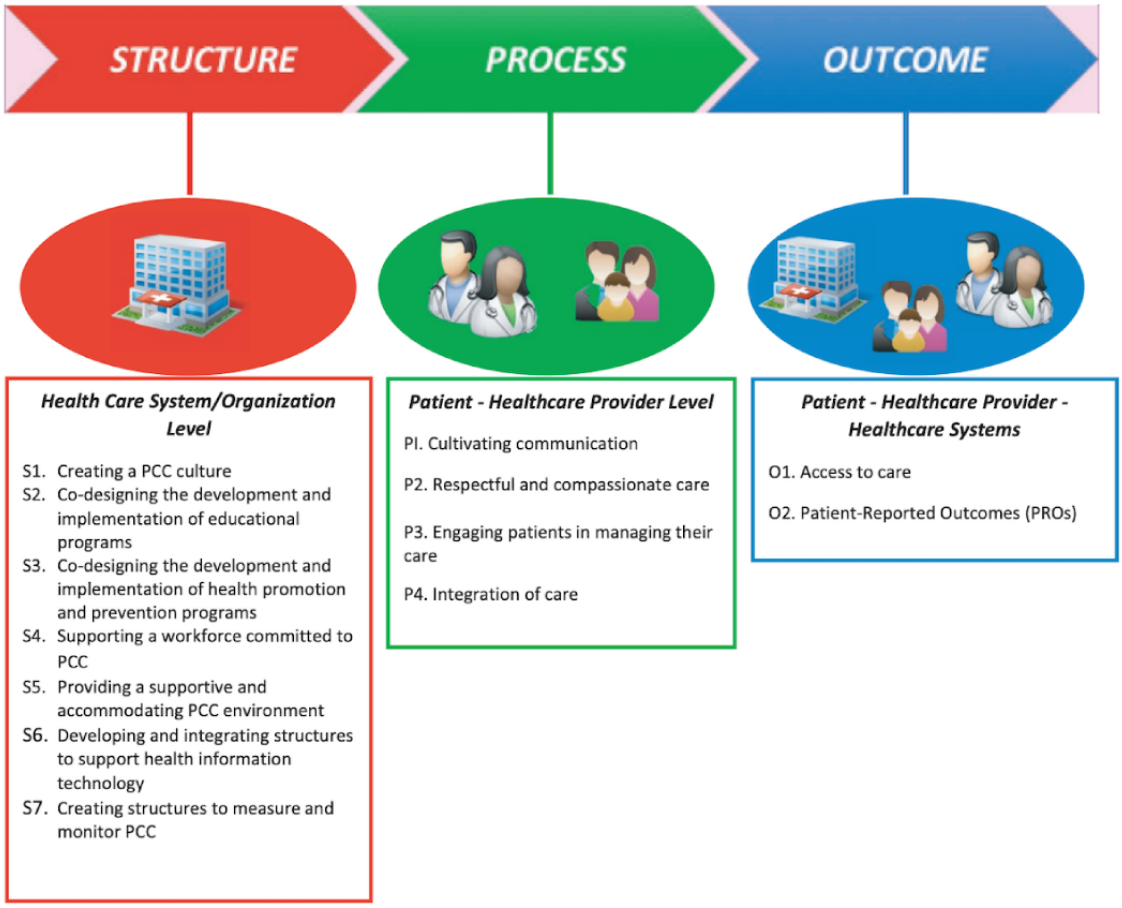
Methods

Framework Model Exploration

A search using Google Scholar was conducted during June 2020 to search for existing frameworks involving or implementing change in health care contexts, particularly with the use of technology. The framework for person-centered care by Santana et al [16] (Figure 1) was identified as having general person-centered domains that could be consistent with a TLP program for older adults, the residential care facility, and its staff. The framework for person-centered care is itself based on the Donabedian Model [17] for the assessment of quality of care, which divides assessment into "structure," "process," and "outcome" components. The framework for person-centered care was selected over other frameworks, such as those applicable to adult education (eg, Kirkpatrick model [18]) or implementation (eg, RE-AIM [19,20]), due to a focus on implementation, person-centeredness, and provider-patient partnership [16]. The framework for person-centered care contains 13 domains split between 3 themes that serve as foundational pillars to implement person-centered care in a health care environment [16].

XSL•FO
RenderX

Figure 1. Framework for person-centered care (PCC) by Santana et al [16]. PRO: patient-reported outcomes.

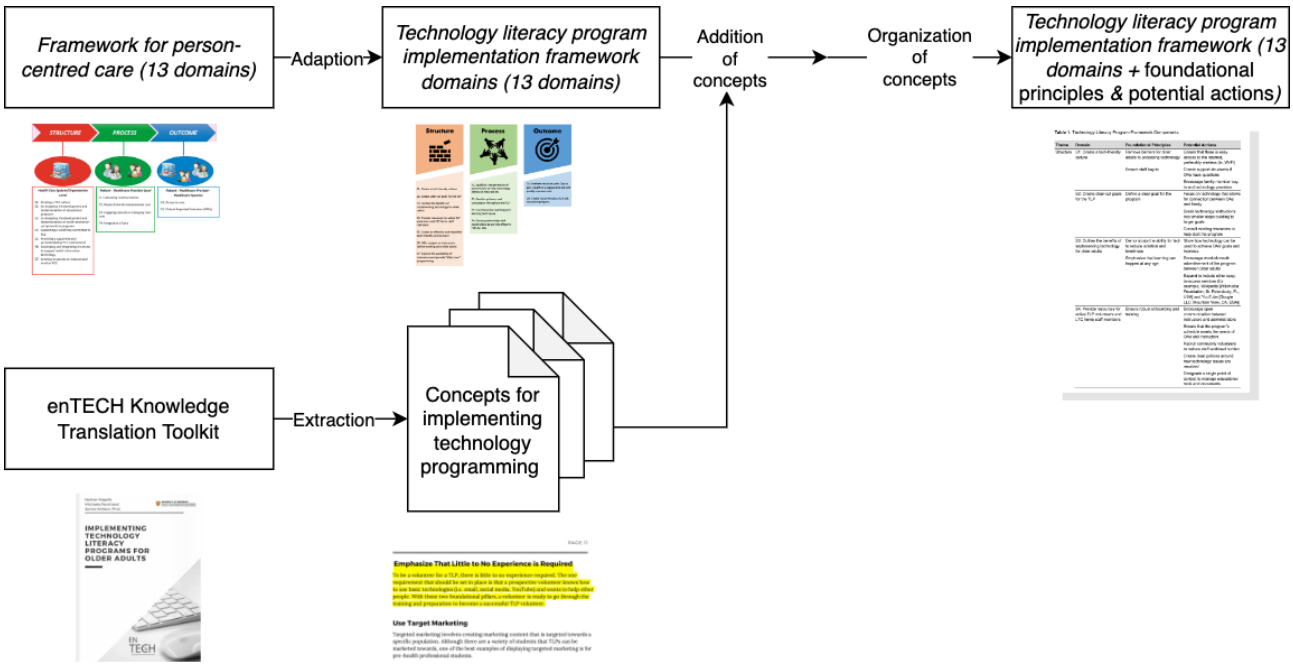


Domain Mapping

The lead author of the knowledge translation toolkit (NN) modified the 13 domains from the framework for person-centered care to reflect the requirements of starting a TLP in a LTC setting. The language in framework for person-centered care by Santana et al [16] was adapted to focus on technology literacy programming, incorporating program implementation, technology education, quality assurance, and stakeholder relations [16]. To identify relevant foundational principles and potential actions for TLPs, the 13 domains for the TLP implementation framework were inserted into a Miro whiteboard (RealtimeBoard, Inc), a cloud-based collaborative whiteboard software. Foundational principles are defined as ideal steps to be taken to implement a successful TLP. Potential actions are defined as action items to help with a particular domain, but they are not completely mandatory for a TLP to be implemented. Author KL identified and extracted every major concept from enTECH's knowledge translation toolkit, focusing on TLP; these concepts were organized under the 13 TLP domains. Authors NN, CW, and KL performed a final check of the domains and associated concepts on the Miro whiteboard. The resulting Miro whiteboard is shown in Multimedia Appendix 1. Authors KL, NN, EM, and CW finalized the categorization of the concepts into either foundational principles or potential actions for each domain. This process is detailed in Figure 2.

XSL•FO
RenderX

Figure 2. Process of developing the technology literacy program implementation framework.



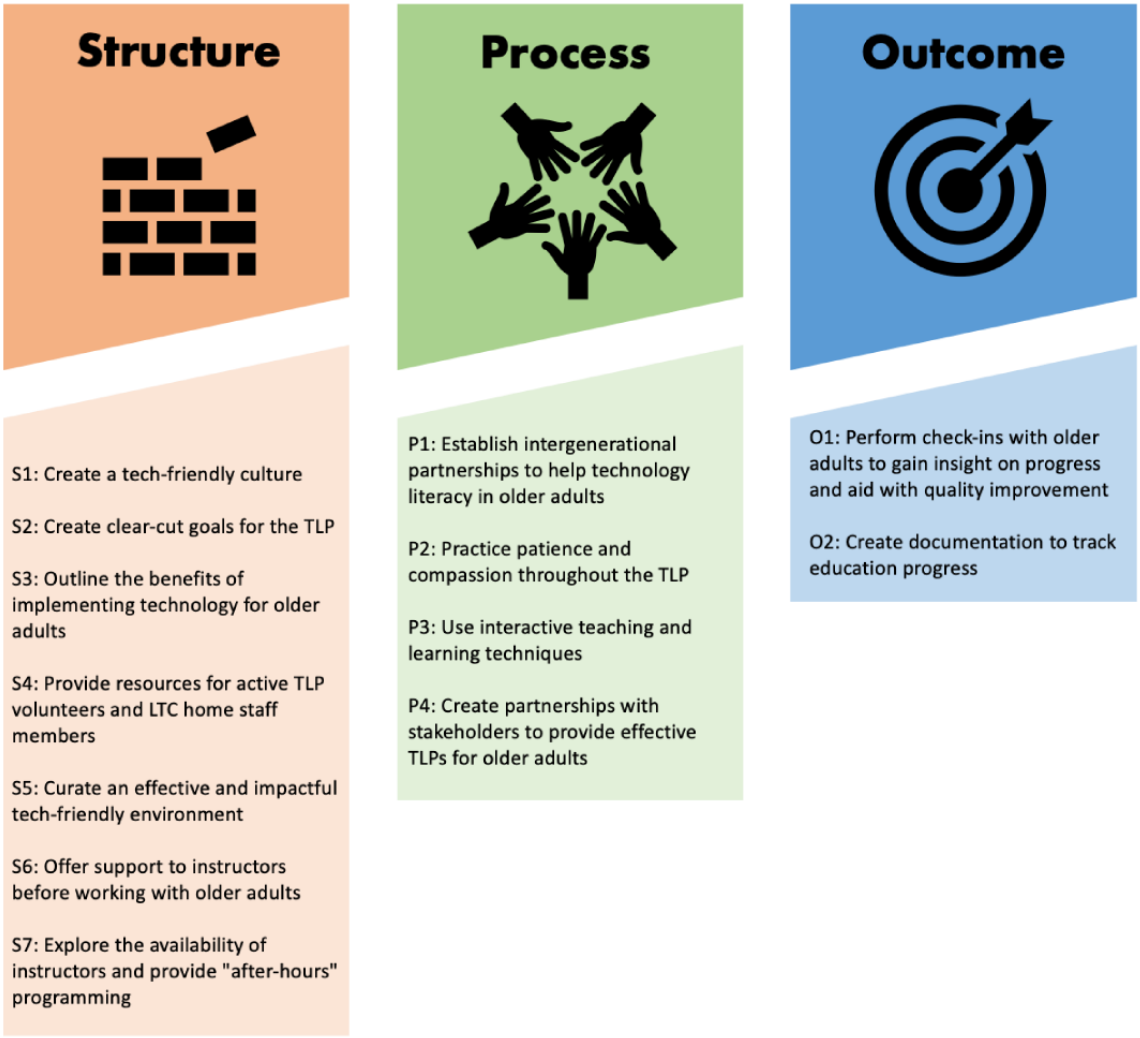
Results

Framework Overview

Similar to the framework for person-centered care [16], the TLP implementation framework uses the 3 main categories of "structure," "process," and "outcome" to sort the 13 domains needed to implement a TLP. Under these 3 categories, 7 domains were placed under "structure," 4 domains were placed under "process," and 2 domains were placed under "outcome." To build this framework, domains were used to guide discussion about foundational principles and potential actions. The 3 main categories and their domains are organized in the order that they should be executed during the implementation process. The final TLP framework model is shown in Figure 3.

XSL•FO
RenderX

Figure 3. Technology literacy program (TLP) framework.



Domains Overview

The TLP framework consists of 3 main categories that are essential to designate the domains and components of the framework; these 3 categories are “structure,” “process,” and “outcome.” The domains and components under each category are detailed in Tables 1-3.

Structure (S1-S7)

According to Donabedian [17], “structure” defines the attributes of the setting in which care takes place. It describes the facilities, financials, human resources, and organizational structure of the place where care occurs. The framework for person-centered care also uses this definition to build their domains for this category [16]. As shown in Table 1, the “structure” category for TLP implementation focuses on developing organizational infrastructure to deliver a successful TLP. This category also includes strategies on how various resources may be used to serve as the foundation of TLP implementation.

XSL•FO
RenderX

Table 1. Structure domains and components of the technology literacy program (TLP) framework.

Structure domains	Foundational principles	Potential actions
• S1: Create a tech-friendly culture	• Remove barriers to accessing technology for older adults • Ensure staff buy-in	• Ensure that there is easy access to the internet, preferably wireless internet (ie, Wi-Fi) • Create support structures if older adults have questions • Encourage family member buy-in and technology provision
• S2: Create clear-cut goals for the TLP	• Define a clear goal for the program	• Focus on technology that allows for connection between older adults and family • Break technology instructions into smaller steps building to larger goals • Consult existing resources to help start the program
• S3: Outline the benefits of implementing technology for older adults	• Demonstrate the ability for tech to reduce isolation and loneliness • Emphasize that learning can happen at any age	• Show how technology can be used to achieve older adults' goals and interests • Encourage word-of-mouth advertisement of the program between older adults • Expand to include other easy-to-access services (eg, Wikipedia and YouTube)
• S4: Provide resources for active TLP volunteers and long-term care homes staff members	• Ensure robust onboarding and training	• Encourage open communication between instructors and administrators • Ensure that the program's schedule meets the needs of older adults and instructors • Recruit community volunteers to reduce staff workload burden • Create clear policies around how technology issues are resolved • Designate a single point of contact to manage educational tools and documents
• S5: Curate an effective and impactful tech-friendly environment	• Reduce paperwork burden on instructors, understanding that some documentation might be needed • Ensure that technology is user-friendly and accounts for differences in strengths between individuals	• Have resources that can be easily accessed by instructors, including printouts • Advertise the benefits of technology to older adults via posters and town halls • Consider a buddy or group system where older adults can be taught the same content together • Listen to and act on what older adults themselves indicate they want to learn or achieve
• S6: Offer support to instructors before working with older adults	• Stress that the requirements to become an instructor are low: only basic technology literacy with common software (ie, email and YouTube) is needed	• Ensure instructors have basic tech literacy skills to help older adults • Create a straightforward onboarding process • Host practice sessions with instructors (eg, on ethical scenarios and common technology problems) • Use a web-based platform (eg, Slack and Microsoft Teams) where instructors can easily communicate, while maintaining confidentiality
• S7: Explore the availability of instructors and provide "after-hours" programming	• Schedule sessions at practical times and create opportunities for learning and follow-up if questions arise when there is no instructor available	• Create a knowledge base that older adults can access if there are no instructors available or for self-study (eg, an easy-to-access website or binder) • Reevaluate instructor availability frequently • Ensure older adults can anonymously leave feedback about programming (eg, an anonymous comment box)

Process (P1-P4)

The "process" category used in the framework for person-centered care includes the processes to deliver person-centered care [16]. Specifically, it describes strategies that can be used by health care practitioners to ensure person-centered care is being provided [16]. These steps closely resemble the action items to begin and maintain a TLP in LTC homes in the "process" category of TLP implementation. As depicted in Table 2, the domains in the "process" category outline procedures undertaken by stakeholders involved (eg, volunteers, club executives, resident facility staff, and older adults) to ensure a smooth transition from conceptualization into action.

XSL•FO
RenderX

Table 2. Process domains and components of the technology literacy program (TLP) framework.

Process domains	Foundational principles	Potential actions
• P1: Establish intergenerational partnerships to help technology literacy in older adults	• There is no set number of instructors needed to help set up a successful TLP • Instruction can be in person, digital, or over video chat or phone	• Encourage older adults to write any questions they have in between teaching sessions • Network with other community groups to identify potential volunteers
• P2: Practice patience and compassion throughout the TLP	• Encourage older adults to protect their personal information, including financial information • Encourage instructors to use their judgement and limit their support to their comfort level • Ensure that instruction allows older adults to learn at their own pace • Instructors can develop or improve their teaching, leadership, and communication skills	• Repetition and practice are key to support learning and are sometimes overlooked
• P3: Use interactive teaching and learning techniques	• Programs can be one-on-one, in small groups, or lecture based, depending on older adults and instructor's comfort and knowledge as well as room capacity limitations • Facilitation must be sensitive to the needs of older adults, including vision, hearing, and mobility challenges • Encourage instructors to adapt to different learning styles to assist older adults as best as possible • Remember that layperson terms for instructors may not be layperson terms for older adults	• Limit session to 1.5- to 2-hour blocks to minimize participant and instructor fatigue • Have diverse learning resources available for older adults and instructors
• P4: Create partnerships with stakeholders to provide effective TLPs for older adults	• Identify what existing TLPs do in other settings to develop curriculum and foster participant engagement • Be open to the potential for collaborating with existing organizations that provide digital services	• Advertise this opportunity broadly; pre-health professional students may be particularly interested in participating in it • Identify existing local supports for community programs, including organizations that may provide technology at a discounted price • Reach out to long-term care or retirement home stakeholders and the community at large to see if any organizations are looking to liquidate older technology • Designate one representative from your organization to liaise with other stakeholders for the purposes of acquiring technology donations and recruiting instructors

Outcome (O1-O2)

In the framework for person-centered care, the "outcome" category describes the effects of person-centered care on patients [16]. It also describes performance indicators of person-centered care in hopes of measuring such indicators to identify areas for improvement to better serve patients [16]. Table 3 reflects how TLP implementation shares similar attributes with the "outcome" category, as it focuses on how TLPs can benefit older adults even with diminished volunteer capacity, and how quality improvement can be used to adjust programming to better serve the needs of the community.

XSL•FO
RenderX

Table 3. Outcome domains and components of the technology literacy program framework.

Outcome domains	Foundational principles	Potential actions
• O1: Perform check-ins with older adults to gain insight on progress and aid with quality improvement	• Older adults' progress on accomplishing goals can be tracked using a web-based spreadsheet like Google Sheets • Focus on the person not the technology; technology should always come second • Validated tools like the Single Ease Question [21] can be used to evaluate the difficulty of a task • Questions or questionnaires to be filled out by older adults can be hosted via the internet, using Google Forms, to encourage older adults to practice their tech skills	• Check in with older adults to see how they feel about their progress • A brief informal postsession feedback question is recommended, even a simple "did you like this session today?"
• O2: Create documentation to track education progress	• Consider curriculum and training documents as "living documents," and iterate on them based on feedback • A "best practices" guide can be used to structure lessons more generally	• Lesson plans can be used to teach content and can optionally be used by residents in the absence of instructors • Provide training documents to instructors • Platforms like Google Docs and Notion (Notion Labs Inc) can be used to organize documents and can be updated immediately • Tracking students' progress and interests can facilitate transition between different instructors

Discussion

Overview

This paper outlines an implementation framework for TLPs in LTC and residential homes and with older adults. Building on a person-centered care framework, the TLP framework starts by outlining how the program should be structured to create a safe learning environment for older adults to learn to use technology. The framework also outlines processes that can be used to build confidence and competence among both older adult learners and instructors, including multigenerational pairings, focusing on web-based safety, and ample hands-on learning opportunities. The framework concludes with 2 domains under "outcome" that are focused on identifying the benefits to older adults and opportunities for quality improvement.

Structure

Domains S1 to S7 are aimed to create a safe learning environment for older adults to learn how to use technology. Having a positive initial experience supported by a well-established, friendly, and supportive learning environment can promote continual participation of the TLP and continual use of technology among older adults, and so domain S1 reinforces this [22]. Domain S3 concentrates on informing older adults of the benefits of implementing technology in their lives. Beyond increasing technology literacy, existing research has suggested that learning new skills and keeping an active mind may maintain cognition and mental health throughout life [23]. In addition, the "structure" category of the framework aims to reduce the social isolation and loneliness that older adults may experience during the pandemic through delivering interactive learning [24,25]. Many of the foundational principles and potential actions of domain S7 were consistent with existing findings, including those suggesting a strong training preference for self-teaching through methods such as reading a printed manual and learning through "trial and error" or "playing around" approach [22,26]. Although self-practicing after a lesson can be beneficial, the TLP framework focuses on the value of instructors to teach older adults and facilitate their learning in TLPs (domains S7 and O1). There is a focus on providing resources for program leaders so that they can effectively support older adults, as emphasized in domains S4 through S7.

Process

Domains P1 to P4 outline the means by which TLP should be operated. This includes engaging older adult learners with staff of TLP within retirement homes and residential care facilities to lead to a person-centered programming. A study conducted during the COVID-19 pandemic suggested that LTC facilities should have dedicated staff to assist residents in using information communication technologies that are already provided by the facilities, such as tablet devices [27]. Therefore, domain P1 has considered this by creating a group of instructors to teach older adult learners and facilitate the TLP operations. Existing cognitive aging research has outlined recommendations on techniques that instructors may find useful when teaching older adults to use computer software applications [28]. Many of the "process" domains align with recommendations based on that research [28]. Additional research identified reasons for older adults' negative attitudes about technology use, which were frequently associated with inconvenience with the technology device, disliking the features of technology, and security and reliability concerns [29]. Given these concerns, foundational principles from domain P2 emphasize the importance of protecting personal and financial information. Social contact or other social interactions have been shown to

XSL•FO
RenderX

be the second most important motivator for older adult learners to participate in learning activities [30]. Therefore, domains P2 and P3 aim to have older adults develop healthy interpersonal relationships with instructors to foster participant engagement. Another study that examined teaching modality themes used by student mentors to help older adults learn technology reinforces the foundational principles and the overall theme of domains P2 and P3 [22]. Finally, other research has suggested social engagements may also maintain cognitive aging and lower mortality outcomes as much as physical activity, depending on the level of social activity [31].

Outcome

Domains O1 and O2 aim to evaluate the progress of the TLP and determine if it is meeting older adults' goals to learn technology. Existing research has recognized the need for technological training programs focused on older adults to evaluate and analyze the effectiveness of teaching technology to them [32]. To enhance the experience of TLP for staff and older adults, TLPs must determine the positive feedback and improvements that can be made. The insights received from following the foundational principles and potential actions of domain O2 allow for more research to be conducted to determine the strengths and improvements needed for TLP.

Strengths and Applicability of the Framework

Governments and health care systems have been encouraging people to remain engaged with their communities via the internet during the COVID-19 pandemic [1-3]. The implementation of TLPs for older adults may help them overcome some of the known barriers to engage with technology. Increased technology literacy, fostered through TLP, may also support older adults in accessing health care digitally, an adoption of which has been accelerating in response to the pandemic [33,34].

It has been suggested by findings of a recent study that using a person-centered care approach to engage older adults with technology is crucial in creating meaningful technology-mediated enrichment experiences for them [35]. The TLP framework supports this approach as it aims to implement fundamentally person-focused TLPs, hence the framework for person-centered care was chosen as the template [16]. By organizing the domains' content into foundational principles and potential actions, we hope that organizations can identify parts of the framework that are readily applicable and can be implemented with little difficulty, which can be beneficial during times of high visitor restrictions. For instance, during visitor restrictions in LTC, organizations may begin with creating a knowledge base that older adults can access with minimal assistance if there are no instructors available, which is a potential action from domain S7. In addition, organizations can plan opportunities to expand their existing TLPs. Lastly, there is a focus on creating and externalizing partnerships to run the program to reduce costs and foster new relationships with others in the community.

Limitations

Despite the benefits of implementing TLPs using this framework, there are some drawbacks to consider. The COVID-19 pandemic has highlighted how complex it can be to support older adults through periods of intense isolation. Periods of visitor and volunteer service restrictions, such as local or global disease outbreaks, can present challenges to the implementation of TLPs. For instance, older adults may have to ask LTC staff outside of TLP for help, and this may interfere with the techniques they learned with their usual TLP instructors. Organizations implementing TLPs could consider incorporating a group of on-site staff who can assist older adults during periods of visitor and volunteer service restrictions if they see fit. Although these events cannot easily be anticipated aside from typical seasonal patterns (ie, increased probability of influenza outbreaks in fall and winter months), ensuring TLP participants are sufficiently trained such that they can accomplish basic communication tasks independently can foster independence and reduce feelings of isolation.

Furthermore, the authors' experience of TLPs are based on implementing TLPs directly and assisting with situating them in existing retirement and LTC settings as an external organization. The authors do not have experience implementing a TLP as an employee of a retirement home or LTC home. EnTECH's experience in TLP implementation is also limited to Ontario, Canada. Thus, the generalizability of this TLP framework may be limited, as the structure of residential care facilities and retirement homes outside of Ontario may differ from those in Ontario. Finally, despite the value of technology, the financial cost of the devices and additional equipment to run a TLP is a concern in both private and publicly funded residential care facilities, and that is addressed in domain P4.

Future Research

Further research should be conducted to validate the utility of this framework in retirement homes and residential care facilities. This could include evaluating the success in implementing a TLP in retirement homes and residential care facilities according to the insights on the progress and quality of the program gained from the check-ins with older adults (domain O2). Additionally, there is an interest in exploring whether older adults are able to independently apply the technology literacy skills they have developed or improved from their TLP to achieve and expand on their technology use. After the implementation of a TLP in retirement homes and residential care facilities, the publication of case studies aimed at both industry and academics is suggested to raise awareness of this framework for residential care employees and stakeholders.

Conclusions

Recognition in staying socially connected through technology has undoubtedly increased throughout the course of the COVID-19 pandemic. A framework for implementing TLPs was developed to support retirement homes and residential care facilities that are interested in starting their own TLP to support older adults' technology competence. The foundation of the framework was constructed from enTECH's knowledge translation toolkit for implementing TLPs for older adults, and it was structurally organized by following the framework for person-centered care by Santana et al [16]. Although modifications to the framework may be required depending on

XSL•FO
RenderX

an organization's needs, starting TLPs for older adults can have potential positive impacts on them in the long term.

Acknowledgments

Sincere thanks are extended to the enTECH Computer Club Spring 2020 executive team for assisting with the organization of extracted activities from enTECH's knowledge translation toolkit under appropriate domains on the Miro whiteboard. This paper was funded by a Network for Aging Research Catalyst grant awarded to KB and KG.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Domains and their foundational principles of the technology literacy program framework on Miro whiteboard.
[PDF File (Adobe PDF File), 724 KB - aging_v5i3e34997_app1.pdf]

References

1. How to protect yourself and others. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html [accessed 2020-06-15]
2. COVID-19: guidance documents. Government of Canada. URL: https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/guidance-documents.html [accessed 2020-06-15]
3. Coronavirus disease (COVID-19). World Health Organization. URL: https://www.who.int/health-topics/coronavirus#tab=tab_2 [accessed 2020-06-15]
4. Government of Canada. COVID-19: Resources for seniors and their caregivers. URL: https://www.canada.ca/en/public-health/services/diseases/coronavirus-disease-covid-19/resources-older-adults-caregivers.html [accessed 2020-06-20]
5. Brooke J, Jackson D. Older people and COVID-19: isolation, risk and ageism. J Clin Nurs 2020 Jul 02;29(13-14):2044-2046. [doi: 10.1111/jocn.15274] [Medline: 32239784]
6. Pandemic experience in the long-term care sector: how does Canada compare with other countries? Canadian Institute for Health Information. 2020 Jun. URL: https://www.cihi.ca/sites/default/files/document/covid-19-rapid-response-long-term-care-snapshot-en.pdf [accessed 2020-08-10]
7. Davidson J, Schimmele C. Evolving internet use among canadian seniors. Statistics Canada, Analytical Studies Branch. URL: https://www150.statcan.gc.ca/n1/en/pub/11f0019m/11f0019m2019015-eng.pdf?st=1PAHc-IV [accessed 2020-07-09]
8. Haase KR, Cosco T, Kervin L, Riadi I, O'Connell ME. Older adults' experiences with using technology for socialization during the COVID-19 pandemic: cross-sectional survey study. JMIR Aging 2021 Apr 23;4(2):e28010 [FREE Full text] [doi: 10.2196/28010] [Medline: 33739929]
9. Marston HR, Genoe R, Freeman S, Kulczycki C, Musselwhite C. Older adults' perceptions of ICT: main findings from the technology In Later Life (TILL) study. Healthcare (Basel) 2019 Jul 04;7(3):86 [FREE Full text] [doi: 10.3390/healthcare7030086] [Medline: 31277387]
10. Armitage R, Nellums LB. COVID-19 and the consequences of isolating the elderly. Lancet Public Health 2020 May;5(5):e256 [FREE Full text] [doi: 10.1016/S2468-2667(20)30061-X] [Medline: 32199471]
11. Czaja SJ, Boot WR, Charness N, Rogers WA, Sharit J. Improving social support for older adults through technology: findings from the prism randomized controlled trial. Gerontologist 2017 Feb 15:467-477 [FREE Full text] [doi: 10.1093/geront/gnw249] [Medline: 28201730]
12. Rogers WA, Mynatt ED. How can technology contribute to the quality of life of older adults? The Technology of Humanity: Can Technology Contribute to the Quality of Life. 2003. URL: https://www.cc.gatech.edu/fce/ahri/publications/rogers_mynatt.pdf [accessed 2020-07-17]
13. Hoang P, Whaley C, Thompson K, Ho V, Rehman U, Boluk K, et al. Evaluation of an intergenerational and technological intervention for loneliness: protocol for a feasibility randomized controlled trial. JMIR Res Protoc 2021 Feb 17;10(2):e23767 [FREE Full text] [doi: 10.2196/23767] [Medline: 33595443]
14. EnTECH Computer Club. EnTECH. URL: http://www.entech.club/ [accessed 2020-05-25]
15. Nagallo N, Panchaud M, Wallace J. Implementing Technology Literacy Programs for Older Adults. EnTECH. 2020. URL: http://www.entech.club/knowledge-translation [accessed 2020-06-01]
16. Santana MJ, Manalili K, Jolley RJ, Zelinsky S, Quan H, Lu M. How to practice person-centred care: a conceptual framework. Health Expect 2018 Apr 19;21(2):429-440 [FREE Full text] [doi: 10.1111/hex.12640] [Medline: 29151269]
17. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260(12):1743-1748. [doi: 10.1001/jama.260.12.1743] [Medline: 3045356]
18. Kirkpatrick JD, Kirkpatrick WK. Kirkpatrick's Four Levels of Training Evaluation. United States: ATD Press; 2016.
19. RE-AIM. URL: https://www.re-aim.org/ [accessed 2021-10-28]

XSL•FO
RenderX

20. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
21. Sauro J, Lewis J. Quantifying the User Experience: Practical Statistics for User Research. Netherlands: Elsevier Science; 2012:214.
22. LoBuono DL, Leedahl SN, Maiocco E. Teaching technology to older adults: modalities used by student mentors and reasons for continued program participation. J Gerontol Nurs 2020 Jan 01;46(1):14-20. [doi: 10.3928/00989134-20191118-02] [Medline: 31895957]
23. Chan MY, Haber S, Drew LM, Park DC. Training older adults to use tablet computers: does it enhance cognitive function? Gerontologist 2016 Dec;56(3):475-484 [FREE Full text] [doi: 10.1093/geront/gnu057] [Medline: 24928557]
24. Berg-Weger M, Morley JE. Editorial: loneliness and social isolation in older adults during the covid-19 pandemic: implications for gerontological social work. J Nutr Health Aging 2020 Apr 14;24(5):456-458 [FREE Full text] [doi: 10.1007/s12603-020-1366-8] [Medline: 32346678]
25. Wu B. Social isolation and loneliness among older adults in the context of COVID-19: a global challenge. Glob Health Res Policy 2020 Jun 5;5(1):27 [FREE Full text] [doi: 10.1186/s41256-020-00154-3] [Medline: 32514427]
26. Mitzner TL, Fausset CB, Boron JB, Adams AE, Dijkstra K, Lee CC, et al. Older adults' training preferences for learning to use technology. Proc Hum Factors Ergon Soc Annu Meet 2008 Sep 01;52(26):2047-2051 [FREE Full text] [doi: 10.1177/154193120805202603] [Medline: 25309139]
27. Schuster AM, Cotten SR. influence on information and communication technologies in long-term care: results from a web-based survey with long-term care administrators. JMIR Aging 2022 Jan 12;5(1):e32442 [FREE Full text] [doi: 10.2196/32442] [Medline: 34878989]
28. Jones BD, Bayen UJ. Teaching older adults to use computers: recommendations based on cognitive aging research. Educ Gerontol 2006 Aug 03;24(7):675-689. [doi: 10.1080/0360127980240705]
29. Mitzner TL, Boron JB, Fausset CB, Adams AE, Charness N, Czaja SJ, et al. Older adults talk technology: technology usage and attitudes. Comput Human Behav 2010 Nov 1;26(6):1710-1721 [FREE Full text] [doi: 10.1016/j.chb.2010.06.020] [Medline: 20967133]
30. Kim A, Merriam SB. Motivations for learning among older adults in a learning in retirement institute. Educ Gerontol 2004 Jun;30(6):441-455. [doi: 10.1080/03601270490445069]
31. Williams KN, Kemper S. Interventions to reduce cognitive decline in aging. J Psychosoc Nurs Ment Health Serv 2010 May;48(5):42-51 [FREE Full text] [doi: 10.3928/02793695-20100331-03] [Medline: 20415290]
32. Heaggans R. The 60's are the new 20's: teaching older adults technology. SRATE 2012;21(2):1-8 [FREE Full text]
33. Gadzinski AJ, Gore JL, Ellimoottil C, Odisho AY, Watts KL. Implementing telemedicine in response to the COVID-19 pandemic. J Urol 2020 Jul;204(1):14-16 [FREE Full text] [doi: 10.1097/JU.0000000000001033] [Medline: 32249678]
34. Webster P. Virtual health care in the era of COVID-19. Lancet 2020 Apr 11;395(10231):1180-1181 [FREE Full text] [doi: 10.1016/S0140-6736(20)30818-7] [Medline: 32278374]
35. Waycott J, Zhao W, Kelly RM, Robertson E. Technology-mediated enrichment in aged care: survey and interview study. JMIR Aging 2022 Apr 12;5(2):e31162 [FREE Full text] [doi: 10.2196/31162] [Medline: 34975014]

Abbreviations

LTC: long-term care
TLP: technology literacy program

Edited by T Leung; submitted 19.11.21; peer-reviewed by S Bhattacharjya, K Colorafi; comments to author 14.03.22; revised version received 31.05.22; accepted 20.06.22; published 19.08.22.

Please cite as:
Li KS, Nagallo N, McDonald E, Whaley C, Grindrod K, Boluk K
Implementing Technology Literacy Programs in Retirement Homes and Residential Care Facilities: Conceptual Framework
JMIR Aging 2022;5(3):e34997
URL: https://aging.jmir.org/2022/3/e34997
doi:10.2196/34997
PMID:35984689

©Karen S Li, Nathan Nagallo, Erica McDonald, Colin Whaley, Kelly Grindrod, Karla Boluk. Originally published in JMIR Aging (https://aging.jmir.org), 19.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction

XSL•FO
RenderX

in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Examining Structural Disparities in US Nursing Homes: National Survey of Health Information Technology Maturity

Gregory L Alexander[1*], RN, PhD; Jianfang Liu[1*], PhD; Kimberly R Powell[2*], RN, PhD; Patricia W Stone[1*], RN, PhD

[1]School of Nursing, Columbia University, New York, NY, United States

[2]Sinclair School of Nursing, University of Missouri, Jefferson City, MO, United States

[*]all authors contributed equally

Corresponding Author:
Gregory L Alexander, RN, PhD
School of Nursing
Columbia University
Room 628
560 W 168th Street
New York, NY, 10032
United States
Phone: 1 573 301 3131
Email: ga2545@cumc.columbia.edu

Abstract

Background: There are 15,632 nursing homes (NHs) in the United States. NHs continue to receive significant policy attention due to high costs and poor outcomes of care. One strategy for improving NH care is use of health information technology (HIT). A central concept of this study is HIT maturity, which is used to identify adoption trends in HIT capabilities, use and integration within resident care, clinical support, and administrative activities. This concept is guided by the Nolan stage theory, which postulates that a system such as HIT moves through a series of measurable stages. HIT maturity is an important component of the rapidly changing NH landscape, which is being affected by policies generated to protect residents, in part because of the pandemic.

Objective: The aim of this study is to identify structural disparities in NH HIT maturity and see if it is moderated by commonly used organizational characteristics.

Methods: NHs (n=6123, >20%) were randomly recruited from each state using Nursing Home Compare data. Investigators used a validated HIT maturity survey with 9 subscales including HIT capabilities, extent of HIT use, and degree of HIT integration in resident care, clinical support, and administrative activities. Each subscale had a possible HIT maturity score of 0-100. Total HIT maturity, with a possible score of 0-900, was calculated using the 9 subscales (3 x 3 matrix). Total HIT maturity scores equate 1 of 7 HIT maturity stages (stages 0-6) for each facility. Dependent variables included HIT maturity scores. We included 5 independent variables (ie, ownership, chain status, location, number of beds, and occupancy rates). Unadjusted and adjusted cumulative odds ratios were calculated using regression models.

Results: Our sample (n=719) had a larger proportion of smaller facilities and a smaller proportion of larger facilities than the national nursing home population. Integrated clinical support technology had the lowest HIT maturity score compared to resident care HIT capabilities. The majority (n=486, 60.7%) of NHs report stage 3 or lower with limited capabilities to communicate about care delivery outside their facility. Larger NHs in metropolitan areas had higher odds of HIT maturity. The number of certified beds and NH location were significantly associated with HIT maturity stage while ownership, chain status, and occupancy rate were not.

Conclusions: NH structural disparities were recognized through differences in HIT maturity stage. Structural disparities in this sample appear most evident in HIT maturity, measuring integration of clinical support technologies for laboratory, pharmacy, and radiology services. Ongoing assessments of NH structural disparities is crucial given 1.35 million Americans receive care in these facilities annually. Leaders must be willing to promote equal opportunities across the spectrum of health care services to incentivize and enhance HIT adoption to balance structural disparities and improve resident outcomes.

(JMIR Aging 2022;5(3):e37482) doi:10.2196/37482

XSL•FO
RenderX

KEYWORDS

nursing homes; health information technology; policy; nursing informatics; electronic health record; electronic data; data sharing; care providers; resident; care; quality of care; structural disparity; clinical support; administration

Introduction

Background

There are 15,632 nursing homes (NHs) in the United States with 1.7 million beds and 1.3 million Americans residing in them [1]. For decades, care delivered in NHs has received significant policy attention due to the poor outcomes of care [2]. Recently, national experts in NH care were charged with examining how our nation delivers, regulates, finances, and measures quality [3]. The committee concluded that the current state of quality of care in NHs is "…ineffective, inefficient, fragmented, and unsustainable" [4]. NH quality has several components that interact to affect residents' health, functional status, and resident outcomes. Researchers have studied these interactions to understand how policy (ie, regulation and reimbursement), clinical interventions, management practices, and individual worker, resident, and family characteristics account for variation in NH quality [5]. A promising strategy for improving NH quality is the use of health information technology (HIT).

In this paper, we define HIT as a system that is used in health care to process, store, and exchange health information in an electronic environment. The use of HIT in NHs has also been recognized by experts in the field as a method to improve NH quality [6]. For example, NH HIT improves timely and secure exchange of electronic data and medical record access, enabling clinicians to gain direct access to clinical information, enhances efficiency, and improves resident outcomes [7,8]. Furthermore, HIT strengthens care coordination processes leading to greater consistency and better accountability among clinicians and staff [9]. Unfortunately, wide differences in NH HIT adoption exist. Researchers have found that 95% of NHs use HIT including electronic health records (EHR), and nearly half (46%) use health information exchange capabilities for resident care. However, they also discovered variations in use of technology, such as an EHR (ie, urban NHs were 2.5 times more likely to have EHR compared to rural NHs) [10]. One element missing from national NH quality reporting systems includes measures of the maturity of HIT adoption, which could help policy makers, researchers, quality improvement specialists, families and residents, and other stakeholders to identify where gaps exist.

Theoretical Approach

A central concept of this study is HIT maturity guided by stage theory by Nolan [11], which postulates that a system of coordinated processes (eg, an EHR) evolves through a series of stages as it matures. HIT maturity models, such as the Health Information Management System Society Electronic Medical Record Adoption Model, are used to assess levels of EHR maturity over time in acute care, and other maturity models are used to assess the general health care environment, mobile health, interoperability, telemedicine, and usability [12-14]. These general models of HIT maturity are not adaptable to NH contexts because NHs provide a different model of care delivery [15]. For instance, NH residents' length of stay is typically much longer than that of a patient in acute care. Furthermore, providers are largely off-site, and the nursing workforce is different in NHs compared with other settings (eg, NHs have higher staffing of licensed practical nurses, and the bulk of the workforce are certified nursing aides) [16]. Other HIT maturity models have been proposed that stress the importance of continuous cycles of reassessment that could be influenced by policy or change in solutions [17]. Previously, we have defined NH HIT maturity in 3 dimensions including HIT capabilities, use, and integration; these dimensions are further defined among the following 3 health care domains: resident care, clinical support (eg, laboratory, pharmacy, and radiology), and administrative activities. Furthermore, NH HIT maturity is categorized into 7 stages ranging from stage 0 "nonexistent HIT or EHR solutions" to stage 6 "integrated HIT systems that generate clinical data to drive self-management" [11]. HIT maturity is best measured longitudinally to enable better estimates of change in adoption over time [18]. Widespread NH HIT maturity has not been achieved due to unbalanced national policies, which have not promoted meaningful use or provided financial incentives in NHs for HIT adoption, similar to other health care sectors [19,20]. This imbalance has created wider structural disparities leading to variation in resource capabilities and use that may influence disparities in resident outcomes. For instance, following the emergence of the COVID-19 pandemic in March 2020, early analyses of telehealth use found that compared to rural NHs, urban NHs were more than 11 times more likely to use telehealth for web-based evaluations, pretransfer arrangements, second opinions, and transfer of diagnostic images following major policy changes [21,22]. Ongoing NH assessments are critical to truly understand the linkages between these types of organizational differences and impacts on quality of NH care, including disparities in resident outcomes.

HIT maturity is an important component of the rapidly changing NH landscape. In previous work, we have found alternating patterns of total HIT maturity over 3 years (2014-2017) among 815 NHs; that is, (n=579, 71%) of NHs exhibited net positive increase in total HIT maturity, (n=155, 19%) had a net negative decrease in total HIT maturity, and (n=58, 10%) had consistently negative patterns of total HIT maturity over time [18,23]. Facilities with a net increase reported increasing HIT capabilities and use as well as greater integration over time. However, NHs, reporting a net decrease in HIT maturity over time, have reduced their capabilities, use, and system integration. For instance, one of the areas that had the most variation over time in HIT maturity was the clinical support dimension associated with laboratory, pharmacy, and radiology technologies used in NHs [23]. Clinically, this makes sense, since NHs that are unaffiliated with hospitals typically do not have a laboratory, pharmacy, or radiology department in house, so they will oftentimes use disparate, stand-alone technologies to facilitate related activities for staff and residents. These disparate, isolated systems may

XSL•FO
RenderX

be easy to remove if they are not found to be efficient, are too costly, or do not meet the expectations of the users.

The development of HIT maturity surveys has allowed researchers to begin exploring the relationship between technology use and NH resident level outcomes. For instance, recent studies have revealed mixed associations between HIT maturity and antibiotic use and urinary tract infections. In one study, linking HIT maturity data with a resident-level minimum data set yielded 219,461 regular resident assessments within 90 days of survey completion on 80,237 unique, older adult long-stay residents. We found that for every 10-point increase in the HIT maturity score, the expected odds of antibiotic use increased by 7% [24]. Although this result was unexpected, NHs with higher HIT maturity may have enhanced systems enabling nurses to monitor when antibiotics are used; therefore, higher levels of awareness could eventually lead to reductions in inappropriate antibiotic use. Additionally, we examined associations between HIT maturity and urinary tract infections. Controlling for NH and resident characteristics, HIT maturity was associated with 10% less urinary tract infections [25].

The COVID-19 pandemic has also influenced changes in HIT use [22,26]. These changes include the expansion of Medicare payment for telehealth, increasing collaborations between academic medical centers and NHs, and the Centers for Medicare & Medicaid Services requiring NHs to report COVID-19 metrics to the Centers for Disease Control electronically [21,27,28]. Following national telehealth policy expansion in Spring 2020, the use of NH telehealth using computer software and web applications surged across the country [22]. However, 16% of NHs were not using telehealth, and this was more likely to occur in rural NHs [22]. Similar findings have been confirmed by other researchers exploring the explosive growth of telehealth since the pandemic started [29]. Technology adoption, especially in the face of emergent conditions, yields positive and negative outcomes that must be recognized, identified, and addressed [10]. The purpose of this study was to identify structural disparities in NH HIT maturity and how HIT maturity is moderated by various commonly used NH characteristics (eg, ownership, location, number of certified beds, chain affiliation, and occupancy rate).

Methods

Ethics Approval

This research was conducted as part of a larger ongoing 3-year national study (2019-2022) exploring trends in NH HIT maturity in the United States. Data were collected in 2019 using an NH survey that measures 3 dimensions of HIT maturity (ie, HIT capabilities, use, and integration) in the 3 domains of resident care, clinical support (ie, laboratory, pharmacy, and radiology), and administrative activities. All methods used in this research were approved by the Columbia University Institutional Review Board (PT-AABR3810).

Sample

Nursing home compare is a publicly available database, maintained by the Centers for Medicare & Medicaid Services, which provides information about every US NH serving beneficiaries of Medicare or Medicaid [1]. Nursing home compare was used to identify NHs in continental US, Alaska, and Hawaii. NHs were excluded from Guam, Puerto Rico, and Virgin Islands. NHs designated as a *special focus facility* were also excluded, because these facilities have a history of serious quality issues and are automatically included in a program to stimulate quality-of-care improvements [30]. Finally, NHs with a hospital-based designation was not included as their HIT maturity is likely different due to national incentives for HIT adoption in acute care [8,31]. After applying the exclusion criteria, the population size was 14,109 (Table 1).

The sample recruitment goal for this study was 10% of all NHs in the United States (N=1570 NH). Based on our previous experience with a 45% response rate, more than 20% (n=6123) of the facilities were randomly recruited from each state. The number of facilities selected in each state was proportional to the number of NHs located in that state. For example, because California has the largest number of homes (n=1241), 248 homes (20% of facilities) were randomly selected from all California NHs. To ensure that responses were received from multiple NHs in each state, we overrecruited in states with smaller numbers to have a minimum of 6 homes per state. Although sampling was stratified by state, facilities were not stratified further by the number of certified beds, ownership, location, and so on prior to recruitment; this is because some states may not have any NHs in some strata. Wyoming, for example, has only 38 homes. In Wyoming, there are fewer large homes in rural areas. By including every NH in the random selection process within each state, every home in each state—regardless of their characteristics—had an equal opportunity to participate.

XSL•FO
RenderX

Table 1. Comparison of national nursing home population vs sample.

Nursing home characteristics	National (N=14,109)	Sample (n=719)	Probability ratio or Cohen *d*[a]	*P* value
Location, n (%)				.07
Metro >50,000	9823 (69.68)	453 (63)	1.11	
Micro 10,000-49,999	1936 (13.73)	114 (15.86)	0.87	
Small town 2500-9999	1414 (10.03)	90 (12.52)	0.80	
Rural <2500	925 (6.56)	62 (8.62)	0.76	
Number of certified beds, n (%)				<.001
<60	2418 (17.14)	150 (20.86)	0.82	
60-120	7582 (53.74)	420 (58.41)	0.92	
>120	4109 (29.12)	149 (20.72)	1.41	
Ownership, n (%)				.89
Nonprofit	3903 (27.66)	193 (26.84)	1.03	
For-profit	10,206 (72.34)	526 (73.16)	0.99	
Chain, n (%)				.43
Yes	10,627 (75.32)	551 (76.63)	0.98	
No	3482 (24.68)	168 (23.37)	1.06	
Occupancy rate, mean (SD)	0.812 (0.2)	0.806 (0.15)	*0.03*	.46

[a]Calculated as the probabilities of national data divided by the sample data for categorical variable or Cohen *d* (italicized) for continuous variable.

Measures

Dependent Variables

Our psychometrically sound NH HIT survey has 9 subscales, from which 7 HIT maturity stages are derived [32]. The composite score has good internal consistency (Cronbach α=.86) [33]. HIT capabilities are scored 0 if the technology is "Not Available" or 1 if HIT is "Available," as indicated by the respondents' feedback. If an NH respondent indicates the availability of HIT capability, the respondent rates the extent of use on a scale of 1 (barely used) to 7 (extensively used). To determine the degree of integration, the respondents indicate the degree of electronic transfer of information among NH systems on a scale of 0 (not at all) to 6 (very much). The maximum range in scores for each HIT maturity dimension and domain is 0 to 100. Total HIT maturity score for all dimensions and domains combined ranges between 0 and 900. The 7 HIT maturity stages are correlated with total HIT maturity [34]. Stage 0 is the lowest stage of HIT maturity. Stage 0 indicates that HIT solutions or EHRs are nonexistent in the NH. Stage 6, the highest stage of HIT maturity, indicates the use of data by residents or residents' representatives to generate clinical data and drive self-management.

Independent Variables

Five NH characteristics were included. Ownership type was collapsed into the 2 categories of "For Profit" and "Nonprofit" (nonprofit included NHs with a government classification in nursing home compare). A binary chain status variable was created. Rural-Urban Commuting Area Codes were used to classify the NHs by ZIP Codes into 4 regional locations including the following: metropolitan >50,000; micropolitan 10,000-49,999; small town 2500-9999; and rural <2500 [35]. NHs were classified into 3 classes based on the number of certified beds, including small (<60 beds), medium (60-120 beds), and large (>120 beds), which are common classifications in other NH research projects. Occupancy rate was calculated as the number of residents divided by the numbers of certified beds in the facility.

Analysis

Probability ratios were computed to compare NH characteristics between the national data and the study sample [36]. Subsequently, weights were computed for each state. Internal consistence measured by Cronbach α and descriptive statistics of the raw and weighted total HIT maturity score as well as the 9 subscales were computed, followed by a table with stage and proportion per stage. The relationship between the NH characteristics and HIT stage was assessed. Since HIT stage is measured as a 7-point Likert scale, ordinal logistic regression was used. Unadjusted cumulative odds ratios (ORs) and associated 95% confidence intervals were computed with each independent variable entered into the regression model separately, and the adjusted cumulative ORs were computed from multivariable ordinal logistic regression where all independent variables were entered into the model. Using the continuous total HIT score as the outcome measure, multiple linear regression was performed as a sensitivity analysis to verify the significant NH factors on the HIT measure. All statistical analyses were performed using SAS version 9.3 (SAS Institute), and PROC SURVEYLOGISTIC, using variance estimators that are appropriate for survey sampled data, was used in the ordinal logistic regression analysis.

XSL•FO
RenderX

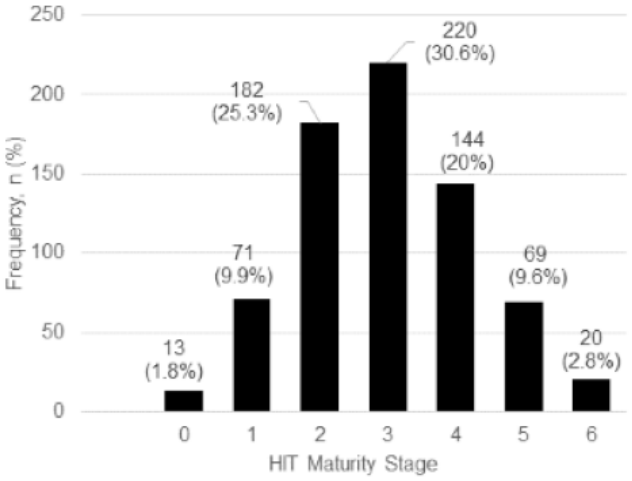
Results

A total of 719 homes completed the survey with all 50 states and the District of Columbia being represented. The comparison of the national population of eligible NHs and the study sample is provided in Table 1, and the differences were very small. However, the sample had a larger proportion of facilities with less than 60 certified beds and a smaller proportion of larger facilities with more than 120 certified beds.

The aggregated raw HIT maturity scores are shown as Table S1 in Multimedia Appendix 1 for the 719 NHs. Median (50th percentile) scores among all 9 subscales for the 719 NHs ranged from a low of 22.22 in integration of clinical support technologies to a high of 68.97 in resident care HIT capabilities, both with a minimum and maximum score possible of 0 to 100. In 7 (78%) out of 9 subscales, at least 1 NH indicated the lowest possible score of 0, the lowest HIT maturity score indicating that HIT systems were nonexistent. Total HIT maturity scores reflect the aggregated score of all 9 subscales. At least one facility reported a total HIT maturity score of 58.32 (minimum=0), while the maximum reported by facility was 869.74 (maximum=900). The median score (440.38) closely approximated the mean score (447.2; SD 158.4), indicating a highly symmetric distribution of total HIT maturity scores. The internal consistency of the HIT maturity scores were validated (Cronbach α>.80).

The weighted HIT maturity scores accounting for number of responses by state are illustrated in Table S2 in Multimedia Appendix 2. Figure 1 includes a description of HIT maturity stages and associated total HIT maturity stages in this sample. In this sample, 1.81% (13/719) of the NHs were at a stage 0, meaning that those facilities had nonexistent HIT solutions or EHRs at the time of the survey. A total of 71/719 (9.87%) facilities were at a stage 1. NHs at stage 1 have disparate or fragmented HIT systems that typically have distinct functionalities, which are not well integrated into care delivery activities. Just over 25% (n=182) of NHs have formal, established HIT leadership involved in governing and coordinating various aspects of HIT systems, putting them at stage 2. The majority of NHs surveyed (220/719, 30.6%) report achieving an HIT maturity of stage 3.

Figure 1. Nursing Home HIT Maturity Stages and Definitions. EHR: electronic health record; HIT: health information technology.



Stage	HIT Maturity Score	Stage Definition
0	0-128	Nonexistent HIT solutions or EHR.
1	129-257	Incomplete or disparate fragmented HIT solutions.
2	258-385	Established HIT leadership that governs and coordinates structures, procedures, processes, and policies.
3	386-514	Automated internal connectivity and reporting.
4	515-643	Automated external connectivity and reporting.
5	643-771	Clinical risk intervention and predictive analytics Use of data by resident or resident.
6	772-900	Representative to generate clinical data and drive self-management.

NHs at stage 3 have internal connectivity and reporting capabilities, meaning that these staff have limited capacity to communicate about care delivery with people outside their facility, such as with staff from external clinics, laboratories, or pharmacies. To be able to communicate electronically with people outside their facilities, NHs must reach a stage 4 or higher, and 32.4% (n=233) reached a stage 4 or higher. Nearly 3% (20/719 NHs) have achieved stage 6, the highest possible stage. In these facilities, data use by residents or residents' representatives are available to generate clinical data and to drive self-management activities.

The results of the ordinal logistic regression models (Tables 2 and 3) demonstrated that the number of certified beds and location were significantly associated with HIT maturity stage, while ownership, chain status, and occupancy rate were not. The results from simple (Table 2) and multivariable (Table 3) ordinal logistic regression models are similar. From the final multivariable models (Table 3), for the number of certified beds

XSL•FO
RenderX

(<60 vs >120), the odds of being at a higher stage increased by 55% for larger homes (cumulative OR=0.45; *P*=.001), when all other variables in the model were held constant. For location, the odds of being at a higher stage decreased by 55% (cumulative OR=0.45; *P*<.001) for rural homes versus metropolitan homes; the odds of being at a higher stage decreased by 43% (cumulative OR=0.57; *P*=.048) for small town homes versus metropolitan homes. The results from sensitivity analysis, using the total HIT maturity score as outcome, demonstrated similar associations between NH characteristics and HIT maturity measure (Table S3 in Multimedia Appendix 3).

Table 2. Simple ordinal logistic regression model assessing the relationship between nursing home characteristics and health information technology maturity stage (n=719).

Nursing home characteristics	Unadjusted cumulative odds ratio	95% CI	*P* value
Bed size (ref: >120)			
60-12	0.93	0.64 (1.35)	.71
<60	0.39	0.25 (0.63)	<.001[a]
Location (ref: metro)			
Micro	0.66	0.42 (1.02)	.06
Rural	0.36	0.23 (0.58)	<.001[a]
Small town	0.53	0.31 (0.91)	.02[a]
For-profit			
Nonprofit vs for-profit	0.80	0.58 (1.10)	.16
Chain			
Chain vs nonchain	1.09	0.77 (1.56)	.63
Occupancy rate	2.21	0.83 (5.90)	.11

[a]*P* value significant at .05 level.

Table 3. Multivariable ordinal logistic regression model assessing the relationship between nursing home characteristics and HIT maturity stage (n=719).

Nursing home characteristics	Adjusted cumulative odds ratio	95% CI	*P* value
Bed size (ref: >120)			
60-12	0.99	0.67 (1.46)	.96
<60	0.45	0.28 (0.73)	.001[a]
Location (ref: metro)			
Micro	0.65	0.41 (1.03)	.07
Rural	0.45	0.29 (0.71)	.001[a]
Small town	0.57	0.33 (1.00)	.045[a]
For-profit			
Nonprofit vs for-profit	0.84	0.61 (1.17)	.30
Chain			
Chain vs nonchain	1.12	0.78 (1.61)	.54
Occupancy rate	1.71	0.63 (4.66)	.30

[a]*P* value significant at .05 level.

Discussion

Principal Findings

The results from this study indicate that structural disparities in HIT maturity exist. Most facilities, nearly 68% (n=486) report being at stage 3 or lower of HIT maturity indicating they are not able to electronically communicate externally with other facilities. This lack of connectivity can result in reduced levels of electronic data sharing, leading to deficiencies in care delivery, substandard care coordination activities, and poorer resident outcomes [37]. Structural disparities in this sample appear to be most evident in HIT maturity domains and dimensions, measuring integration of clinical support technologies used for laboratory, pharmacy, and radiology services. Clinically, this makes sense, since many NHs without

XSL•FO
RenderX

a hospital designation do not have these services available on site. However, that should not deter leaders from adopting systems that support higher integration and data sharing opportunities. NH leaders are challenged by a lack of financial and other incentives to adopt HIT systems that support these clinical activities and optimize care delivery [38]. Other challenges affecting outcomes of adoption supported in the literature relate to human factors design and usability issues such as excessive data entry, information overload, and slow system response times [39]. In addition to these influences, our study revealed structural disparities influenced by organizational characteristics, including NH size and location, which place some residents—usually the most vulnerable—at a disadvantage for receiving optimal care [40].

Ongoing assessments and characterization of structural disparities in NHs is crucial given 1.35 million Americans receive care in these facilities annually [41]. Without determining where disparities exist and what factors influence them, it is difficult for policy-setting organizations that oversee NH quality and care delivery to act effectively. This includes recommending and implementing strategies to reduce differences in care delivery across settings that can have a positive effect on resident outcomes. Nevertheless, leaders who focus on health care policy and disparities must be willing to promote equal opportunities across the spectrum of health care services to incentivize and enhance HIT adoption in all settings to balance these types of structural disparities to maximize resident outcomes. Otherwise, facilities such as NHs, which historically have not had the same support for promoting HIT infrastructure as other health care facilities, will certainly experience wider structural disparities and likely poorer resident outcomes.

Clearly, when incentives are provided or barriers are removed from HIT adoption, facilities will respond in ways that reduce structural disparities and promote better care delivery. To some extent, current incentive programs through meaningful use appear to be influencing HIT adoption in NHs and information-sharing practices with other clinical settings such as hospitals [8]. However, decisions to integrate electronic data sources is also dependent on organizational characteristics. For example, contrary to our findings, Burke et al [31] found that there were lower odds (OR=0.11; *P*=.04) of formal data integration between NHs and hospitals if an NH were for-profit versus not-for-profit. In a related work, Adler-Milstein et al [8] reported that higher odds (OR=1.96; *P*=.008) of sharing more complete resident information occur between Hospitals and NHs in metropolitan versus rural locations. Disparities enhanced by the size or location of a facility are likely related to resources including knowledgeable staff available to support technology implementation throughout its lifecycle. Policy makers have begun to address these deficits. For example, the Office of the National Coordinator provided funding to develop a toolkit called the Usability Change Package to support organizations that did not have ready access to usability experts and resources for EHR adoption and maintenance [42], a frequent occurrence in NHs in the United States. It is not clear how well the uptake has been or how effective these tools are for the NH industry.

Limitations

Our survey uses broad constructs to describe structural disparities in this sample of NHs. However, we have used rigorous methods to be sure our measures have been informed by highly experienced and qualified members of the NH community [32]. One limitation, however, may be a response bias for NHs choosing not to participate. Some NH administrators may not participate because they have no technology and do not perceive relevance, which could result in an overall higher level of HIT maturity. Some NHs may not join because administrators do not have the knowledge to complete the survey. We offered help to overcome barriers by providing our contact information and answering questions as administrators participated. Our team's increased availability and responsiveness may have reduced respondent burden, which in turn may have increased participation. Although we found some areas indicating significant differences in HIT maturity and stage when comparing some commonly used organizational characteristics, we cannot assume that lack of significance means that structural deficiencies are not present.

Conclusion

In this national sample, we identified important structural disparities in NHs that are likely impacting the quality of care their residents are receiving. The majority of these NHs have lower HIT maturity levels, reporting a gap in connectivity with external facilities that might otherwise enhance health data sharing across different organizations. These differences could be due to inadequate infrastructures, availability of a knowledgeable workforce, or financial resources to promote higher levels of adoption. It is crucial that we begin to consistently identify a means to address these disparities, first by increasing transparency and public reporting about the trends in NH HIT maturity in the United States, followed by implementing national policies to level these deficits.

Practice Implications

Increasingly, at the forefront of policies affecting NH care delivery is the awareness that structural disparities can place undue burden on practicing NH leaders and staff to provide high-quality care to residents. However, underneath this problem is a lack of structured and standardized means to identify and report existing structural disparities in NHs in the United States. In the absence of systematic reporting mechanisms to identify existing structural disparities in NHs, these issues will go undetected, and leaders, staff, and residents will continue to suffer the consequences.

Acknowledgments

This project was supported by grant number R01HS022497 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare

XSL•FO
RenderX

Research and Quality. The authors wish to thank our Project Coordinator (Keely Wise) and Recruitment Team (Brooke Schrimpf, Joy Bennett, Chris Duren, and Laurie Janeczko) who have worked diligently toward our goal of improving care of older adults in nursing homes.

Data Availability

The data sets generated and analyzed during this study are not publicly available due to the project not being completed but are available from the corresponding author upon reasonable request following completion of the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Table S1.
[DOCX File , 14 KB - aging_v5i3e37482_app1.docx]

Multimedia Appendix 2
Table S2.
[DOCX File , 18 KB - aging_v5i3e37482_app2.docx]

Multimedia Appendix 3
Table S3.
[DOCX File , 14 KB - aging_v5i3e37482_app3.docx]

References

1. Medicare and Medicaid (CMS) Care Compare. Medicare Provider Data. URL: https://data.cms.gov/provider-data/ [accessed 2022-08-09]
2. Reinhard SR, Stone R. Promoting quality in nursing homes: The Wellspring model. The Commonwealth Fund. 2001. URL: https://www.commonwealthfund.org/publications/fund-reports/2001/jan/promoting-quality-nursing-homes-wellspring-model [accessed 2022-08-09]
3. National Academy of Science Engineering and Medicine. The national imperative to improve nursing home quality: Honoring our commitment to residents, families, and staff. Washington DC, US: The National Academies Press; 2022.
4. National Academy of Science Engineering and Medicine. Summary and Introduction. In: The National Imperative to Improve Nursing Home Quality: Honoring Our Commitment to Residents, Families, and Staff. Washington DC, US: National Academies Press; 2022:1-46.
5. Ingber MJ, Feng Z, Khatutsky G, Wang JM, Bercaw LE, Zheng NT, et al. Initiative to reduce avoidable hospitalizations among nursing facility residents shows promising results. Health Aff (Millwood) 2017 Mar 01;36(3):441-450. [doi: 10.1377/hlthaff.2016.1310] [Medline: 28264945]
6. National Academy of Science Engineering and Medicine. Health information technology. In: The National Imperative to Improve Nursing Home Quality: Honoring Our Commitment to Residents, Families, and Staff. Washington DC, US: The National Academies Press; 2022:453-494.
7. Wong SP, Jacobson HN, Massengill J, White HK, Yanamadala M. Safe interorganizational health information exchange during the COVID-19 pandemic. J Am Med Dir Assoc 2020 Dec;21(12):1808-1810 [FREE Full text] [doi: 10.1016/j.jamda.2020.10.022] [Medline: 33162358]
8. Adler-Milstein J, Raphael K, O'Malley TA, Cross DA. Information sharing practices between US hospitals and skilled nursing facilities to support care transitions. JAMA Netw Open 2021 Jan 04;4(1):e2033980 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.33980] [Medline: 33443582]
9. Gill E, Dykes PC, Rudin RS, Storm M, McGrath K, Bates DW. Technology-facilitated care coordination in rural areas: What is needed? Int J Med Inform 2020 May;137:104102 [FREE Full text] [doi: 10.1016/j.ijmedinf.2020.104102] [Medline: 32179256]
10. Powell KR, Deroche CB, Alexander GL. Health data sharing in US nursing homes: a mixed methods study. J Am Med Dir Assoc 2021 May;22(5):1052-1059 [FREE Full text] [doi: 10.1016/j.jamda.2020.02.009] [Medline: 32224261]
11. Nolan RL. Managing the computer resource. Commun. ACM 1973 Jul;16(7):399-405. [doi: 10.1145/362280.362284]
12. Analytics: Electronic medical record adoption model. Health Information Management Systems Society. 2017. URL: http://www.himssanalytics.org [accessed 2019-09-16]
13. Electronic healthcare maturity model (eHMM). Quintegra. URL: http://www.quintegrasolutions.com/eHMM%20White%20Paper.pdf [accessed 2022-08-08]

XSL•FO
RenderX

14. Carvalho JV, Rocha Á, Abreu A. Maturity models of healthcare information systems and technologies: a literature review. J Med Syst 2016 Jun 15;40(6):131. [doi: 10.1007/s10916-016-0486-5] [Medline: 27083575]
15. Dykes PC, Samal L, Donahue M, Greenberg JO, Hurley AC, Hasan O, et al. A patient-centered longitudinal care plan: vision versus reality. J Am Med Inform Assoc 2014 Nov 01;21(6):1082-1090 [FREE Full text] [doi: 10.1136/amiajnl-2013-002454] [Medline: 24996874]
16. National Academy of Science Engineering and Medicine. The nursing home workforce. In: The National Imperative to Improve Nursing Home Quality: Honoring Our Commitment to Residents, Families, and Staff. Washington DC, US: The National Academies; 2022:221-302.
17. Cresswell K, Sheikh A, Krasuska M, Heeney C, Franklin BD, Lane W, et al. Reconceptualising the digital maturity of health systems. The Lancet Digital Health 2019 Sep;1(5):e200-e201. [doi: 10.1016/s2589-7500(19)30083-4]
18. Alexander GL, Madsen R, Deroche CB, Alexander R, Miller E. Ternary trends in nursing home information technology and quality measures in the United States. J Appl Gerontol 2020 Oct 17;39(10):1134-1143 [FREE Full text] [doi: 10.1177/0733464819862928] [Medline: 31311420]
19. Kruse CS, Mileski M, Alaytsev V, Carol E, Williams A. Adoption factors associated with electronic health record among long-term care facilities: a systematic review. BMJ Open 2015 Jan 28;5(1):e006615-e006615 [FREE Full text] [doi: 10.1136/bmjopen-2014-006615] [Medline: 25631311]
20. Electronic health records incentives and certification. HealthIT.gov. URL: https://www.healthit.gov/providers-professionals/meaningful-use-definition-objectives [accessed 2022-08-09]
21. CMS releases telehealth toolkit for general practitioners and end-stage renal disease (ESRD) providers. Medicaid.gov. 2020. URL: https://www.cms.gov/files/document/general-telemedicine-toolkit.pdf [accessed 2020-04-08]
22. Alexander G, Powell K, Deroche C. An evaluation of telehealth expansion in U.S. nursing homes. J Am Med Inform Assoc 2021 Feb 15;28(2):342-348 [FREE Full text] [doi: 10.1093/jamia/ocaa253] [Medline: 33164054]
23. Alexander GL, Madsen R, Newton M. Analyzing change in nursing home information technology sophistication: a 2-year survey. J Gerontol Nurs 2017 Jan 01;43(1):17-21 [FREE Full text] [doi: 10.3928/00989134-20161215-05] [Medline: 28091687]
24. Cohen CC, Powell K, Dick AW, Deroche CB, Agarwal M, Stone PW, et al. Examining nursing home information technology maturity and antibiotic use among long-term care residents. J Am Med Dir Assoc 2022 Jun;23(6):1019-1024.e2. [doi: 10.1016/j.jamda.2022.01.052] [Medline: 35172166]
25. Cohen C, Powell K, Dick A, Stone P, Alexander G. The association between nursing home information technology maturity and urinary tract infection among long-term residents. J Appl Gerontol 2022 Jul;41(7):1695-1701 [FREE Full text] [doi: 10.1177/07334648221082024] [Medline: 35439093]
26. Powell KR, Alexander GL. Consequences of rapid telehealth expansion in nursing homes: promise and pitfalls. Appl Clin Inform 2021 Aug;12(4):933-943. [doi: 10.1055/s-0041-1735974] [Medline: 34614517]
27. Nursing home Covid-19 data dashboard. National Healthcare Safety Network. 2021. URL: https://www.cdc.gov/nhsn/covid19/ltc-report-overview.html [accessed 2022-08-09]
28. Archbald-Pannone LR, Harris DA, Albero K, Steele RL, Pannone AF, Mutter JB. COVID-19 collaborative model for an academic hospital and long-term care facilities. J Am Med Dir Assoc 2020 Jul;21(7):939-942 [FREE Full text] [doi: 10.1016/j.jamda.2020.05.044] [Medline: 32563752]
29. Mehrotra A. Looking ahead: what should the telemedicine regulatorypayment landscape look like after the pandemic? Committee on WaysMeans, Subcommittee on Health. Ways and Means Committee Democrats. 2021. URL: https://waysandmeans.house.gov/sites/democrats.waysandmeans.house.gov/files/documents/Ateev%20Mehrotra%20Testimony.pdf [accessed 2022-08-09]
30. Special Focus Facility Program. Centers for Medicare and Medicaid Nursing Homes. URL: https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/Downloads/SFFList.pdf [accessed 2022-08-10]
31. Burke RE, Phelan J, Cross D, Werner RM, Adler-Milstein J. Integration activities between hospitals and skilled nursing facilities: a national survey. J Am Med Dir Assoc 2021 Dec;22(12):2565-2570.e4. [doi: 10.1016/j.jamda.2021.05.005] [Medline: 34062148]
32. Alexander G, Powell K, Deroche C, Popejoy L, Mosa ASM, Koopman R, et al. Building consensus toward a national nursing home information technology maturity model. J Am Med Inform Assoc 2019 Jun 01;26(6):495-505 [FREE Full text] [doi: 10.1093/jamia/ocz006] [Medline: 30889245]
33. Alexander GL, Deroche CB, Powell KR, Mosa ASM, Popejoy L, Koopman RJ, et al. Development and pilot analysis of the nursing home health information technology maturity survey and staging model. Res Gerontol Nurs 2022 Mar;15(2):93-99. [doi: 10.3928/19404921-20220218-04] [Medline: 35312439]
34. Alexander GL, Deroche C, Powell K, Mosa ASM, Popejoy L, Koopman R. Forecasting content and stage in a nursing home information technology maturity instrument using a Delphi method. J Med Syst 2020 Feb 05;44(3):60. [doi: 10.1007/s10916-020-1528-6] [Medline: 32020390]
35. Cromartie J. Rural Urban Commuting Area Codes. Food and Agriculture Organization of the United Nations. URL: https://agris.fao.org/agris-search/search.do?recordID=US2019X00125 [accessed 2019-10-24]

XSL•FO
RenderX

36. Chen H, Cohen P, Chen S. How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. Communications in Statistics - Simulation and Computation 2010 Mar 31;39(4):860-864. [doi: 10.1080/03610911003650383]
37. Alexander GL, Madsen RW. A report of information technology and health deficiencies in U.S. nursing homes. J Patient Saf 2017 May 31;17(6):e483-e489. [doi: 10.1097/pts.0000000000000390]
38. Ko M, Wagner L, Spetz J. Nursing home implementation of health information technology: review of the literature finds inadequate investment in preparation, infrastructure, and training. Inquiry 2018 Jun 11;55:46958018778902 [FREE Full text] [doi: 10.1177/0046958018778902] [Medline: 29888677]
39. Kroth PJ, Morioka-Douglas N, Veres S, Babbott S, Poplau S, Qeadan F, et al. Association of electronic health record design and use factors with clinician stress and burnout. JAMA Netw Open 2019 Aug 02;2(8):e199609 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.9609] [Medline: 31418810]
40. Sloane PD, Yearby R, Konetzka RT, Li Y, Espinoza R, Zimmerman S. Addressing systemic racism in nursing homes: a time for action. J Am Med Dir Assoc 2021 Apr;22(4):886-892. [doi: 10.1016/j.jamda.2021.02.023] [Medline: 33775548]
41. Harris-Kotjetin L, Sengupta M, Lendon J, Rome V, Valverde R, Caffrey C. Long-term care providers and services users in the United States, 2015-2016. National Centre for Health Statistics. 2015. URL: https://www.cdc.gov/nchs/data/series/sr_03/sr03_43-508.pdf [accessed 2022-08-09]
42. ONC Change Package for Improving EHR Usability. HealthIT.gov. URL: https://www.healthit.gov/sites/default/files/playbook/pdf/usability-change-plan.pdf [accessed 2022-08-09]

Abbreviations

EHR: electronic health records
HIT: health information technology
NH: nursing home
OR: odds ratio

Edited by T Leung, J Wang; submitted 22.02.22; peer-reviewed by G Davlyatov, A Landmark; comments to author 26.04.22; revised version received 07.06.22; accepted 14.07.22; published 23.08.22.

Please cite as:
Alexander GL, Liu J, Powell KR, Stone PW
Examining Structural Disparities in US Nursing Homes: National Survey of Health Information Technology Maturity
JMIR Aging 2022;5(3):e37482
URL: https://aging.jmir.org/2022/3/e37482
doi:10.2196/37482
PMID:35998030

©Gregory L Alexander, Jianfang Liu, Kimberly R Powell, Patricia W Stone. Originally published in JMIR Aging (https://aging.jmir.org), 23.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effects of a Digital Well-being Intervention on Older Adults: Retrospective Analysis of Real-world User Data

Eliane Boucher[1], PhD; Ryan Honomichl[1], PhD; Haley Ward[1], BA; Tyler Powell[1], MSc; Sarah Elizabeth Stoeckl[1], BS; Acacia Parks[1], PhD

Twill, New York, NY, United States

Corresponding Author:
Eliane Boucher, PhD
Twill
114 Fifth Avenue
10th floor
New York, NY, 10011
United States
Phone: 1 432 258 5233
Email: eliane@twill.health

Abstract

Background: Digital interventions have been shown to be effective for a variety of mental health disorders and problems. However, few studies have examined the effects of digital interventions in older adults; therefore, little is known about how older adults engage with or benefit from these interventions. Given that adoption rates for technology among people aged ≥65 years remain substantially lower than in the general population and that approximately 20% of older adults are affected by mental health disorders, research exploring whether older adults will use and benefit from digital interventions is needed.

Objective: This study aimed to examine the extent to which older adults engaged with a digital well-being intervention (Happify) and whether engaging with this program led to improvements in both subjective well-being and anxiety symptoms.

Methods: In this retrospective analysis, we analyzed data from 375 real-world Happify users aged ≥65 years who signed up for the platform between January 1, 2019, and December 23, 2021. Changes in well-being and anxiety symptoms across 42 to 182 days were assessed using responses to the in-app assessment, which users were prompted to take every 2 weeks, and were compared among users who engaged with the program at the recommended level (ie, 2 or more activities per week) or below the recommended level.

Results: In all, 30% (113/375) of the sample engaged with the platform at the recommended level (ie, completed an average of 2 or more activities per week), and overall, users completed an average of 43.35 (SD 87.80) activities, ranging from 1 to 786, between their first and last assessment. Users were also active on the platform for an average of 19.36 (SD 27.16) days, ranging from 1 to 152 days. Moreover, older adults who engaged at the recommended level experienced significantly greater improvements in subjective well-being (P=.002) and anxiety symptoms (P<.001) relative to those who completed fewer activities.

Conclusions: These data provide preliminary evidence that older adults engage with and benefit from digital well-being interventions. We believe that these findings highlight the importance of considering older adult populations in digital health research. More research is needed to understand potential barriers to using digital interventions among older adults and whether digital interventions should be modified to account for this population's particular needs (eg, ensuring that the intervention is accessible using a variety of devices). However, these results are an important step in demonstrating the feasibility of such interventions in a population that is assumed to be less inclined toward digital approaches.

(JMIR Aging 2022;5(3):e39851) doi:10.2196/39851

KEYWORDS

mobile apps; mental health; older adults; technology adoption; digital health; mobile phone

XSL•FO
RenderX

Introduction

Background

According to the United Nations, the proportion of the global population >65 years is increasing and is expected to reach 1 in 6 people by 2050, compared with 1 in 11 in 2019 [1]. In the United States, the US Census Bureau estimated that by 2030, the population of older adults will exceed that of children for the first time, with 1 in 5 Americans being of retirement age [2]. Although the aging population has spurred discussions surrounding the added burden of chronic illness and complex medical conditions in this cohort [3-6], comparatively less attention has been dedicated to understanding the implications of poor mental health in this "silver tsunami."

The World Health Organization reports that 1 in 5 adults aged ≥60 years is affected by a mental or neurological disorder, excluding headaches, and these disorders account for 17.4% of the years lived with disability [7]. The most common mental health disorders in this age group are depression and anxiety, affecting approximately 7% and 3.8% of older adults worldwide, respectively [7]. In the United States and Australia, the prevalence of depressive symptoms alone among older adults is estimated to be 9.8% [8]. Poor mental health in older adults subsequently contributes to elevated health care costs [9], directly and indirectly by worsening comorbid health conditions [10]. Moreover, research suggests that approximately 1 in 4 older adults with mental health disorders report experiencing discrimination, including racism and ageism, compared with 1 in 10 older adults without mental health disorders. Among those with mental health disorders, this discrimination was also more likely to occur within health care settings, thus increasing the likelihood that the individual will delay treatment or not seek treatment at all [11].

In fact, older adults are less likely than middle-aged adults to seek mental health care [12]. In a study, 6.5% of older adults self-reported some level of mental health care in the previous 12 months, but 65.9% of respondents with clinical levels of depression and 72.5% of those with anxiety never received treatment [13]. Research suggests that the lack of treatment seeking may not be related to perceptions of mental health care but to access. Specifically, a study found that older adults were more likely to indicate that access to mental health care was important but less likely to indicate that they had access to such care [14]. Although lower rates of treatment seeking may be because older adults are less likely to report having insurance coverage for mental health services compared with younger adults [14], it is further compounded by the shortage of mental health professionals specializing in geriatric populations [15].

Given these difficulties in accessing mental health care, the need to explore scalable and affordable options for mental health care is imperative as a growing proportion of the population enters old age and requires more services. Over the past 2 decades, a number of digital interventions have been developed to address the general unmet need for mental health care [16], and research suggests that these can effectively help improve mental health, including depression, anxiety, and stress [17-20]. However, research on whether older adults will engage with or benefit from digital interventions is limited.

We should be careful not to assume that the evidence suggesting that digital interventions are usable and effective within the general population applies to older adults. Although an increasing number of older adults reports owning a smartphone and using the internet, the proportion of older adults owning smartphones or having access to broadband services at home is still lagging behind that of younger age groups. For example, a national survey of adults in the United States conducted by the Pew Research Center in 2021 showed that 85% of respondents indicated owning a smartphone, whereas among older adults, only 61% reported owning smartphones. Among those aged ≥75 years, only 43% owned smartphones [21]. Age-related issues with manual dexterity and vision as well as a lack of confidence in using new technologies may contribute to the slower adoption of digital interventions and tools among older adults [22].

However, the few studies conducted with older adults suggest that those who engage with digital interventions show improvements in mental health outcomes. A meta-analysis of 9 studies exploring the effects of internet-based cognitive behavioral therapy (CBT) in older adults (mean age 66 years) found that these programs were generally effective at reducing depressive symptoms, although there was some evidence that their effectiveness was negatively related to the user's age [23]. However, limited research on the impact of digital interventions on loneliness has shown no significant improvement in loneliness among older adults [24]. Other research suggests that engagement with digital interventions may improve with age [25-27]. One study of patients prescribed internet-based CBT in Australia found that patients aged ≥60 years were more likely to complete all treatment modules than younger patients, and improvements in psychological distress and disability were consistent across age groups [25]. However, researchers have argued that these studies provide little information about the uptake and engagement of digital interventions among older adults outside of controlled research conditions [28].

Objectives

The purpose of this study was to explore whether older adults engaged with a publicly available digital intervention, the *Happify* wellness program, and the extent to which engaging with this intervention led to improvements in mental health over time. Happify is a self-guided wellness program that aims to improve psychological well-being by delivering brief gamified activities adapted from evidence-based activities from various therapeutic approaches. Previous research has shown that completing 2 or more activities via Happify per week led to significant improvements in subjective well-being and anxiety symptoms after 6 weeks [29-31], but none of these studies examined the effects of age or focused specifically on older adults. Therefore, in the current retrospective analysis, we analyzed data from real-world Happify users who self-reported being aged ≥65 years to determine whether completion of intervention activities was related to changes in subjective well-being and anxiety symptoms after at least six weeks of use.

XSL•FO
RenderX

Methods

Study Design

This study was a retrospective analysis of real-world Happify users who signed up for the program between January 1, 2019, and December 23, 2021.

Sample Selection

When signing up, all users were prompted to complete an onboarding questionnaire after downloading the Happify application or accessing the website. This questionnaire inquired about demographic information, including age group, gender, relationship status, and employment status, intended to help tailor the program for individual users. Upon completing this questionnaire, users must agree to the terms of the service and privacy policy before creating their account, which includes language indicating that their data may be used for research purposes. All data presented here were generated by real-world users as part of the standard user experience and stored on secure company servers, and only deidentified data were extracted for analysis.

Data from all users located in the United States who selected "65 or older" as their age category when responding to the onboarding questionnaire and who completed at least two in-app assessments were initially considered. To be included in the analysis, users also had to meet the following criteria: (1) complete at least two in-app assessments within 182 days, (2) the time between their first and last assessments was no less than 42 days, and (3) complete at least one Happify activity between their first and last assessments.

Ethics Approval

The use of Happify consumer data for retrospective analyses, such as this one, where data represent only those of users who naturally sign up for Happify and engage with the generic version of the program (ie, where no content or assessments have been changed for the purposes of research), was reviewed by IntegReview, an independent institutional review board, and labeled as exempt research (HLS-018).

Materials and Procedure

Digital Well-being Intervention

Happify is a self-guided wellness program that draws on various theoretical approaches to improve well-being including CBT [32], mindfulness-based stress reduction [33], positive psychology [34], acceptance and commitment therapy [35], and behavioral activation [36]. Activities based on these therapeutic approaches were developed by identifying activities within each evidence-based approach (ie, demonstrated effectiveness in at least two different studies and with different samples) [37]. These activities are then organized into "tracks," which are meant to help users address a specific area of concern, such as coping with stress or improving sleep (Figure 1). Each track consists of 4 parts, and users progress through the track by completing a percentage of the activities within each part (Figure 2). Users can change tracks at any time, and they can also access activities outside tracks via the instant play feature. A more detailed description of the Happify program is available in a previous research [31].

XSL•FO
RenderX

Figure 1. Screenshot of a Happify track on the smartphone app version.

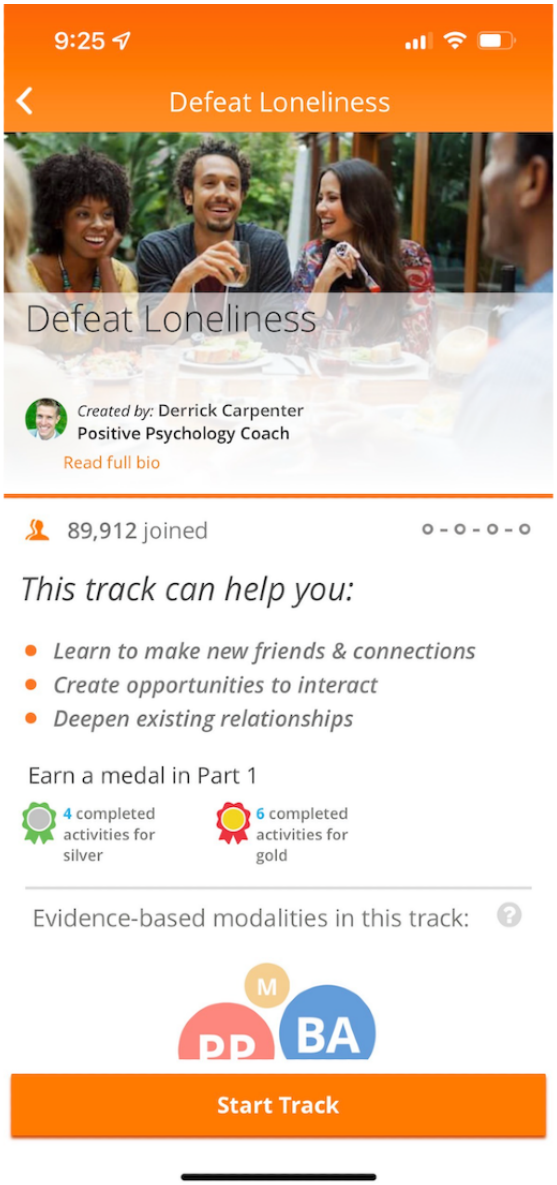
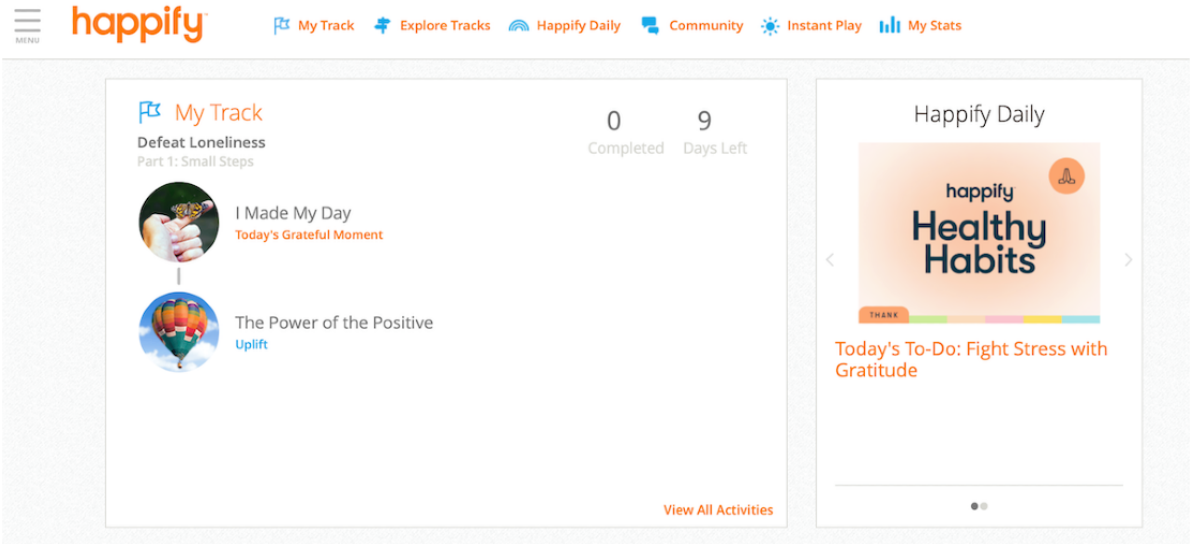


Figure 2. Screenshot of Happify track part on the web-based version.



XSL•FO
RenderX

Assessment

As part of the regular Happify program, users are prompted to complete an in-app assessment the second time they log into the platform and every 2 weeks thereafter. This assessment consists of two measures: a proprietary measure of subjective well-being, the Happify Scale, and the Generalized Anxiety Disorder 2 (GAD-2) scale [38].

The in-app assessment is optional; therefore, users may choose to skip or delay the assessment when prompted. Owing to the optional nature of the assessment, many users do not provide this outcome data. Among those who completed the assessments, the time between assessments and the number of assessments also varied across users. Thus, the time between assessments was treated continuously in our analysis. For the analysis, assessments completed within 182 days after the first assessment were included. Consequently, the potential range of time between a user's first and last assessment was 42 to 182 days.

Happify Scale

The Happify Scale is a 9-item, proprietary measure of subjective well-being. This scale was designed to measure two primary components of subjective well-being: positive emotionality and life satisfaction [31]. A total of 4 items assessed the frequency of positive emotions over the past month (eg, "In the past month, how often have you felt joyous, exuberant, inspired, and/or awestruck?") and were based on the Positive and Negative Affect Schedule, a widely used measure of positive and negative emotions [39]. These items are rated on a 5-point scale ranging from 0 (*never*) to 4 (*very often [almost every day]*). The 5 remaining items assess user satisfaction across various life domains (eg, "How satisfied do you feel with the relationships in your life?") and were based on the Satisfaction with Life Scale [40], which is a widely used measure of life satisfaction. These items were rated on a 7-point scale ranging from 0 (*very dissatisfied*) to 6 (*very satisfied*). Scores on each subscale were computed by summing the ratings and computing a percentage score ranging from 0 to 100. A composite score was then generated by averaging the 2 percentages, with higher scores indicating greater subjective well-being.

As described elsewhere [29], an unpublished internal validation study was conducted on the Happify Scale with 559 adults recruited from the general population using Amazon MTurk. In this study, the Happify Scale was shown to have good internal consistency (α=.89), and the internal consistency for each subscale was at least adequate (positive emotions: α=.72; life satisfaction: α=.88). The subscales were also strongly associated with the scales they were based on (r=0.76-0.80), and the total Happiness Scale score was strongly correlated with the Subjective Happiness Scale (r=0.78) and the Center for Epidemiological Studies Depression Scale (r=−0.75), in the predicted direction.

GAD-2 Scale

The GAD-2 [38] is a 2-item screening tool for generalized anxiety disorder, consisting of the first 2 items from the longer 7-item GAD scale. Respondents indicate the extent to which they have been bothered by each of the issues over the past 2 weeks on a scale from 0 (*not at all*) to 3 (*nearly every day*). The ratings were summed so that higher scores indicated greater anxiety symptoms. Although the GAD-2 is typically used as a screening tool, where a score of ≥3 indicates the likelihood of an anxiety disorder, it has also been used as a continuous scale to assess changes in anxiety in response to treatment [41].

Statistical Analysis

To examine changes in well-being and anxiety symptoms, we fit linear mixed effects models for each outcome, treating days from the first assessment to each subsequent assessment as a fixed effect. We selected this approach because of its ability to handle a varied number of assessments across participants at various time points. Models were fitted with R (version 4.1.0 [42]), using the nlme package [43]. Models with random intercepts only and those with both random intercepts and random slopes were evaluated using the Akaike Information Criterion. The final models were fitted using random intercepts. Owing to the variability in the number of assessments and time between assessments, we fitted a continuous autoregressive error structure, conditional autoregressive (1) [44]. For each outcome, we fit a main effects model and then a model with an added interaction between time and use level (recommended vs below recommended). Each model included time, use (recommended level vs below recommended level), gender (dummy coded as "woman"), relationship status (dummy coded as "in a relationship"), number of chronic conditions, number of activities completed before the first assessment, and initial scores on the other outcome variables as predictors. Model comparison and selection were then made using Akaike Information Criterion; the interaction model better fit the data for both well-being and anxiety symptoms and thus, are reported here. All statistical tests were 2-tailed with an α criterion of .050.

Results

Sample and Demographics

During the qualifying period, 1292 new users reported being ≥65 years, residing in the United States, and completing at least two in-app assessments. After excluding participants whose second assessment was >182 days from their first assessment (n=89), whose last assessment was <42 days from their first (n=476), who completed no activities between their first and last assessments (n=306), or who had missing demographic information (n=46), our final sample consisted of 375 older adults. The sample demographics are presented in Table 1.

Users completed an average of 4.17 (SD 2.63) assessments, ranging from 2 to 13 assessments, with an average of 49.89 (SD 38.52) days between assessments. Initial well-being was not correlated with the number of completed assessments (r=.06) or with the average number of days between assessments (r=−.07), whereas initial anxiety symptoms were significantly correlated with both (r=−0.17 and 0.21, respectively; P<.001).

XSL•FO
RenderX

Table 1. Baseline sample characteristics (N=375).

Characteristic	Value
Gender[a], n (%)	
Woman	297 (79.2)
Man	76 (20)
Other	2 (1)
Relationship status, n (%)	
In a relationship	265 (70.7)
Single	110 (29)
Chronic conditions	
Users with at least one chronic condition, n (%)	269 (71.7)
Number of chronic conditions, mean (SD)	1.39 (1)
Self-reported chronic physical conditions, n (%)	
Arthritis	56 (15)
Asthma	17 (5)
Cancer	12 (3)
Chronic pain	70 (19)
Diabetes	37 (10)
Eczema	8 (2)
Heart disease	4 (1)
High blood pressure and/or cholesterol	127 (33.9)
Insomnia	80 (21)
Migraine	19 (5)
Multiple sclerosis	2 (1)
Psoriasis	4 (1)
Rheumatoid arthritis	10 (3)
Other conditions	76 (20)

[a]During onboarding, users are asked “Everyone’s Different: Tell Us Your Gender.” Before October 2020, response options were “male,” “female,” and “none of the above”; after this time, response options were changed to “man,” “woman,” and “none of the above.” Users who selected “male” or “man” are both represented in this table as “man,” whereas those who selected “female” or “woman” are both represented under “woman.”

Use

We were able to verify how 326 of the 375 users (86.9%) accessed the Happify program. A small proportion of these users (54/326, 16.6%) accessed the program exclusively using a computer. Older adults were more likely to access Happify either using a smartphone or tablet (129/326, 34.3%) or a mix of devices (143/326, 43.8%). Among those who used either a smartphone or a tablet, the program was accessed primarily via a smartphone (117/129, 90.5%) compared with the tablet (12/129, 9.5%). Among those who used all 3 devices to access the program, access via a smartphone was the most common (mean 50.78%, SD 22.46%; range 0% -75%), followed by access via a computer (mean 36.12%, SD 10.63%; range 20%-50%) and a tablet (mean 13.10%, SD 18.25%; range 0%-25%).

The sample use statistics are listed in Table 2. Overall, older adult users completed an average of 43.35 (SD 87.80) activities, ranging from 1 to 876 activities, between their first and last assessment. On average, more activities were completed within a dedicated track (mean 35.47, SD 67.31; range 0-415) compared with activities completed via the instant play feature (mean 7.88, SD 37.56; range 0-558). A total of 113 (30.1%) of the 375 retained users engaged with the program at the recommended level of 2 activities per week during that period, which is consistent with other Happify research with a different sample [31]. The older adults in our sample also had an average of 19.36 (SD 27.16) active days on Happify, ranging from 1 to 152 days between their first and last assessment.

XSL•FO
RenderX

Table 2. Characteristics of engagement with Happify program.

	Value, mean (SD)	Value, range
Number of in-app assessments	4.17 (2.63)	2-13
Days between first and last assessments	104.54 (46.55)	42-182
Number of days between assessments	49.89 (38.52)	13.75-177
Activities completed before first assessment	1.22 (1.75)	0-19
Activity between first and last assessment		
Total activities completed	43.35 (87.80)	1-786
Track activities completed	35.47 (67.31)	0-415
Instant play activities completed	7.88 (37.56)	0-588
Days between first and last activity	92 (63.20)	1-152
Active days[a]	19.36 (27.16)	1-152

[a]Any day when a user logged on to the Happify platform and completed an activity but does not include days when the user may have logged on without completing an activity (including completing the assessment).

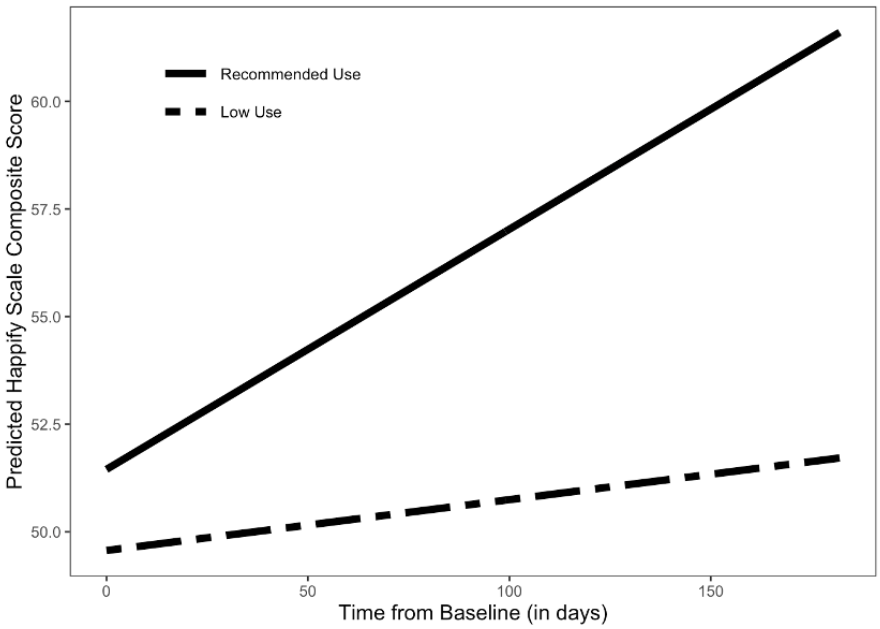
Subjective Well-being

Across the sample, the mean well-being score on the first assessment was 52.56 (SD 19.82), ranging from 5 to 99. This is below the 50th percentile of the Happify Scale in the general population (ie, a score of 61-63) [29]. We found that older adults with a higher number of chronic conditions had lower Happify Scale scores overall, B=−1.58 (95% CI −2.72 to −0.45; *P*=.007), which is consistent with research showing that health status is correlated with subjective well-being [45]. Not surprisingly, older adults with higher levels of anxiety symptoms on their first assessment also had lower levels of subjective well-being overall, B=−5.24 (95% CI −6.10 to −4.38; *P*<.001). These effects were consistent for both the main effects and the interaction models.

We also found significant main effects for both use (B=4.38; 95% CI 1.00-7.77; *P*=.011) and time (B=0.03; 95% CI 0.01-0.04; *P*<.001). However, these effects were qualified by a significant time×use interaction when added to the model (B=0.04; 95% CI 0.02-0.07; *P*=.002), and the main effects were no longer significant for use (B=−1.884; 95% CI −1.874 to 5.641; *P*=.33) or time (B=−0.012; 95% CI −0.005 to 0.029; *P*=.18).

As shown in Figure 3, older adults who completed an average of 2 or more activities per week while on Happify reported significantly greater improvements in subjective well-being than did those who completed fewer activities. More specifically, those who completed an average of 2 or more activities per week had an average improvement of 24.5% (SD 79.3%) in their Happify Scale scores compared with 11.7% (SD 45.5%) among those who engaged below the recommended level. No other effects were significant in either model.

Figure 3. Changes in subjective well-being (as measured by the Happify Scale) over time as moderated by use (recommended use: average of ≥2 activities per week; low use: average of <2 activities per week).



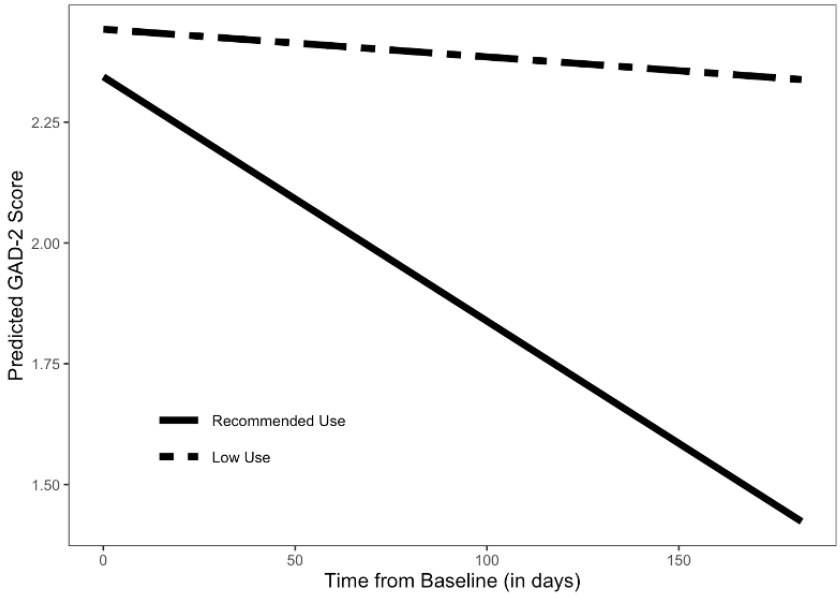
XSL•FO
RenderX

Anxiety

The mean GAD-2 scores on the first assessment were 2.14 (SD 1.82), ranging from 0 to 6, and 116 of the 375 (30.9%) users scored above the cutoff (ie, scores between 3 and 6) for likely anxiety disorder. In addition to the significant main effects of use (B=−0.31, 95% CI −0.60 to −0.02; *P*=.03) and time (B=−0.002, 95% CI: −0.004 to −0.001; *P*<.001), the only other significant main effect was for initial Happify Scale scores. Older adults with higher Happify Scale scores on their first assessment also had significantly lower levels of anxiety symptoms overall (B=−0.04, 95% CI −0.05 to −0.04; *P*<.001).

As with subjective well-being, both main effects for use and time were qualified by a significant use×time interaction (B=−0.004, 95% CI −0.007 to −0.002; *P*<.001) and were no longer significant once this interaction was added to the model (use (B=−0.098, 95% CI −0.409 to 0.212; *P*=.54); time (B=−0.001, 95% CI −0.002 to 0.001; *P*=.05). As depicted in Figure 4, older adults who completed an average of 2 or more activities per week reported significantly greater improvements in anxiety symptoms than those who completed fewer activities. More specifically, among older adults who engaged with the program at the recommended level, there was a 25.6% (SD 58.3%) improvement in GAD-2 scores compared with 10.5% (SD 88.3%) improvement among those who engaged below the recommended level.

Figure 4. Changes in anxiety symptoms (as measured by the Generalized Anxiety Disorder 2 [GAD-2] scale) over time as moderated by use (recommended use: average of ≥2 activities per week; low use: average of <2 activities per week).



Discussion

Principal Findings

The purpose of this analysis was to explore whether older adults would engage with a digital well-being intervention and the extent to which engaging with this intervention led to improvements in subjective well-being and anxiety symptoms. Our results show that approximately one-third of older adults who qualified for our analysis engaged with the program at the recommended level, which is consistent with the rates of engagement reported in other analyses of Happify users that did not include older adults [30].

Our results further showed that when older adults engaged at this optimal level, they experienced significant improvements in both subjective well-being and anxiety symptoms over time. These improvements are also comparable (or better) to those reported in other populations. For instance, a real-world analysis of Happify users found a 10.47% improvement in positive emotions (a subscale from the Happify Scale) over 8 weeks [31], whereas a more recent analysis of Happify users with self-reported migraines, where <1% of the sample was represented by older adults, found an average improvement of 23.5% in subjective well-being and 26.05% in anxiety symptoms [46]. The magnitude of change observed in the current analysis is consistent with those observed in the migraine analysis and better than what was reported in the study by Carpenter et al [31], with improvements in well-being averaging 24.5% and improvements in anxiety symptoms averaging 25.6%. These findings add to the growing body of literature suggesting that although technology adoption remains lower among older adults [21], they nevertheless respond to digital interventions [23,25-27,47]. By examining data from real-world users, our data provide evidence that even outside controlled research environments, older adults will sign up for digital well-being interventions and engage with them at comparable rates to younger users. This is important to understanding whether digital interventions are a viable means of addressing the unmet need for mental health care among older adults, regardless of their efficacy. If older adults do not engage with digital interventions, they cannot improve access to care.

XSL•FO
RenderX

Improving Uptake of Digital Interventions Among Older Adults

Our data show that older adults can benefit from digital well-being interventions when they engage with them and that those who adopt this technology appear to engage at rates similar to what we find in younger age groups. However, it is worth noting that the uptake of these digital interventions was lower than that in other age groups. In the case of the Happify program, older adults who qualified for our analysis made up <3% of the general user base that met all other criteria for inclusion. Therefore, it is important to explore methods to increase older adults' willingness to use these interventions.

One potential explanation for this lower uptake is that older adults are less comfortable with emerging technologies. Indeed, research has shown that older adults who are more confident in their ability to use computers and the internet are more willing to adopt new technologies [48] and become long-term users of new digital programs [49]. Some researchers suggest that tablets may be the answer to increased adoption of technology among older adults [50]; however, our data suggest that only a small proportion of older adults on Happify used tablets to access the program. Generally, older adults still seemed to use smartphones, followed by computers, to access the program. Although more research is needed to explore the impact of optimizing digital interventions for tablets, these data suggest that the mode of delivery for digital interventions is not the root cause of lower uptake.

Rather, the issue may have more to do with a lack of familiarity with the interventions themselves than with technology. For example, qualitative research with adults aged ≥50 years suggests that the most common barrier to uptake of digital interventions in this age group is a lack of understanding [28]. More specifically, many participants were unaware of digital interventions, although they developed positive attitudes toward such interventions when introduced to them during the study session. However, other research suggests that awareness alone may not be sufficient to overcome older adults' skepticism about how digital interventions can help improve their mental health [51]. Consequently, education to improve awareness among older adults may need to be coupled with support while learning new technologies to increase use [52].

Designing Digital Interventions With Older Adults in Mind

Beyond efforts to introduce older adults into digital interventions, we also need to consider the unique needs of this population, which might make digital interventions designed for younger groups impractical for older adults. For instance, qualitative research with older adults suggests that although participants felt there were numerous benefits to technology, many reported concerns with usability based on age- or health-related changes in abilities (eg, difficulties with small screens and manual dexterity) [51,53]. Certain features, such as audio and voice recognition technology, may be required to increase the usability of digital interventions in this population [22]. Despite the proliferation of digital interventions available on the market, few, if any, have been developed specifically with older adults in mind. Given the increasing need for scalable mental health solutions among older adults, and the unique barriers to engaging with digital tools in this age group, product development that actively includes end user feedback will be imperative to the success of digital interventions among older adults, both in terms of uptake and efficacy. Indeed, other researchers have called for patient-centered or user-focused research with older adults as part of digital intervention development [54] or even exploring opportunities for participatory co-design [55]. Although qualitative research on older adults' general perspectives on digital interventions and technology exists [28], more user-centered work on older adults as they engage with specific programs is needed.

Strengths and Limitations

Although the strength of this study is its ability to provide insights into the real-world uptake of digital well-being interventions among older adults, it also has several limitations. First, given the lack of a control group, we could not determine whether the changes in well-being or anxiety symptoms were directly related to the intervention. We found that changes in outcomes were significantly different based on use, such that older adults who engaged with Happify at a minimal level showed less improvement in both well-being and anxiety symptoms compared with those who completed an average of at least two activities per week. This moderating effect of use suggests that the completion of Happify activities contributed, at least in part, to changes in outcomes. However, research with a control group is required to determine the causality. In particular, given recent criticisms that the effects of digital interventions are much weaker when compared to active controls [56], future research should include a rigorous control that would account for potential placebo effects as well as time.

Second, because of the naturalistic design, our analyses were limited to users who signed up for the program on their own. Consequently, it is likely that the older adults included in our analysis were not representative of all older adults. In particular, our sample predominantly consisted of older adults who identified as women. Although this is often the case in research on digital interventions and we found no significant effects of gender in either of our models, some research suggests that women may show greater improvements in mental health and well-being outcomes after engaging in digital interventions [57]. Consequently, these findings may not be generalizable to men.

This is compounded by the fact that our analysis included those users who engaged with the intervention for at least six weeks, who completed activities, and who completed at least two in-app assessments. Thus, our sample likely represents older adults who are early adopters of digital interventions and are more comfortable with the technology overall. Although it is important to understand how this group of older adults will respond to digital interventions, to determine whether these interventions are a viable solution to address the unmet need for mental health care among older adults, we need to test the usability of these interventions with a broader population of older adults. In particular, it is important to test the impact of digital interventions among older adults who may experience more barriers to engaging with these technologies, including those less familiar or comfortable with technology, those with

XSL•FO
RenderX

conditions that might interfere with their ability to use digital tools (eg, cognitive deficits and mobility concerns), and those from diverse backgrounds.

Finally, although we were able to isolate users aged ≥65 years, because the question about age in the onboarding questionnaire is categorical, it is impossible to examine the continuous effects of age within this cohort. Research suggests that technology adoption may be even lower among adults aged ≥75 years [21] and that the benefits of digital interventions may be negatively related to age among older adults [23]. Conceivably, our effects may be driven by younger older adults, and in future research, it will be important to determine users' age more precisely.

Conclusions

As the population ages, the increasing need for mental health care coupled with the shortage of mental health professionals specializing in geriatric populations presents important concerns regarding unmet care needs. Although digital interventions have been presented as one way to address unmet needs in the general population, few studies have specifically examined the impact of such interventions on older adults. The current data add to the growing body of evidence suggesting that although older adults are less likely to begin using digital interventions without efforts to familiarize themselves with these interventions, those who engage with these interventions show corresponding improvements in their mental health. This suggests that digital interventions may present a viable opportunity to improve access to mental health care among older adults. Importantly, digital health programs may also help foster a sense of independence among older adults [58], offering them opportunities to address mental health concerns without feeling like they burden others [51]. Given the potential benefits, developing digital interventions specifically for older adults to address their unique needs and to provide education surrounding digital interventions to improve awareness of and comfort with these tools among older adults should be a priority.

Acknowledgments

We would like to thank Robert Vlisides-Henry for his assistance with follow-up analyses on this project.

Data Availability

The data used for the analyses described in this paper are owned by Twill Inc and cannot be shared publicly without violating the terms, conditions and privacy policy agreed to by the users. Data may be shared upon request by contacting the corresponding authors.

Authors' Contributions

EB generated ideas for real-world analysis and led to the writing of the manuscript. TP was responsible for pulling the data used for the analysis. RH was responsible for analyzing the data. HW, SES, and AP provided support throughout this study.

Conflicts of Interest

EB, TP, and SES are full-time employees of Twill, the company owning the digital program described in this paper, and hold stock options. AP, HW, and RH were employed by Twill when the data were analyzed and continue to hold stock options. None of the authors were compensated based on the outcomes presented herein.

References

1. World Population Prospects 2019: Highlights. Department of Economic and Social Affairs, Population Division, United Nations. 2019 Jun 17. URL: https://www.un.org/development/desa/publications/world-population-prospects-2019-highlights.html#:~:text=The%20world's%20population%20is%20expected,United%20Nations%20report%20launched%20today [accessed 2022-04-07]
2. Vespa J, Armstrong DM, Medina L. Demographic Turning Points for the United States: Population Projections for 2020 to 2060. United States Census Bureau. 2020 Feb. URL: https://www.census.gov/library/publications/2020/demo/p25-1144.html [accessed 2022-04-07]
3. Dall TM, Gallo PD, Chakrabarti R, West T, Semilla AP, Storm MV. An aging population and growing disease burden will require a large and specialized health care workforce by 2025. Health Aff (Millwood) 2013 Nov;32(11):2013-2020. [doi: 10.1377/hlthaff.2013.0714] [Medline: 24191094]
4. Rowe JW, Fulmer T, Fried L. Preparing for better health and health care for an aging population. JAMA 2016 Oct 25;316(16):1643-1644. [doi: 10.1001/jama.2016.12335] [Medline: 27668895]
5. Odden MC, Coxson PG, Moran A, Lightwood JM, Goldman L, Bibbins-Domingo K. The impact of the aging population on coronary heart disease in the United States. Am J Med 2011 Sep;124(9):827-33.e5 [FREE Full text] [doi: 10.1016/j.amjmed.2011.04.010] [Medline: 21722862]
6. Wroe PC, Finkelstein JA, Ray GT, Linder JA, Johnson KM, Rifas-Shiman S, et al. Aging population and future burden of pneumococcal pneumonia in the United States. J Infect Dis 2012 May 15;205(10):1589-1592. [doi: 10.1093/infdis/jis240] [Medline: 22448012]

XSL•FO
RenderX

7. Mental Health of Older Adults: Fact Sheets. World Health Organization. 2017 Dec 12. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2022-04-08]
8. Mohebbi M, Agustini B, Woods RL, McNeil JJ, Nelson MR, Shah RC, ASPREE Investigator Group. Prevalence of depressive symptoms and its associated factors among healthy community-dwelling older adults living in Australia and the United States. Int J Geriatr Psychiatry 2019 Aug;34(8):1208-1216 [FREE Full text] [doi: 10.1002/gps.5119] [Medline: 30989707]
9. Vasiliadis HM, Dionne PA, Préville M, Gentil L, Berbiche D, Latimer E. The excess healthcare costs associated with depression and anxiety in elderly living in the community. Am J Geriatr Psychiatry 2013 Jun;21(6):536-548. [doi: 10.1016/j.jagp.2012.12.016] [Medline: 23567409]
10. Choi S, Lee S, Matejkowski J, Baek YM. The relationships among depression, physical health conditions and healthcare expenditures for younger and older Americans. J Ment Health 2014 Jun;23(3):140-145. [doi: 10.3109/09638237.2014.910643] [Medline: 24803220]
11. Temple JB, Brijnath B, Enticott J, Utomo A, Williams R, Kelaher M. Discrimination reported by older adults living with mental health conditions: types, contexts and association with healthcare barriers. Soc Psychiatry Psychiatr Epidemiol 2021 Jun;56(6):1003-1014. [doi: 10.1007/s00127-020-01914-9] [Medline: 32696302]
12. Crabb R, Hunsley J. Utilization of mental health care services among older adults with depression. J Clin Psychol 2006 Mar;62(3):299-312. [doi: 10.1002/jclp.20231] [Medline: 16400646]
13. Garrido MM, Kane RL, Kaas M, Kane RA. Use of mental health care by community-dwelling older adults. J Am Geriatr Soc 2011 Jan;59(1):50-56 [FREE Full text] [doi: 10.1111/j.1532-5415.2010.03220.x] [Medline: 21198461]
14. Robb C, Haley WE, Becker MA, Polivka LA, Chwa HJ. Attitudes towards mental health care in younger and older adults: similarities and differences. Aging Ment Health 2003 Mar;7(2):142-152. [doi: 10.1080/1360786031000072321] [Medline: 12745392]
15. Bartels SJ, Naslund JA. The underside of the silver tsunami--older adults and mental health care. N Engl J Med 2013 Feb 07;368(6):493-496. [doi: 10.1056/NEJMp1211456] [Medline: 23343039]
16. Lattie EG, Stiles-Shields C, Graham AK. An overview of and recommendations for more accessible digital mental health services. Nat Rev Psychol 2022 Jan 26;1(2):87-100. [doi: 10.1038/s44159-021-00003-1]
17. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health-a systematic review and meta-analysis. NPJ Digit Med 2019 Dec 2;2:118 [FREE Full text] [doi: 10.1038/s41746-019-0188-8] [Medline: 31815193]
18. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
19. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
20. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
21. Perrin A. Mobile Technology and Home Broadband 2021. Pew Research Center. 2021 Jun 3. URL: https://www.pewresearch.org/internet/2021/06/03/mobile-technology-and-home-broadband-2021/ [accessed 2022-04-08]
22. Evangelista L, Steinhubl SR, Topol EJ. Digital health care for older adults. Lancet 2019 Apr 13;393(10180):1493 [FREE Full text] [doi: 10.1016/S0140-6736(19)30800-1] [Medline: 30983579]
23. Xiang X, Wu S, Zuverink A, Tomasino KN, An R, Himle JA. Internet-delivered cognitive behavioral therapies for late-life depressive symptoms: a systematic review and meta-analysis. Aging Ment Health 2020 Aug;24(8):1196-1206 [FREE Full text] [doi: 10.1080/13607863.2019.1590309] [Medline: 30913898]
24. Shah SG, Nogueras D, van Woerden HC, Kiparoglou V. Evaluation of the effectiveness of digital technology interventions to reduce loneliness in older adults: systematic review and meta-analysis. J Med Internet Res 2021 Jun 04;23(6):e24712 [FREE Full text] [doi: 10.2196/24712] [Medline: 34085942]
25. Mewton L, Sachdev PS, Andrews G. A naturalistic study of the acceptability and effectiveness of internet-delivered cognitive behavioural therapy for psychiatric disorders in older australians. PLoS One 2013 Aug 12;8(8):e71825 [FREE Full text] [doi: 10.1371/journal.pone.0071825] [Medline: 23951253]
26. Reinwand DA, Schulz DN, Crutzen R, Kremers SP, de Vries H. Who follows eHealth interventions as recommended? A study of participants' personal characteristics from the experimental arm of a randomized controlled trial. J Med Internet Res 2015 May 11;17(5):e115 [FREE Full text] [doi: 10.2196/jmir.3932] [Medline: 25963607]
27. Graham SA, Stein N, Shemaj F, Branch OH, Paruthi J, Kanick SC. Older adults engage with personalized digital coaching programs at rates that exceed those of younger adults. Front Digit Health 2021 Aug 6;3:642818 [FREE Full text] [doi: 10.3389/fdgth.2021.642818] [Medline: 34713112]
28. Pywell J, Vijaykumar S, Dodd A, Coventry L. Barriers to older adults' uptake of mobile-based mental health interventions. Digit Health 2020 Feb 11;6:2055207620905422 [FREE Full text] [doi: 10.1177/2055207620905422] [Medline: 32110429]

XSL•FO
RenderX

29. Parks AC, Williams AL, Kackloudis GM, Stafford JL, Boucher EM, Honomichl RD. The effects of a digital well-being intervention on patients with chronic conditions: observational study. J Med Internet Res 2020 Jan 10;22(1):e16211 [FREE Full text] [doi: 10.2196/16211] [Medline: 31922491]
30. Montgomery RM, Boucher EM, Honomichl RD, Powell TA, Guyton SL, Bernecker SL, et al. The effects of a digital mental health intervention in adults with cardiovascular disease risk factors: analysis of real-world user data. JMIR Cardio 2021 Nov 19;5(2):e32351 [FREE Full text] [doi: 10.2196/32351] [Medline: 34806986]
31. Carpenter J, Crutchley P, Zilca RD, Schwartz HA, Smith LK, Cobb AM, et al. Seeing the "Big" picture: big data methods for exploring relationships between usage, language, and outcome in internet intervention data. J Med Internet Res 2016 Aug 31;18(8):e241 [FREE Full text] [doi: 10.2196/jmir.5725] [Medline: 27580524]
32. Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive Therapy of Depression. New York, NY, USA: Guilford Press; 1987.
33. Praissman S. Mindfulness-based stress reduction: a literature review and clinician's guide. J Am Acad Nurse Pract 2008 Apr;20(4):212-216. [doi: 10.1111/j.1745-7599.2008.00306.x] [Medline: 18387018]
34. Sin NL, Lyubomirsky S. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. J Clin Psychol 2009 May;65(5):467-487. [doi: 10.1002/jclp.20593] [Medline: 19301241]
35. Hayes SC, Strosahl KD, Wilson KG. Acceptance and Commitment Therapy: The Process and Practice of Mindful Change. 2nd edition. New York, NY, USA: Guilford Press; 2011.
36. Dimidjian S, Barrera Jr M, Martell C, Muñoz RF, Lewinsohn PM. The origins and current status of behavioral activation treatments for depression. Annu Rev Clin Psychol 2011;7:1-38. [doi: 10.1146/annurev-clinpsy-032210-104535] [Medline: 21275642]
37. Chambless DL, Hollon SD. Defining empirically supported therapies. J Consult Clin Psychol 1998 Feb;66(1):7-18. [doi: 10.1037//0022-006x.66.1.7] [Medline: 9489259]
38. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
39. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988 Jun;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
40. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]
41. Staples LG, Dear BF, Gandy M, Fogliati V, Fogliati R, Karin E, et al. Psychometric properties and clinical utility of brief measures of depression, anxiety, and general distress: the PHQ-2, GAD-2, and K-6. Gen Hosp Psychiatry 2019;56:13-18 [FREE Full text] [doi: 10.1016/j.genhosppsych.2018.11.003] [Medline: 30508772]
42. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2021. URL: https://www.r-project.org/ [accessed 2022-04-01]
43. Pinheiro J, Bates D, R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. The Comprehensive R Archive Network. 2022 Jun 15. URL: https://cran.r-project.org/web/packages/nlme/index.html [accessed 2022-04-01]
44. Jones RH. Longitudinal Data with Serial Correlation: A State-space Approach. New York, NY, USA: CRC Press; 1993.
45. Ngamaba KH, Panagioti M, Armitage CJ. How strongly related are health status and subjective well-being? Systematic review and meta-analysis. Eur J Public Health 2017 Oct 01;27(5):879-885. [doi: 10.1093/eurpub/ckx081] [Medline: 28957478]
46. Weiner L, Boucher EM, Zucker M, Parks AC. The effects of a digital mental health intervention in migraine patients: a pilot study. In: American Academy of neurology Annual Meeting. 2022 Presented at: AAN '22; April 2-7, 2022; Seattle, WA, USA.
47. Stockwell S, Schofield P, Fisher A, Firth J, Jackson SE, Stubbs B, et al. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: a systematic review and meta-analysis. Exp Gerontol 2019 Jun;120:68-87. [doi: 10.1016/j.exger.2019.02.020] [Medline: 30836130]
48. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. Innov Aging 2018 Jan;2(1):igy002 [FREE Full text] [doi: 10.1093/geroni/igy002] [Medline: 30480129]
49. Mitzner TL, Savla J, Boot WR, Sharit J, Charness N, Czaja SJ, et al. Technology adoption by older adults: findings from the PRISM trial. Gerontologist 2019 Jan 09;59(1):34-44 [FREE Full text] [doi: 10.1093/geront/gny113] [Medline: 30265294]
50. Tsai HY, Shillair R, Cotten SR, Winstead V, Yost E. Getting grandma online: are tablets the answer for increasing digital inclusion for older adults in the U.S.? Educ Gerontol 2015;41(10):695-709 [FREE Full text] [doi: 10.1080/03601277.2015.1048165] [Medline: 26877583]
51. Andrews JA, Brown LJ, Hawley MS, Astell AJ. Older adults' perspectives on using digital technology to maintain good mental health: interactive group study. J Med Internet Res 2019 Feb 13;21(2):e11694 [FREE Full text] [doi: 10.2196/11694] [Medline: 30758292]
52. Seifert A, Reinwand DA, Schlomann A. Designing and using digital mental health interventions for older adults: being aware of digital inequality. Front Psychiatry 2019 Aug 9;10:568 [FREE Full text] [doi: 10.3389/fpsyt.2019.00568] [Medline: 31447716]

XSL•FO
RenderX

53. Pirhonen J, Lolich L, Tuominen K, Jolanki O, Timonen V. “These devices have not been made for older people's needs” – older adults' perceptions of digital technologies in Finland and Ireland. Technol Soc 2020 Aug;62:101287. [doi: 10.1016/j.techsoc.2020.101287]
54. Singh K, Drouin K, Newmark LP, Lee J, Faxvaag A, Rozenblum R, et al. Many mobile health apps target high-need, high-cost populations, but gaps remain. Health Aff (Millwood) 2016 Dec 01;35(12):2310-2318. [doi: 10.1377/hlthaff.2016.0578] [Medline: 27920321]
55. Cosco TD, Fortuna K, Wister A, Riadi I, Wagner K, Sixsmith A. COVID-19, social isolation, and mental health among older adults: a digital catch-22. J Med Internet Res 2021 May 06;23(5):e21864 [FREE Full text] [doi: 10.2196/21864] [Medline: 33891557]
56. Goldberg SB, Lam SU, Simonsson O, Torous J, Sun S. Mobile phone-based interventions for mental health: a systematic meta-review of 14 meta-analyses of randomized controlled trials. PLOS Digit Health 2022;1(1):e0000002 [FREE Full text] [doi: 10.1371/journal.pdig.0000002] [Medline: 35224559]
57. Przybylko G, Morton D, Morton J, Renfrew M. The influence of gender and age on the outcomes of and adherence to a digital interdisciplinary mental health promotion intervention in an Australasian nonclinical setting: cohort study. JMIR Ment Health 2021 Nov 11;8(11):e29866 [FREE Full text] [doi: 10.2196/29866] [Medline: 34762058]
58. Kruse CS, Mileski M, Moreno J. Mobile health solutions for the aging population: a systematic narrative analysis. J Telemed Telecare 2017 May;23(4):439-451. [doi: 10.1177/1357633X16649790] [Medline: 27255207]

Abbreviations

CBT: cognitive behavioral therapy
GAD-2: Generalized Anxiety Disorder 2

Edited by T Leung, J Wang; submitted 25.05.22; peer-reviewed by X Xiang, J Andrews, J Sharit; comments to author 15.06.22; revised version received 24.06.22; accepted 28.06.22; published 02.09.22.

Please cite as:
Boucher E, Honomichl R, Ward H, Powell T, Stoeckl SE, Parks A
The Effects of a Digital Well-being Intervention on Older Adults: Retrospective Analysis of Real-world User Data
JMIR Aging 2022;5(3):e39851
URL: https://aging.jmir.org/2022/3/e39851
doi:10.2196/39851
PMID:36053569

©Eliane Boucher, Ryan Honomichl, Haley Ward, Tyler Powell, Sarah Elizabeth Stoeckl, Acacia Parks. Originally published in JMIR Aging (https://aging.jmir.org), 02.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Scaling Up Citizen Workshops in Public Libraries to Disseminate and Discuss Primary Care Research Results: Quasi-Experimental Study

José Massougbodji[1], MSc, MPH, MD; Hervé Tchala Vignon Zomahoun[1], BSc, MSc, PhD; Evehouenou Lionel Adisso[1], BSc, MSc; Jasmine Sawadogo[2], MSc; Valérie Borde[3], MBA; Cynthia Cameron[4], MD; Hélène Moisan[5], MSI; Jean-Sébastien Paquette[4], MD, MSc; Zamzam Akbaraly[6]; Lëa-Kim Châteauneuf[7], MSI; Geneviève David[8], BSc, MSc, MPA; Bibliothèques2 Group[9]; France Légaré[10], BSc Arch, MD, MSc, PhD, CCFP, FCFP

[1]Department of Social and Preventive Medicine, Laval University, Québec, QC, Canada

[2]First Nations of Quebec and Labrador Health and Social Services Commission, Québec, QC, Canada

[3]Centre d'excellence sur le Dialogue entre les scientifiques et le public, Québec, QC, Canada

[4]Department of Family Medicine and Emergency Medicine, Laval University, Québec, QC, Canada

[5]Bibliothèque de Québec, Québec, QC, Canada

[6]Patient and Public Partnership Research Strategy Component, Quebec SPOR-SUPPORT Unit, Laval University, Québec, QC, Canada

[7]Direction des bibliothèques, Service de la culture - Ville de Montréal, Montreal, QC, Canada

[8]Centre d'excellence sur le partenariat avec les patients et le public, Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montreal, QC, Canada

[9]See Acknowledgments, Québec, QC, Canada

[10]VITAM – Centre de recherche en santé durable, Centre intégré universitaire de santé et services sociaux de la Capitale-Nationale, Québec, QC, Canada

Corresponding Author:
France Légaré, BSc Arch, MD, MSc, PhD, CCFP, FCFP
VITAM – Centre de recherche en santé durable
Centre intégré universitaire de santé et services sociaux de la Capitale-Nationale
2480, chemin de la Canardière
Québec, QC, G1G 2G1
Canada
Phone: 1 418 663 5713
Email: France.Legare@mfa.ulaval.ca

Abstract

Background: Little is known about engaging patients and stakeholders in the process of scaling up effective knowledge translation interventions targeting the public.

Objective: Using an integrated knowledge translation approach, we aimed to scale up and evaluate an effective pilot program to disseminate research results in public libraries.

Methods: We conducted a scaling-up study targeting the public. On the basis of our successful pilot project, we codeveloped and implemented a large-scale program of free citizen workshops in public libraries, in a close research partnership with stakeholders and patient representatives. Citizen workshops, each facilitated by 1 participating physician and 1 science communicator, consisted of a 45-minute computer-assisted presentation and a 45-minute open exchange. The intervention outcome was knowledge gained. The scale-up outcomes were satisfaction, appropriateness, coverage, and costs. An evaluation questionnaire was used to collect data of interest. Both quantitative and qualitative analyses were performed.

Results: The workshop theme chosen by the patient and stakeholder representatives was the high prevalence of medication overuse among people aged ≥65 years. From April to May 2019, 26 workshops were conducted in 25 public libraries reaching 362 people. The mean age of participants was 64.8 (SD 12.5) years. In total, 18 participating physicians and 6 science communicators facilitated the workshops. Participants reported significant knowledge gain (mean difference 2.1, 95% CI 2.0-2.2; P<.001). The median score for overall public satisfaction was 9 out of 10 (IQR 8-10). The public participants globally rated the workshops as having a high level of appropriateness. Coverage was 92% (25/27) of the total number of public libraries targeted. Costs were CAD $6051.84 (US $4519.69) for workshop design and CAD $22,935.41 (US $17,128.85) for scaling them up.

XSL•FO
RenderX

Conclusions: This project successfully established a large-scale and successful implementation science or knowledge translation bridge among researchers, clinicians, and citizens via public libraries. This study provides a model for a dissemination practice that benefits the public by both engaging them in the dissemination process and targeting them directly.

(JMIR Aging 2022;5(3):e39016) doi:10.2196/39016

KEYWORDS

scaling up; knowledge translation; dissemination strategies; integrated knowledge translation; public libraries; citizen workshops; potentially inappropriate medicines

Introduction

Scale Up of Health Interventions

Much research remains on the shelf. The average delay for integrating research findings into healthcare delivery is still estimated in units of decades, despite recent advances made in implementation science or knowledge translation (both hereafter referred to as KT) [1-3]. While KT attempts to address this gap, most KT interventions target health professionals and ignore the public [4]. Within the clinical context of primary care, the public and patients are key end users of research findings. They should be informed about new evidence that could benefit them and be involved in any KT process that targets them [5,6]. Without patient and stakeholder involvement in judging the relevance of the knowledge being transferred, the new knowledge may not be patient-centered and remains in the hands of the professionals delivering care. Any patient engagement in KT is still mostly *low level* engagement [7]. Most health intervention pilot projects, even if proven effective, remain on the shelf. One way to bring effective pilot projects off the shelf is to scale them up so that their benefits reach a broader population. Scaling up is becoming an important motor of KT and is developing into a science unto itself [8,9]. The process of *scaling up* can be defined as "deliberate efforts to increase the impact of successfully tested health innovations so as to benefit more people and to foster policy and program development on a lasting basis." [10].

KT interventions rarely target the public directly, who are their potential if not actual patients. Even web-based surveys are unrepresentative of the public, as they only reach people with educational and technological resources [11]. Meanwhile, public libraries are known for their extensive population reach, as they attract homeless and other marginalized patrons [12-14]. Their patrons also see public libraries as a valuable resource for medical information [15,16]. Furthermore, unlike other service-providing institutions (eg, medical and some social welfare institutions), libraries are widely trusted by the public [14]. Therefore, they can be an excellent avenue for disseminating accurate medical information to users. Ultimately, this could lead to increased public expectations and demands for care that is more patient-centered, thus changing the dynamics of care between patients and providers [17] by fostering positive behaviors such as shared decision-making by both partners in the care relationship [18].

Workshops in Public Libraries as Effective KT Interventions

In 2017, we established a proof of concept on the dissemination of research results to the public through workshops in public libraries. These pilot workshops, designed to raise awareness of new knowledge in primary care research, were conducted across 9 public libraries in Quebec City, Quebec, Canada. We demonstrated evidence of their effectiveness by measuring the acquisition of knowledge among participants [19]. First, we hypothesized that one of the reasons for our pilot's positive results was the library setting. Public libraries are free community-based civic institutions associated with increasing knowledge at one's own pace and in accordance with one's interests. This setting thus helped reduce the usual power differential between health professionals and patients, as these potential patients had freely chosen to be present instead of being obliged to hear messages from health professionals. Second, we give credence to our communication strategy, whereby research findings were delivered by physicians who were credible messengers and by a science communicator using plain language accessible to lay people. Third, we successfully mobilized several key stakeholders, such as physicians, a science communicator, and a public library manager, and attracted the public. Overall, our successful pilot workshops appeared to be an appropriate candidate for scaling up, according to a World Health Organization guide to scale up [10]. Evidence of their effectiveness was sound, observable, and documented. They had already been tested in a setting similar to the target setting. We succeeded in maintaining comparable participation rates for workshops across public libraries, which was a good indication of the generalizability of our project. The model was easily transferable, it matched the values of the target institutions (such as libraries), and similar logistics could be applied. On the basis of our body of evidence, our next step was to investigate how these results would hold on a larger scale; that is, by targeting more public libraries and delivering more workshops. Ultimately, we expected that reaching a larger public and increasing their knowledge would greatly impact population health.

However, there is no point in scaling up KT intervention projects that are not relevant to their target populations. Knowledge must also be accessible to end users. To this end, the integrated knowledge translation (IKT) approach has been increasingly adopted in implementation studies [20]. IKT aims to gather the views of all stakeholders, including knowledge users, throughout the research process in an inclusive, engaging, and interactive manner [21]. It is based on research partnership, the equitable sharing of power, and mutual respect among all stakeholders.

XSL•FO
RenderX

The benefits of this approach have been widely demonstrated in the literature [22]. In this scaled-up version of our pilot, we planned to engage patients and stakeholders from start to finish, involving men and women at a *high level* of engagement [23]; that is, in choosing the theme, defining its content, and evaluating its outcomes while maintaining or improving workshop effectiveness.

Therefore, we aimed to scale up an effective pilot program to disseminate research results to the public through citizen workshops in public libraries using an IKT approach, while maintaining fidelity and with equal or improved effectiveness.

Methods

As no specific reporting guidelines for scaling up studies are available, we used an adapted version of the Standards for Reporting Implementation Studies guidelines to report our study [24]. We also relied on the Template for Intervention Description and Replication reporting guidelines for the description of the intervention [25].

Study Design

As with the pilot project, we conducted a pre-post intervention study. Participants self-reported both preintervention and postintervention outcome measures only after the intervention was completed (reducing response shift bias for the outcome measures and the burden on participants) [26,27]. Using an IKT approach, we adapted the pilot methodology to engage patients and stakeholders throughout this scaling up study. Therefore, this study was not registered.

Context

While the pilot project took place in Quebec City, Quebec, Canada, the scaled-up intervention was extended to Montreal, Quebec, Canada, which, similar to Quebec City, is largely Francophone, so the culture and language were similar. However, it should be noted that Montreal has a higher immigrant population and includes the city of Westmount, Quebec, Canada, which is more Anglophone. There are also economic differences between localities, with Westmount being richer than most districts of Montreal and Quebec City.

Targeted Sites and Population

Convenience sampling was conducted to select libraries in Quebec City, Montreal, and Westmount, which were able to include a citizen workshop in their spring 2019 agenda and had the necessary amenities (ie, video projector, laptop, speaker, and room for 30 people). The target population for our study was public library patrons aged ≥18 years. Their participation was voluntary. To maximize the number of participants in the libraries, and also to ensure a variety of profiles (eg, sex, age, and education level) among all participants, libraries were free to schedule citizen workshops on the dates and times they deemed most convenient (ie, during the workday or in the evenings and weekdays or weekends).

Planning to Scale Up the Intervention

Establishing a Committee

This scale-up study began with the formation of a preliminary project steering committee and was informed by the Canadian Institutes of Health Research Integrated and End-of-grant knowledge translation frameworks [28]. The proposal for scaling up the pilot intervention, that is, a presentation with layman- and user-friendly content followed by an exchange period, was consensually retained by the committee.

Name

We named this intervention *citizen workshops* because of the strong involvement and responsibility of all stakeholders in the process: primary care researchers would produce results and make them available to physicians, science communicators, and patient and research partners for dissemination to the public; the public would identify the most relevant results; public libraries would host the workshops; and science communicators would facilitate them.

Recruitment

Through a convenience selection, we recruited stakeholder representatives, including 4 experts in patient-oriented research, 1 science communicator, 1 primary care physician, and 2 public library officials. The primary health care researcher, whose results were selected for dissemination, and a patient expert (ie, a patient or informal caregiver trained in research), who was a caregiver of a patient facing the health problem addressed, joined the committee for the remaining stages. The library officials on the committee arranged for invitations to be sent to all public libraries in Montreal, Westmount, and Quebec City and then helped to identify the libraries that would host the citizen workshops. In addition to the libraries' usual information channels (programming pamphlets, websites, and social media platforms), posters, a dedicated website [29] and radio advertisements, social media platforms (Facebook), and newsletters were used to reach the participants. These means of promotion were designed and approved by all final committee members including the patient expert.

The committee decided that each citizen workshop would be moderated by a team consisting of a family physician as the speaker and a science communicator in charge of facilitating and articulating the message in plain language. Thus, researchers and physicians on the committee issued a letter to be included in primary care professional and research organization newsletters, inviting any willing primary care physician (emergency and family medicine) or resident to participate in the project. The only prerequisite was that they had to have good knowledge and practical experience of the health problem addressed. Science communicators were selected by the science communicator member of the steering committee according to their ability to communicate orally in plain language, to lead a constructive discussion with an audience, and to manage the unforeseen (inappropriate questions, speaker forgetting important details or explaining key concepts poorly, or technical or operational mishaps) and their respectful and empathetic attitude.

XSL•FO
RenderX

IKT Strategies

We have involved different stakeholders as research partners at all steps of our research process, except the patient caregiver and the representative of winning researchers who were identified and involved after the selection of the research results to be disseminated. The research results selected allowed us to determine the health problem addressed and therefore the relevant profile of patients to be involved as research partners in our study. The relevant patient profile we determined required the permanent support of a caregiver. Therefore, we included as the research partner a caregiver having substantial life experience with the patient having a health problem of interest.

As research partners, the stakeholders contributed to the research process using the following strategies: (1) the members of the preliminary steering committee discussed and approved the aim of the citizen workshops, which was to inform and raise awareness of the research results that would be selected. (2) Then, they identified the theme for which research results would be disseminated in the citizen workshops: a call for research results was issued to all primary health care research teams in the province with a letter, validated by all committee members, encouraging them to submit their research results for free dissemination. The physicians and researchers on the committee helped to identify the best means for disseminating the call for research results throughout the province of Quebec (dissemination networks of primary care research centers, university hospital centers, faculties of medicine, pharmacy, and nursing care in universities in Quebec). (3) Applications were evaluated exclusively by the preliminary committee members who reflected the voice of the main stakeholders, including librarians, primary care physicians, experts in patient-oriented research, and science communicators. (4) All final committee members, including the patient expert and the owner of the selected research results, participated in writing and approving the script for the workshops. (5) Patients and stakeholders on the committee were involved in all stages of the implementation, including workshop observation, data collection, and outcomes evaluation. (6) Meetings among all actors, including the patient expert, were organized on a bimonthly basis. A progress report was sent every 2 months over 18 months. (7) Patients and stakeholders on the committee were assigned to observe all workshops. Their role was to distribute and collect evaluation questionnaires completed by the participants of the citizen workshop. They ensured the smooth running of the workshop and counted the number of participants at the beginning of each citizen workshop. (8) Preliminary results of the data analysis for the project were discussed with stakeholders, and their comments and suggestions were considered in the final interpretation of the results.

Implementing the Scaled-up Intervention

Preparation

To ensure consistency of citizen workshops across libraries, materials for moderation (ie, PowerPoint [Microsoft Inc] presentation, handouts, and notes for each individual moderator) were sent 6 weeks earlier to all participating moderators. They had 2 weeks to familiarize themselves with the materials. Then, two 1-hour preparatory meetings, spaced 2 weeks apart, were held by the committee and all participating moderators. During these meetings, the committee gathered moderators' feedback on the documents for consideration and sought their approval.

Workshop Content

Each citizen workshop was divided into two 45-minute equivalent parts: the first was a computer-assisted presentation of the results, and the second was a knowledge exchange between participants and physician presenters moderated by a science communicator. The knowledge exchange included not only the question-and-answer round but also knowledge sharing through the lenses of scientific evidence, beliefs, personal experiences, and values disseminated about the topic. First, an introductory part raised public awareness of the health issues related to the results by defining terms and providing context. Then followed the actual results of the selected study and a detailed description of their direct impact on the public and potential repercussions on their health. In the knowledge exchange session, the science communicator ensured that any questions from the participants did not seek a personalized medical consultation. This format was identical to that of the pilot phase workshops, except for one major adaptation: the addition of a video clip to the presentation in the first part of the workshop with the testimony of the patient expert associated with the project. We made this change because, in the pilot study, the workshops generated greater gains in knowledge among young people than among older people. We hypothesized that older people might need information presented in a different format to reach them better. On completion of the workshops, participants were left with a handout outlining the research results along with additional documents and resources about the health problem addressed. Detailed information on the content of the intervention and handouts can be available upon request.

Maintaining Fidelity

Except for the addition of the video clip to the presentation, efforts were made to maintain fidelity to the piloted workshop concept and content. Workshops were given in French in all libraries, even in areas that were predominantly Anglophone. The same content was offered with moderators having comparable profiles. To maintain fidelity, we had to add some elements to the new contexts; for example, some public libraries did not have projectors for the slide presentations with sound, so we purchased our projection materials.

Evaluation

Outcomes of Scaling Up

Outcomes of scaling up were related to selected aspects of acceptability and appropriateness of citizen workshops among participants, workshop coverage, time, and costs.

According to the taxonomy of implementation outcomes by Proctor et al [30], acceptability is the perception among stakeholders that an innovation is agreeable or satisfactory while appropriateness is the perceived fit, relevance, or compatibility of the innovation. These outcomes were measured using 12 closed-ended questions regarding participants' opinions of the workshop. Acceptability was measured using 3 questions that

XSL•FO
RenderX

focused on the structure of the activity, 3 questions on workshop facilitation, and 2 questions on whether the workshop met their expectations and whether they would recommend it to others. Participants also indicated their overall satisfaction with the workshop using a discrete 11-point scale where 0 corresponded to unsatisfied and 10 corresponded to fully satisfied.

Appropriateness, on the other hand, was measured using 4 questions on the workshop quality and relevance. Answers to all questions except general satisfaction were chosen from a 4-point Likert scale (ranging from 1, "totally disagree," to 4, "totally agree"). Qualitative data on participants' acceptability were also collected from open-ended questions in the evaluation form.

Coverage was determined by determining the ratio between the numerator (ie, the number of libraries that hosted the workshops) and the denominator (ie, the number of libraries targeted for participation).

A partial economic evaluation focusing solely on costs was conducted separately for the workshop design costs and the scaling-up costs to distinguish between modifiable costs related to the scale-up strategy and nonmodifiable costs related to the intervention. Costs for scaling up included remuneration of steering committee members, medical moderators, science communicators, and patient observers; purchase of the necessary equipment; and actual delivery of workshops. Expenses related to designing the citizen workshops included fees for steering committee members for designing and writing the workshop script, as well as filming the video clip incorporated into the presentation.

Intervention Outcomes

The main outcome of the intervention was knowledge gain, as perceived by the participants about the health problem addressed. To assess this, we adapted the self-administered questionnaire used in our pilot study [19]. This questionnaire was administered to participants at the beginning of the workshop, and they were invited to complete it at the end (Multimedia Appendix 1). They rated their knowledge using a discrete scale from 0 (very low) to 10 (very high).

Data were also collected on participants' sociodemographic characteristics, such as age, sex, and highest level of education. Finally, other variables pertaining to workshop characteristics were collected by direct observation during their delivery: the time of day during which the workshops were held, the presence or absence of the patient expert as an observer during the workshop, and whether the speaker was a physician or resident.

Analysis

First, we performed a descriptive analysis of the participants according to their sociodemographic characteristics and the workshops they attended, their opinions, and their levels of satisfaction and knowledge.

We used a 2-tailed paired *t* test to compare self-reported pre- and postknowledge levels [31,32]. Comparative analyses of the knowledge gain were then conducted according to the characteristics of not only the participants but also the workshop in which they participated. To this end, univariate linear regression models of knowledge gain were constructed [33]. To assess how knowledge gains would vary across public libraries, comparisons were also made according to the workshops' moderators (ie, each facilitator, speaker, and pair of moderators) using an ANOVA test [34]. However, given the skewed distributions of knowledge levels and gain, sensitivity analyses were performed: first, the Wilcoxon signed rank comparison test was used to compare before and after median knowledge levels [35]. Second, unmatched rank tests on the median and nonparametric multiple comparisons were performed using the SAS NPAR1WAY procedure [36]. Statistical significance was defined as $P<.05$ (2-sided test).

All analyses were performed in the SAS software (version 9.4; SAS Institute Inc). Qualitative data collected through open-ended questions were transcribed by 1 author (JS) and analyzed using an iterative deductive method discussed with team members. For the economic evaluation, we calculated the sum of expenses separately for the scaling-up strategies and for the design of citizen workshops. The cost results are presented in Canadian dollars. On April 5, 2019, CAD $1 was equal to US $0.75.

Ethics Approval and Consent to Participate

Ethics approval was granted from the *Comité d'éthique du Centre intégré universitaire de santé et services sociaux de la Capitale-Nationale* under project 2019-1513. Informed consent was first obtained verbally from the study participants at the beginning of each conference as the conferences were recorded. Written consent was obtained from the participants who agreed to complete the conference evaluation forms. The study protocol was approved by the ethics committee.

Results

Overview

Following the committee's call for research results, 5 research teams submitted their results. The results selected that responded to public or patient interest, according to the selection committee, addressed the high prevalence of the use of potentially inappropriate medicines among people aged ≥65 years in Quebec [37].

Population

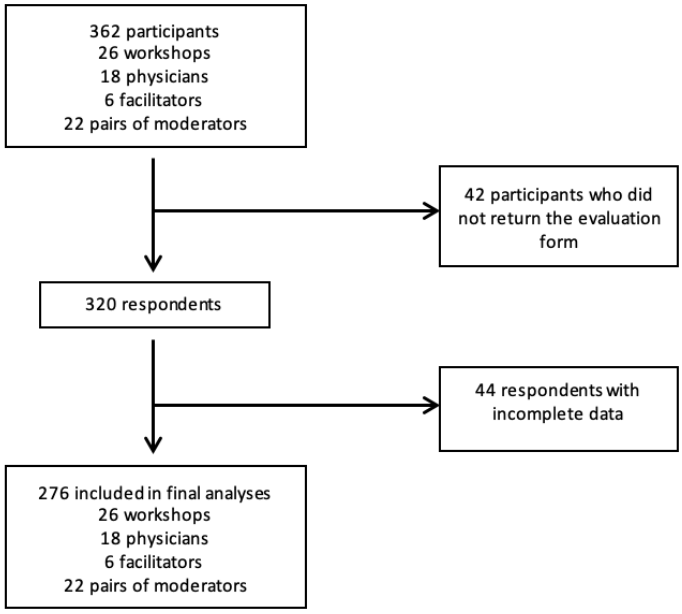
A total of 25 libraries, including 9 in Quebec City and 16 in Montreal, agreed to host the citizen workshops. From April 4 to May 29, 2019, 26 workshops were offered in Montreal, including 1 workshop in Westmount and 10 workshops in Quebec City, with 1 library agreeing to host 2 workshops. A total of 18 physicians were mobilized to present the selected findings, and 6 facilitators were recruited. Consequently, 22 distinct pairs of moderators were assembled.

The citizen workshops drew 362 participants, with a mean of 13.9 (SD 6.0) participants per workshop. The evaluation questionnaire was returned by 320 participants (Figure 1). Table 1 presents the sociodemographic characteristics of the participants and characteristics of the workshops in which they participated. The mean age of the participants from the public was 64.8 (SD 12.5) years. Women accounted for 71.6%

XSL•FO
RenderX

(229/320) of the participants from the public, and half had a university-level education (172/320, 53.8%). Approximately half (150/320, 46.9%) attended workshops in the evening, and 18.1% (58/320) had a patient partner present at their workshop. Most participants (279/320, 87.2%) had a physician as the speaker.

Figure 1. A flowchart of participants in citizen workshops.



XSL•FO
RenderX

Table 1. Distribution of participants and citizen workshops and their characteristics.

	Montreal and Westmount (n=178)	Quebec (n=142)	Total (N=320)
Participant characteristics			
Sex, n (%)			
Female	128 (71.9)	101 (71.1)	229 (71.6)
Male	38 (21.4)	38 (26.8)	76 (23.8)
Missing data	12 (6.7)	3 (2.1)	15 (4.7)
Highest educational level, n (%)			
Secondary or lower	29 (16.3)	25 (17.6)	54 (16.8)
College	42 (23.6)	34 (23.9)	76 (23.8)
University	94 (52.8)	78 (54.9)	172 (53.8)
Missing data	13 (7.3)	5 (3.5)	18 (5.6)
Age (years)			
Value, mean (SD)	65.5 (12.4)	64 (12.6)	64.8 (12.5)
Missing data (participants)	13	5	18
Workshop characteristics			
Time of day, n (%)			
Morning	34 (19.1)	47 (33.1)	81 (25.3)
Afternoon	81 (45.5)	8 (5.6)	89 (27.8)
Evening	63 (35.4)	87 (61.3)	150 (46.9)
Presence of the patient partner, n (%)			
Present	58 (32.6)	0 (0)	58 (18.1)
Absent	120 (67.4)	142 (100)	262 (81.9)
Qualification of physician speaker, n (%)			
Physician only	163 (91.6)	116 (81.7)	279 (87.2)
Physician + resident	0 (0)	20 (14.1)	20 (6.6)
Resident only	15 (8.4)	6 (4.2)	21 (6.6)

Outcomes

Outcomes of Scaling Up

Coverage

Of the 27 public libraries initially planned for the citizen workshops, 25 held workshops, corresponding to a coverage of 92%.

Acceptability and Appropriateness of Citizen Workshops, According to the Public

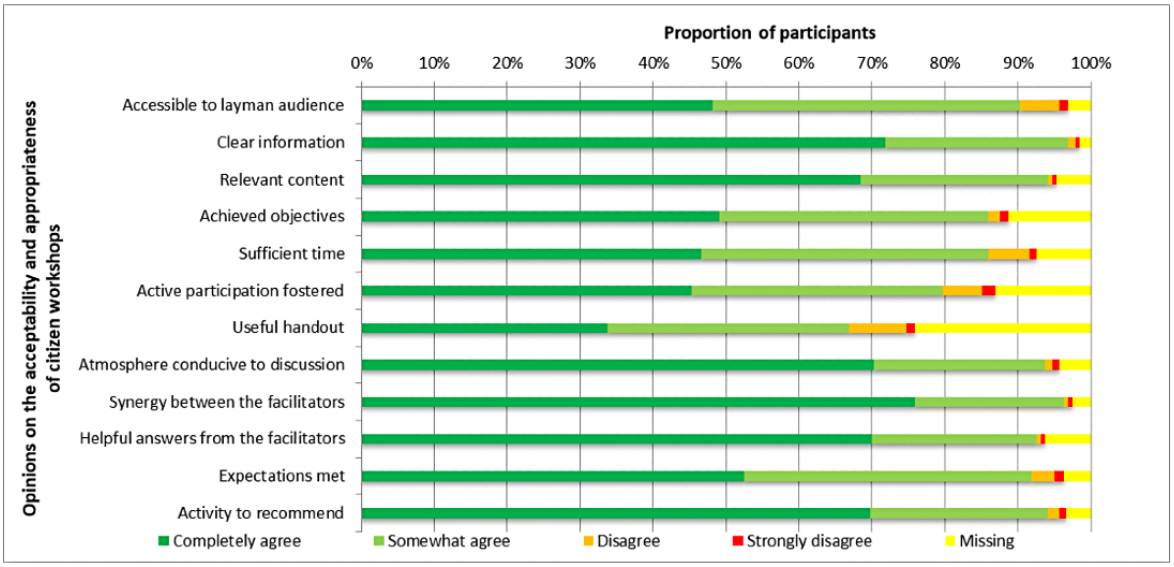
The median level of overall satisfaction was 9 (IQR 8.0-10) out of 10. With regard to qualitative data, participants pointed out the good quality of the PowerPoint presentation. They particularly liked the inclusion of the interview with the patient partner in the layout of the presentation. This could be considered an indicator of the value of the patient caregiver in IKT. Many participants also perceived and praised the effort to communicate the research results in plain language through the PowerPoint presentation and during workshop facilitation. However, participants expressed some negative impressions, notably that several libraries were open-plan concept and therefore did not have dedicated rooms for this type of activity. Although most participants found the length of the workshops adequate (275/320, 86%), some found that there was not enough time to discuss their concerns. The lowest approval score was obtained for an item that assessed whether their active participation had been encouraged (255/320, 79.7%). However, for the same item, a high rate of missing responses (42/320, 13.1%) was noted. Regarding the moderation of the workshops, most participants reported that the moderators provided an atmosphere conducive to discussion (299/320, 93.5%) and gave them useful answers (296/320, 92.5%). They also appreciated the enthusiasm of the moderators and their complementarity (308/320, 96.2%). Finally, most participants felt that the workshop met their expectations (294/320, 91.9%), and 94.1% (301/320) recommended the activity to others (Figure 2).

In terms of appropriateness, more than 9 out of 10 participants found that the citizen workshops were accessible to a layman audience and that the information presented to them was clear and relevant. However, a low agreement was obtained regarding the usefulness of the documentation provided to them (214/320, 66.9%). This was also the item for which the proportion of

XSL•FO
RenderX

missing responses was the highest (79/320, 24.8%). However, many participants found that information in the handouts was too brief, and 1 participant suggested a more substantial document with more information such as examples, useful websites, and a detailed outline of the presentation.

Figure 2. Public participants' opinions on citizen workshops (N=320).



Cost and Time

Workshop Design

In total, 16 people were mobilized to participate in the committee. Regarding the design of citizen workshops, costs were mainly the fees of the science communicator member of the steering committee for the writing of the workshop's script and for the shooting of video clips embedded in the presentation. These costs were CAD $6051.84 (US $4519.69). The patient caregiver who had worked for the Quebec Strategy for Patient-Oriented Research Support for People and Patient-Oriented Research and Trials Unit did not receive additional compensation in our study context. The script revision and video clip editing were free, as they were performed by other members of the steering committee with the tools already at their disposal in their workplaces.

Scaled-up Workshop Delivery

None of the steering committee members billed for their time since they were professionals who were already paid in their respective workplaces, except the science communicator and the patient expert. Their fee for scale up was CAD $3511.05 (US $2622.15). A software was purchased for the posters and the website creation at a cost of CAD $453.10 (US $338.39). The preparatory meetings for the scaled-up workshops, in terms of travel, per diem, and food, cost CAD $4380.12 (US $3271.20). For the scaled-up delivery of the citizen workshops, 7 external observers were mobilized in addition to the 24 moderators (18 physicians and 6 facilitators). The external observers were research assistants and graduate students. They were mandated to give and collect the evaluation questionnaire completed by participants. They also counted the number of participants and noted any incident occurred during the citizen workshop. The per diem, travel, and accommodation expenses of observers and moderators totaled CAD $13,620.65 (US $10,172.31). The material used during the workshops (office supplies, recorders, and pointers) was evaluated at CAD $970.49 (US $724.79). Total costs for scaling up the intervention were CAD $22,935.41 (US $17,128.85). Therefore, the overall cost for the project was CAD $28,987.25 (US $21,648.55).

The duration of the scaling-up process using the IKT approach, from the creation of the steering committee to the beginning of the citizen workshops, was 17 months and 8 months, respectively, longer than that of the pilot project.

Intervention Outcomes

Knowledge Gain

The final analyses were carried out on 276 participants after removing those whose information on their level of knowledge about potentially inappropriate medicines either before or after the citizen workshops was missing. On a knowledge scale of 0 to 10, participants reported that they were, on average, fairly well informed about MIPs before the citizen workshops (mean 6.2, SD 1.8) and more so afterwards (8.2, SD 1.4). This represented a significant (P<.001) mean increase in knowledge of 2.1 (95% CI 2.0-2.2). Neither the range of participants' sociodemographic profiles, the workshop characteristics, nor the variety of workshop moderators (as individuals or as pairs) appeared to modify the effect of the workshop on knowledge gain (Table 2).

These results were confirmed in our sensitivity analysis (Multimedia Appendix 2).

XSL•FO
RenderX

Table 2. Comparison of knowledge gain among participants in citizen workshops (N=276)[a].

Characteristics	Population, n	β (95% CI)	*P* value[b]
Participant characteristics			
Sex			.97
Male (reference female)	208	.01 (−0.46 to 0.48)	
Age (years)	276	−.02 (−0.02 to 0.00)	.06
Highest educational level			.28
Up to secondary (reference university)	48	.01 (−0.54 to 0.56)	
College (reference university)	70	.37 (−0.10 to 0.85)	
Workshop characteristics			
Time of day			.59
Morning (reference evening)	72	−0.18 (−0.68 to 0.31)	
Afternoon (reference evening)	76	.10 (−0.39 to 0.58)	
Presence of the patient partner			.38
Present (reference absent)	51	.23 (−0.29 to 0.75)	
Qualification of physician speaker			.22
Physician + resident (reference physician only)	18	−0.72 (−1.54 to 0.10)	
Resident only (reference physician only)	19	.05 (−0.75 to 0.84)	
According to the physician speaker	18[c]	N/A[d]	.63[e]
According to the facilitator	6[f]	N/A	.47[e]
According to the pair of moderators	22[g]	N/A	.60[e]

[a]N=276 (after deletion of observations with missing variables).

[b]*P* value of linear bivariate regression.

[c]Number of physician's groups.

[d]N/A: not applicable.

[e]*P* value of ANOVA test.

[f]Number of facilitator's groups.

[g]Number of pair of moderator's groups.

Harms

No harm was reported from stakeholders or workshop participants.

Discussion

Principal Findings

We aimed to evaluate the scaling up of an effective pilot program to disseminate research results through citizen workshops in public libraries. The main departure of the scaled-up intervention from strict fidelity to the pilot intervention was that we adopted an IKT approach to ensure that the citizen workshops faithfully reflected the needs and interests of patients and other stakeholders at every step of the intervention. We achieved high coverage of the project to scale up the workshops, which generated high levels of satisfaction among participants and high levels of acceptability and appropriateness. Participants in the scaled-up citizen workshops also reported an increase in their knowledge level of the subject being discussed. These findings lead us to make the following observations.

First, we achieved high coverage for the scaling-up citizen workshop. This finding could be explained, in part, by the topic being disseminated that was of great interest to most public library users (old people). Another explanation could be that the citizen workshop was integrated into the conference program of participating public libraries. Therefore, there was no additional logistical management that could limit the participation of public libraries.

Second, our scaled-up citizen workshops led to an increase in knowledge among participants. Interactive workshops have been established as ideal for sharing knowledge across professional and sectoral boundaries [38]. In this project, the interactive aspect was emphasized as much as possible by adding a video clip to the initial format of the workshops to better communicate the patient's perspective. Although the participants in the *scaling-up* audience were much older than those in the pilot project audience, our scaled-up citizen workshops, in addition to being highly satisfying, led to an improvement in

XSL•FO
RenderX

knowledge among these participants. These results confirm the importance of designing a more detailed and inclusive format for citizens' workshops, regardless of the topic under discussion, to increase knowledge among all age ranges within the audience. However, it should also be noted that these results did not allow us to assess the extent to which an increase in knowledge among public participants produced behavioral changes. A study in the United States that evaluated the midterm impact of after-school nutrition workshops in a public library setting and that targeted adolescents and their parents, a program deemed by the authors to be of low intensity even though it consisted of 5 workshops, did not produce any lasting behavioral change after just 3 months [39]. Our citizen workshops, which were one-time events, sought primarily to raise awareness, with behavioral change as an indirect goal. The next step would be to evaluate the immediate and midterm impacts of citizen workshops among the public by assessing health outcome data related to the themes, both at the time of the workshops and at intervals afterward.

Third, adopting an IKT approach improved our scaling-up results in the following ways: (1) the involvement of library network stakeholders in identifying participating libraries could explain the high coverage of our scaling-up project; (2) prioritizing the public's perspective to identify the results to be disseminated, adopting a co-constructive approach to designing the workshops, and holding preparatory meetings to allow workshop moderators to make the content of the message their own are all reasons that could explain our positive results in terms of acceptability and appropriateness among the public. These positive findings are also consistent with those of our (non-IKT) pilot project. However, interestingly, they also turned out to be of equal magnitude [19] despite the differences between the pilot project and the scaling-up project. This last observation also holds true for the increase in knowledge. This maintenance of improved outcomes despite the change in subject matter, the involvement of various workshop moderators, and the sociodemographic and linguistic differences within the participating public libraries is likely due to the modification of the intervention by incorporating an IKT approach from start to finish.

Fourth, to the best of our knowledge, this is the first scaling-up study to address such high levels of patient and other stakeholder engagement. Our scaled-up version of the workshop achieved fidelity in terms of being true to the concept and content from one site to another and largely true to the concept implemented in the pilot trial, with the addition of a patient-designed video clip. In this video clip, a caregiver having substantial life experience with the patient told the patient's story. We involved this caregiver in the rest of the research process once the research results to be disseminated were selected. Therefore, the caregiver contributed as a research partner to the research team meetings; workshop content development, planning, and scaling up; and the revision of different research documents (eg, materials related to citizen workshop). However, our pilot project did not use an IKT approach; thus, in theory, our scaled-up version of the intervention did not meet the strictest fidelity requirements of adhering to the intervention, as outlined in the original pilot design. This raises an interesting question about KT. If new knowledge emerges between the pilot program and the scaling-up phase (eg, evidence about the importance of high-level patient engagement), should the scaled-up intervention maintain fidelity at all costs or should this new knowledge be integrated into the scaled-up version? The science of scaling up must not restrict researchers to reproduce interventions at a scale that excludes important new knowledge. Indeed, we propose that, going forward, the IKT approach should be, as far as possible, an essential and integral dimension of scaling up. At first glance, IKT appears to be a cumbersome approach because it requires constant consultation and adaptation that could slow the process of scaling up [40,41]. However, it ensures that the effectiveness of the interventions would not be diluted with scaling up and that the interventions are worth scaling up because they respond to the real needs and interests of patients and other stakeholders. In this sense, IKT can also be perceived as a necessary regulator of the upscaling process.

Fifth, as Milat et al [42] suggested, before scaling up an intervention, evidence of its effectiveness should ideally be provided through randomized controlled trials. In our case, it was impossible to manipulate exposure to the intervention, and so our evidence was from a natural experiment performed in the real world. Therefore, we skipped the randomized controlled trial step and went straight from our pilot project, a feasibility study, to the scale-up phase. However, the results of the pilot phase had already provided us information on scalability elements. Scaling up has been taking place, under different names, for several decades (especially in low- and middle-income countries for quickly stemming the spread of infectious diseases) [43], and current scale-up efforts in low- and middle-income countries show that scale-up strategies must be sufficiently flexible to respond to emerging questions [44]. Scaling up is still a new science and, as Milat et al [42] concede, must build flexibility in its application to real-world interventions.

Sixth, Quebec City, where our pilot took place, is almost unilingual and Francophone. Our workshops were scaled up to include libraries in Montreal, which has more immigrants and is more culturally diverse, and Westmount, which is more Anglophone. Although we did not measure these contextual differences in our sociodemographic questionnaires, the positive and consistent effect of citizen workshops on knowledge gain is a good indication that extending our model to more diverse populations will maintain acceptability and knowledge acquisition levels. However, this does not preclude the importance of adapting to different sociodemographic profiles when scaling up. Further adaptations may depend on the theme addressed, the target population, and the social situation. For example, the modalities of mass gatherings have changed dramatically during the COVID-19 pandemic. As a result, modifications in delivery will have to be made to our citizen workshops to follow public health recommendations.

Finally, we lacked the opportunity to conduct a complete economic (cost-effectiveness) analysis. However, our partial cost evaluation could be useful in the future for scaling-up studies, which so far have rarely included economic evaluations [43]. In addition, costs are considered an essential reporting

XSL•FO
RenderX

item in the proposed guidelines for reporting on scaling-up studies [45]. Full economic evaluations in the real context of scaling up will also help choose efficient strategies involving the high-level engagement of patients and stakeholders across the scaling-up process and predicting the economic and human resource costs of further scale up.

Limitations

The limitation of our study was, first, the fact that it had no comparison group. However, our earlier pilot project results helped us better understand some of the findings. It would be interesting to compare the costs of using an IKT approach to scale up our model without integrating patients and stakeholders, although it is ethically questionable. Second, we included only 1 patient caregiver in the research process. This could have limited the variety of patient perspectives in our research process. However, our patient caregiver, as the daughter of the patient, had not only a great life experience with the patient but also a substantial experience of health system use. Indeed, the caregiver supported her mother during different phases of her disease progression. Third, participants in the citizen workshops were self-selected citizens who responded to an advertisement for the workshop. However, self-selection sampling has some advantages: it reduces recruitment time, and self-selected participants are more likely to be committed to participate in the study (eg, more willing to spend time filling in the questionnaire) and to provide insights into the theme [46]. Nevertheless, we failed to meet the more vulnerable populations with lower literacy levels: half of the public in the workshops were university graduates and therefore not representative of Quebec's overall older adult population literacy level. Fourth, the data were collected using self-reporting tools; however, the impact of this on the effectiveness analysis should be, if anything, an underestimation of the knowledge gain among participants.

Conclusions

This project successfully established a large-scale and successful KT bridge among researchers, clinicians, and citizens via public libraries. We found that scaling up a program of citizen workshops in public libraries resulted in high levels of knowledge gain, content appropriateness, and acceptability. The addition of an IKT approach involving patients and other stakeholders as research partners throughout the process and remunerating them improved the final product without harming the scale-up outcomes. These findings, based on citizen workshops integrating a computer-assisted presentation on scientific evidence and patient video clips plus a knowledge exchange session, highlight that an IKT approach and patient-oriented research should no longer be optional. This study provides a model for a dissemination practice that benefits the public by targeting and directly engaging them in the dissemination process. Public libraries are free and power-neutral educational institutions, and this simple and reproducible intervention is a ground-breaking knowledge translation model.

Acknowledgments

The authors express their most devout and profound thoughts for their late colleague Sylvie Frechette.

The authors thank the Bibliothèques2 Group: Mame-Awa Ndiaye, Patient and public partnership research strategy component of the Quebec Strategy for Patient-Oriented Research Support for People and Patient-Oriented Research and Trials (SPOR-SUPPORT) Unit, Université Laval; Léa Langlois, VITAM: Centre de recherche en santé durable, Quebec; Priscille-Nice Sanon, Patient and public partnership research strategy component of the Quebec SPOR-SUPPORT Unit, Université Laval; Sarah Osama, Direction des bibliothèques, Service de la culture - Ville de Montréal; Danielle Gilet, Hôpital Hôtel-Dieu de Sorel, Centre Intégré de Santé et de Services sociauxde la Montérégie-Est; Frantz-Daniel Lafortune, Department of Family Medicine and Emergency Medicine, Université Laval; Valérie Levée, Centre DECLIC (Dialogue Entre les sCientifiques et Le publiC); Charles Desjardins, Department of Family Medicine and Emergency Medicine, Université Laval; Annik Desfossés, Department of Family Medicine and Emergency Medicine, Université de Montréal; Louise-Isabelle Rivard, Department of Family Medicine and Emergency Medicine, Université de Montréal.

They also thank Audrey-Ann Lavoie, Alexandre Laurin, Hélène Landry, Céline Gauthier, Olivia Nguyen, Marie-Frédérique Thibault, Catherine Belanger, Frantz-Daniel Lafortune, Minh Vu, Maude Soulieres, Géraldine Layani, Sylvie Rheault, Hannah Beattie, Michel Rochon, Marine Corniou, Carine Monat, and Renaud Manuguerra for their contribution to the implementation of the citizen workshops. The authors also thank Tatiana Agbadje, Ella Diendéré, Nathalie Rheault, Amédé Gogovor, Louise Légaré, Lobna Khadraoui, and Sylvain Thivierge for their contribution during the citizen workshop implementation and Louisa Blair for her editorial support.

This study was conducted by the Quebec SPOR-SUPPORT Unit. The unit was supported by the Canadian Institutes of Health Research and provincial partners, including the *Ministère de la Santé et des Services sociaux du Québec* and the *Fonds de recherche du Québec—Santé*. They had no role in the design of the study and the collection, analysis, and interpretation of the data, or in writing the manuscript. The information provided and the views expressed in this paper are the responsibility of the authors alone.

Data Availability

All data and materials used during the study will be available from the corresponding author (FL) upon request.

XSL•FO
RenderX

Authors' Contributions

JM, HTVZ, ELA, VB, CC, HM, GD, JS, JSP, ZA, LKC, and FL contributed to the conception and design of the study. JM, HTVZ, ELA, and JS contributed to the analysis and interpretation of data. JM, HTVZ and FL drafted the manuscript. All authors have read, substantively revised, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Evaluation forms.
[PDF File (Adobe PDF File), 211 KB - aging_v5i3e39016_app1.pdf]

Multimedia Appendix 2
Comparison of knowledge gain among participants in citizen workshops (using nonparametric tests).
[PDF File (Adobe PDF File), 93 KB - aging_v5i3e39016_app2.pdf]

References

1. Atkins MS, Rusch D, Mehta TG, Lakind D. Future directions for dissemination and implementation science: aligning ecological theory and public health to close the research to practice gap. J Clin Child Adolesc Psychol 2016;45(2):215-226 [FREE Full text] [doi: 10.1080/15374416.2015.1050724] [Medline: 26155972]
2. Balas EA, Weingarten S, Garb CT, Blumenthal D, Boren SA, Brown GD. Improving preventive care by prompting physicians. Arch Intern Med 2000 Feb 14;160(3):301-308. [doi: 10.1001/archinte.160.3.301] [Medline: 10668831]
3. Curran JA, Grimshaw JM, Hayden JA, Campbell B. Knowledge translation research: the science of moving research into policy and practice. J Contin Educ Health Prof 2011;31(3):174-180. [doi: 10.1002/chp.20124] [Medline: 21953658]
4. Schipper K, Bakker M, De Wit M, Ket JC, Abma TA. Strategies for disseminating recommendations or guidelines to patients: a systematic review. Implement Sci 2016 Jun 07;11(1):82 [FREE Full text] [doi: 10.1186/s13012-016-0447-x] [Medline: 27268061]
5. Kitson A, Powell K, Hoon E, Newbury J, Wilson A, Beilby J. Knowledge translation within a population health study: how do you do it? Implement Sci 2013 May 21;8:54 [FREE Full text] [doi: 10.1186/1748-5908-8-54] [Medline: 23694753]
6. Kothari A, Wathen CN. A critical second look at integrated knowledge translation. Health Policy 2013 Feb;109(2):187-191. [doi: 10.1016/j.healthpol.2012.11.004] [Medline: 23228520]
7. Bombard Y, Baker GR, Orlando E, Fancott C, Bhatia P, Casalino S, et al. Engaging patients to improve quality of care: a systematic review. Implement Sci 2018 Jul 26;13(1):98 [FREE Full text] [doi: 10.1186/s13012-018-0784-z] [Medline: 30045735]
8. Kristensen N, Nymann C, Konradsen H. Implementing research results in clinical practice- the experiences of healthcare professionals. BMC Health Serv Res 2016 Feb 10;16:48 [FREE Full text] [doi: 10.1186/s12913-016-1292-y] [Medline: 26860594]
9. Bégin M, Eggertson L, Macdonald N. A country of perpetual pilot projects. CMAJ 2009 Jun 09;180(12):1185-1E89 [FREE Full text] [doi: 10.1503/cmaj.090808] [Medline: 19506272]
10. Nine steps for developing a scaling-up strategy. Department of Reproductive Health and Research, World Health Organization. Geneve, Switzerland: World Health Organization; 2010. URL: http://apps.who.int/iris/bitstream/handle/10665/44432/9789241500319_eng.pdf;jsessionid=91D2056D7F94844FFD29EDADF393296A?sequence=1 [accessed 2020-05-01]
11. Khazaal Y, van Singer M, Chatton A, Achab S, Zullino D, Rothen S, et al. Does self-selection affect samples' representativeness in online surveys? An investigation in online video game research. J Med Internet Res 2014 Jul 07;16(7):e164 [FREE Full text] [doi: 10.2196/jmir.2759] [Medline: 25001007]
12. Cabello M, Butler SM. How public libraries help build healthy communities. Brookings. 2017 Mar 30. URL: https://www.brookings.edu/blog/up-front/2017/03/30/how-public-libraries-help-build-healthy-communities/ [accessed 2020-05-05]
13. Morgan AU, Dupuis R, D'Alonzo B, Johnson A, Graves A, Brooks KL, et al. Beyond books: public libraries as partners for population health. Health Aff (Millwood) 2016 Nov 01;35(11):2030-2036. [doi: 10.1377/hlthaff.2016.0724] [Medline: 27834243]
14. Philbin MM, Parker CM, Flaherty MG, Hirsch JS. Public libraries: a community-level resource to advance population health. J Community Health 2019 Feb;44(1):192-199 [FREE Full text] [doi: 10.1007/s10900-018-0547-4] [Medline: 29995303]
15. Wood FB, Lyon B, Schell MB, Kitendaugh P, Cid VH, Siegel ER. Public library consumer health information pilot project: results of a National Library of Medicine evaluation. Bull Med Libr Assoc 2000 Oct;88(4):314-322 [FREE Full text] [Medline: 11055298]
16. Arnott-Smith C, Keselman A. Meeting Health Information Needs Outside Of Healthcare: Opportunities and Challenges. Amsterdam, The Netherlands: Elsevier; Aug 4, 2015.

XSL•FO
RenderX

17. Kravitz RL, Epstein RM, Feldman MD, Franz CE, Azari R, Wilkes MS, et al. Influence of patients' requests for direct-to-consumer advertised antidepressants: a randomized controlled trial. JAMA 2005 Apr 27;293(16):1995-2002 [FREE Full text] [doi: 10.1001/jama.293.16.1995] [Medline: 15855433]
18. Shepherd HL, Butow PN, Tattersall MH. Factors which motivate cancer doctors to involve their patients in reaching treatment decisions. Patient Educ Couns 2011 Aug;84(2):229-235. [doi: 10.1016/j.pec.2010.10.018] [Medline: 21112174]
19. Adisso EL, Borde V, Saint-Hilaire ME, Robitaille H, Archambault P, Blais J, et al. Can patients be trained to expect shared decision making in clinical consultations? Feasibility study of a public library program to raise patient awareness. PLoS One 2018 Dec 12;13(12):e0208449 [FREE Full text] [doi: 10.1371/journal.pone.0208449] [Medline: 30540833]
20. Gagliardi AR, Berta W, Kothari A, Boyko J, Urquhart R. Integrated knowledge translation (IKT) in health care: a scoping review. Implement Sci 2016 Mar 17;11:38 [FREE Full text] [doi: 10.1186/s13012-016-0399-1] [Medline: 26988000]
21. Bowen S, Graham ID. Integrated knowledge translation. In: Stauss SE, Tetroe J, Graham ID, editors. Knowledge Translation in Health Care. Hoboken, NJ, USA: John Wiley & Sons; 2013:14-23.
22. Allaire JF, St-Martin KA, Massougbodji J, Zomahoun HT, Langlois L. S'outiller pour favoriser la participation des usagers, des proches, des citoyens et des communautés à l'amélioration continue de la qualité des soins et des services: recueil d'idées inspirantes. Institut universitaire de première ligne en santé et services sociaux. 2018. URL: https://numerique.banq.qc.ca/patrimoine/details/52327/3487091 [accessed 2020-05-01]
23. Pomey MP, Flora L, Karazivan P, Dumez V, Lebel P, Vanier MC, et al. Le « Montreal model » : enjeux du partenariat relationnel entre patients et professionnels de la santé. Sante Publique 2015;1 Suppl:41-50. [Medline: 26414256]
24. Pinnock H, Barwick M, Carpenter CR, Eldridge S, Grandes G, Griffiths CJ, StaRI Group. Standards for Reporting Implementation Studies (StaRI) statement. BMJ 2017 Mar 06;356:i6795 [FREE Full text] [doi: 10.1136/bmj.i6795] [Medline: 28264797]
25. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
26. Geldhof GJ, Warner DA, Finders JK, Thogmartin AA, Clark A, Longway KA. Revisiting the utility of retrospective pre-post designs: the need for mixed-method pilot data. Eval Program Plann 2018 Oct;70:83-89. [doi: 10.1016/j.evalprogplan.2018.05.002] [Medline: 30029016]
27. Pratt CC, McGuigan WM, Katzev AR. Measuring program outcomes: using retrospective pretest methodology. Am J Eval 2000;21(3):341-349. [doi: 10.1177/109821400002100305]
28. Guide to knowledge translation planning at CIHR: integrated and end-of-grant approaches. Canadian Institutes of Health Research. 2012. URL: https://cihr-irsc.gc.ca/e/45321.html#a12 [accessed 2017-06-01]
29. Quebec SPOR-SUPPORT Unit - Health and Social Services Systems, Knowledge Translation and Implementation Component. 2019. URL: https://express.adobe.com/page/fv4fgQ5oKVRrP/ [accessed 2018-03-01]
30. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
31. Student. The probable error of a mean. Biometrika 1908 Mar;6(1):1-25. [doi: 10.2307/2331554]
32. Norman G. Likert scales, levels of measurement and the "laws" of statistics. Adv Health Sci Educ Theory Pract 2010 Dec;15(5):625-632. [doi: 10.1007/s10459-010-9222-y] [Medline: 20146096]
33. Pandis N. Using linear regression for t tests and analysis of variance. Am J Orthod Dentofacial Orthop 2016 May;149(5):769. [doi: 10.1016/j.ajodo.2016.02.007] [Medline: 27131260]
34. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965 Dec 01;52(3/4):591-611. [doi: 10.2307/2333709]
35. West ED. The signed-rank (Wilcoxon) test. Lancet 1969 Mar 08;1(7593):526. [doi: 10.1016/s0140-6736(69)91616-x] [Medline: 4179592]
36. Stokes ME, Davis CS, Koch GG. Categorical Data Analysis Using the SAS System. 2nd edition. Cary, NC, USA: SAS Institute; Dec 26, 2001.
37. Pelletier G, Panneton L. Prevalence and risk factors for prescribing potentially inappropriate medications (PIM) to individuals aged 65 and over. Rev Gériatr 2017 Oct;8:455-463. [doi: 10.21019/pharmacotherapyfirst.meduse_overview]
38. Rushmer RK, Hunter DJ, Steven A. Using interactive workshops to prompt knowledge exchange: a realist evaluation of a knowledge to action initiative. Public Health 2014 Jun;128(6):552-560. [doi: 10.1016/j.puhe.2014.03.012] [Medline: 24854761]
39. Freedman MR, Nickell A. Impact of after-school nutrition workshops in a public library setting. J Nutr Educ Behav 2010;42(3):192-196. [doi: 10.1016/j.jneb.2009.07.003] [Medline: 20434074]
40. Elliott DS, Mihalic S. Issues in disseminating and replicating effective prevention programs. Prev Sci 2004 Mar;5(1):47-53. [doi: 10.1023/b:prev.0000013981.28071.52] [Medline: 15058912]
41. Anyon Y, Roscoe J, Bender K, Kennedy H, Dechants J, Begun S, et al. Reconciling adaptation and fidelity: implications for scaling up high quality youth programs. J Prim Prev 2019 Feb;40(1):35-49. [doi: 10.1007/s10935-019-00535-6] [Medline: 30659405]

XSL•FO
RenderX

42. Milat AJ, Newson R, King L, Rissel C, Wolfenden L, Bauman A, et al. A guide to scaling up population health interventions. Public Health Res Pract 2016 Jan 28;26(1):e2611604 [FREE Full text] [doi: 10.17061/phrp2611604] [Medline: 26863167]
43. Ben Charif A, Zomahoun HT, LeBlanc A, Langlois L, Wolfenden L, Yoong SL, et al. Effective strategies for scaling up evidence-based practices in primary care: a systematic review. Implement Sci 2017 Nov 22;12(1):139 [FREE Full text] [doi: 10.1186/s13012-017-0672-y] [Medline: 29166911]
44. Bennett S, Mahmood SS, Edward A, Tetui M, Ekirapa-Kiracho E. Strengthening scaling up through learning from implementation: comparing experiences from Afghanistan, Bangladesh and Uganda. Health Res Policy Syst 2017 Dec 28;15(Suppl 2):108 [FREE Full text] [doi: 10.1186/s12961-017-0270-0] [Medline: 29297353]
45. Gogovor A, Zomahoun HT, Ben Charif A, McLean RK, Moher D, Milat A, et al. Essential items for reporting of scaling studies of health interventions (SUCCEED): protocol for a systematic review and Delphi process. Syst Rev 2020 Jan 11;9(1):11 [FREE Full text] [doi: 10.1186/s13643-019-1258-3] [Medline: 31926555]
46. Sharma G. Pros and cons of different sampling techniques. Int J Appl Res 2017;3(7):749-752.

Abbreviations

IKT: integrated knowledge translation
KT: implementation science or knowledge translation
SPOR-SUPPORT: Strategy for Patient-Oriented Research Support for People and Patient-Oriented Research and Trials

Edited by J Wang; submitted 27.04.22; peer-reviewed by A McKinnon; comments to author 18.05.22; revised version received 31.05.22; accepted 12.06.22; published 19.08.22.

Please cite as:
Massougbodji J, Zomahoun HTV, Adisso EL, Sawadogo J, Borde V, Cameron C, Moisan H, Paquette JS, Akbaraly Z, Châteauneuf LK, David G, Bibliothèques2 Group, Légaré F
Scaling Up Citizen Workshops in Public Libraries to Disseminate and Discuss Primary Care Research Results: Quasi-Experimental Study
JMIR Aging 2022;5(3):e39016
URL: https://aging.jmir.org/2022/3/e39016
doi:10.2196/39016
PMID:35690963

©José Massougbodji, Hervé Tchala Vignon Zomahoun, Evehouenou Lionel Adisso, Jasmine Sawadogo, Valérie Borde, Cynthia Cameron, Hélène Moisan, Jean-Sébastien Paquette, Zamzam Akbaraly, Lëa-Kim Châteauneuf, Geneviève David, Bibliothèques2 Group, France Légaré. Originally published in JMIR Aging (https://aging.jmir.org), 19.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Smartwatch Step-Counting App for Older Adults: Development and Evaluation Study

George Boateng[1], MS; Curtis L Petersen[2], MPH, MS, PhD; David Kotz[2], PhD; Karen L Fortuna[2], PhD; Rebecca Masutani[3], MD; John A Batsis[2,4,5,6], MD

[1]Department of Management, Technology and Economics, ETH Zurich, Zurich, Switzerland

[2]Geisel School of Medicine, Dartmouth College, Hanover, NH, United States

[3]Division of Geriatrics and Palliative Care, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[4]Division of Geriatric Medicine, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[5]Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[6]The Dartmouth Institute for Health Policy & Clinical Practice, Geisel School of Medicine at Dartmouth, Hanover, NH, United States

Corresponding Author:
John A Batsis, MD
Division of Geriatric Medicine
School of Medicine
University of North Carolina at Chapel Hill
5017 Old Clinic Building
Chapel Hill, NC, 27599
United States
Phone: 1 919 843 4096
Email: john.batsis@gmail.com

Abstract

Background: Older adults who engage in physical activity can reduce their risk of mobility impairment and disability. Short amounts of walking can improve quality of life, physical function, and cardiovascular health. Various programs have been implemented to encourage older adults to engage in physical activity, but sustaining their motivation continues to be a challenge. Ubiquitous devices, such as mobile phones and smartwatches, coupled with machine-learning algorithms, can potentially encourage older adults to be more physically active. Current algorithms that are deployed in consumer devices (eg, Fitbit) are proprietary, often are not tailored to the movements of older adults, and have been shown to be inaccurate in clinical settings. Step-counting algorithms have been developed for smartwatches, but only using data from younger adults and, often, were only validated in controlled laboratory settings.

Objective: We sought to develop and validate a smartwatch step-counting app for older adults and evaluate the algorithm in free-living settings over a long period of time.

Methods: We developed and evaluated a step-counting app for older adults on an open-source wrist-worn device (Amulet). The app includes algorithms to infer the level of physical activity and to count steps. We validated the step-counting algorithm in the lab (counting steps from a video recording, n=20) and in free-living conditions—one 2-day field study (n=6) and two 12-week field studies (using the Fitbit as ground truth, n=16). During app system development, we evaluated 4 walking patterns: normal, fast, up and down a staircase, and intermittent speed. For the field studies, we evaluated 5 different cut-off values for the algorithm, using correlation and error rate as the evaluation metrics.

Results: The step-counting algorithm performed well. In the lab study, for normal walking (R^2=0.5), there was a stronger correlation between the Amulet steps and the video-validated steps; for all activities, the Amulet's count was on average 3.2 (2.1%) steps lower (SD 25.9) than the video-validated count. For the 2-day field study, the best parameter settings led to an association between Amulet and Fitbit (R^2=0.989) and 3.1% (SD 25.1) steps lower than Fitbit, respectively. For the 12-week field study, the best parameter setting led to an R^2 value of 0.669.

Conclusions: Our findings demonstrate the importance of an iterative process in algorithm development before field-based deployment. This work highlights various challenges and insights involved in developing and validating monitoring systems in real-world settings. Nonetheless, our step-counting app for older adults had good performance relative to the ground truth (a commercial Fitbit step counter). Our app could potentially be used to help improve physical activity among older adults.

XSL•FO
RenderX

(JMIR Aging 2022;5(3):e33845) doi:10.2196/33845

KEYWORDS

step tracking; step counting; pedometer; wearable; smartwatch; older adults; physical activity; machine learning; walking; mHealth; mobile health; mobile app; mobile application; app; uHealth

Introduction

Older adults are faced with an increased risk of developing multiple comorbid medical conditions, social isolation, and reduced physical function, which can lead to an increased risk of disability [1]. An inability to engage in activities of daily living may increase mortality risk and premature nursing home placement [2]. Hence, it is critical to encourage older adults with multimorbidity to engage in interventions that improve health, including physical activity. In fact, simple community-based walking programs and resistance-based programs [3] have effectively demonstrated reductions in the long-term risk of disability [4]. Even short bouts of walking can improve quality of life, physical function, and cardiovascular fitness in older adults [5].

Traditional consumer-based health-promoting endeavors, such as Silver Sneakers [6], have been scaled and widely disseminated across the United States. Randomized control trials have also shown the short- and long-term benefits of physical activity. However, sustained motivation continues to be a challenge for many individuals. Simple encouragement by clinicians may enhance engagement [7]. Yet, a study of accelerometry data demonstrated that only 7.6% of older adults aged 65 years meet Physical Activity Guidelines for Americans [8]. These pragmatic challenges demonstrate the need to overcome the barriers of traditional health promotion to enhance self-efficacy and behavioral change.

Older adults are the fastest-growing group of technology users; one survey suggested that 61% of older adults use smartphones [9], an increase from 23% in 2013 [10]. In fact, over 50% of older adults use social media [9]. Remote monitoring using fitness devices has now become ubiquitous in many countries where technology is readily available. In both consumer-based and academic-focused trials, it continues to be challenging to find a balance between clinical accuracy and ease of use. Current algorithms in consumer devices (such as Fitbit) are proprietary and often are not tailored to the group being evaluated in a clinical setting, such as older adults in free-living conditions, and have been shown to have wide error rates in such contexts [11,12]. A few smartwatch-based step-counting algorithms have been developed using data from young adults and validated in controlled settings only [13,14]. Matthies et al [15] developed a smartwatch step-counting app for older adults who use a walking frame equipped with wheels, which was evaluated outdoors, but only in a controlled setting, with 5 older adults. To the best of our knowledge, a smartwatch step-counting app for older adults has not been developed and validated in free-living settings over a long period of time with a large sample.

We previously developed GeriActive, an app that measures the daily activity levels (low, moderate, or vigorous) of older adults [16]. We aimed to develop and validate a smartwatch-based step-counting algorithm for older adults that runs as an app on the Amulet device. The Amulet is an open-source wrist-worn device that has been used for various mobile health studies, such as stress and physical activity monitoring [17,18]. The Amulet platform enables developers to write energy- and memory-efficient apps.

Methods

Study Overview

We validated the step-counting algorithm with older adults in the lab (using videorecordings as ground truth) and in 2 free-living studies (using the Fitbit as ground truth) lasting 2 and 12 days.

Overview of Step Counter App

Our step-counting Amulet app estimates the number of steps taken over the course of a day and displays the information on the Amulet screen, similar to the functionality of a pedometer or of other wearable devices (Figure 1). The app continuously estimates the number of steps for each 5-second window, updating the count viewed by the user. It uses a 2-step process: machine learning is used to determine if physical activity occurred in the most recent 5 seconds of data, and then, the number of steps is estimated by counting the number of peaks.

XSL•FO
RenderX

Figure 1. The Step Counter App with the step count displayed at the bottom.



Activity-Detection Algorithm

We used a linear support vector machine that classifies each 5-second window of accelerometer data into low, moderate, or vigorous activity [16]. We trained the algorithm on data collected from older adults who performed various physical activities: sitting, standing, lying down, walking, and running [16,19]. Our evaluation of the algorithm produced an accuracy of 91.7% using leave-one-subject-out cross-validation. If the output of the algorithm is moderate or vigorous, the step-detection algorithm is run to determine the number of steps in the 5-second window. This 2-phase approach was necessary to reduce false positives by preventing various actions such as random hand motions from being counted as steps.

Step-Detection Algorithm

The step-detection algorithm estimates the number of steps in 5 seconds of acceleration data. The algorithm uses the magnitude of the acceleration of the 3-axis accelerometer of the Amulet. It is a 3-stage process consisting of zero-meaning, filtering, and peak detection, using an approach similar to that described in [13]. First, to ensure the data have a mean of zero, for each sample, we subtract the average of the preceding 20 data samples. Subsequently, a moving-average filter is applied, that is, each sample is replaced with the average of the 3 preceding samples. Finally, peaks in the filtered signal are identified by checking for change of slope. If the slope changes from positive to negative, and the peak value is above a certain threshold, then the peak is counted as a step. The cut-off threshold was initially empirically determined and then tuned.

Ethics

Studies were approved by the Committee for the Protection of Human Subjects at Dartmouth College and the Dartmouth-Hitchcock Institutional Review Board (28905). All participants provided signed informed consent.

Participants

Participants were recruited through the Center for Health and Aging at Dartmouth and primary care practices at Dartmouth-Hitchcock using flyers, listservs, and posters. This was a convenience sample; our results may not necessarily be applicable to other groups.

Laboratory Study

Data for the development and evaluation of the step-detection algorithm were collected at our Dartmouth campus laboratory. Older adults (n=20, age ≥65 years) were asked to perform different types of walking (normal, fast, up and down a staircase, and intermittent) while wearing an Amulet. The Amulet collected raw acceleration data at a frequency of 20 Hz and logged the magnitude, which we later used to develop the step algorithm. The participants were videotaped. The video was independently reviewed to count steps by 2 individuals independently (JAB, RKM) and any differences were later reconciled. We used these data for the preliminary development of the step-detection algorithm and evaluated the algorithm using the error rate (the percentage difference between the algorithm's estimated step count and the ground-truth step count measured from the video).

2-Day Field Study

We conducted a 2-day field study in which older adults (n=7, age ≥65 years) wore an Amulet device (running our step counter app) and a Fitbit Flex 2 device (Fitbit Inc) on the same wrist for 2 days. We compared each participant's step count estimated by their Amulet (exploring 5 different cut-off values) with their step count reported by the Fitbit (downloaded from the

XSL•FO
RenderX

individual's research-based Fitbit account). The error rate between Fitbit's steps and Amulet's steps for each of the 5 peak cut-off values was computed.

12-Week Field Study

We conducted a field study (2 cohorts, 12 weeks each) to test the step-detection algorithm with the target population—older adults with obesity. This study was a subset of a larger study that evaluated the impact of a combined weight loss and exercise intervention in older adults with obesity [20]. The goal of this analysis was to compare the Amulet's step-count estimate with the Fitbit's step count over a long period in real-world settings. Participants (Table 1) from both rounds wore an Amulet and Fitbit on the same wrist for 12 weeks. The Amulet logged the summary steps on an SD memory card hourly and at the end of each day. A research assistant copied the data from each participant's SD card on a biweekly basis. The Fitbit logged a summary of each day's step count (computed by a proprietary algorithm) to the Fitbit app on the participant's Android tablet, which uploaded the data to the Fitbit servers; we subsequently used the Fitbit research API to download participants' data. After 10 weeks of monitoring data from the first cohort, we modified the step-detection peak cut-off value to minimize the error rate relative to ground-truth step count data from the Fitbit.

Table 1. Participant characteristics.

Characteristic	Value (n=16)
Age (years)	
Mean (SD)	74.1 (5.6)
Range	66-87
Sex, n (%)	
Male	4 (25)
Female	12 (75)
Marital status, n (%)	
Married	7 (44)
Divorced	8 (50)
Widowed	1 (6)
Smoking history, n (%)	
None	13 (81)
Formerly smoked	3 (19)
Education, n (%)	
High school	2 (12)
Some college	5 (31)
College degree	3 (19)
Postcollege degree	6 (38)
Weight (kg), mean (SD)	97.06 (18.2)
Body mass index (kg/m^2), mean (SD)	36.8 (4.9)
Multimorbidity, n (%)	14 (87)
Comorbidities, n (%)	
Diabetes	2 (12)
Fibromyalgia	1 (6)
High cholesterol	6 (38)
Hypertension	9 (56)
Osteoarthritis	6 (38)

Statistical Analysis

We calculated the mean with standard deviation for continuous measures and count with percentage for categorical measures. Although participants were instructed to wear the Amulet and Fitbit simultaneously, not all participants did so the entire time. We thus limited the data to days for which it was safe to assume that they were worn for nearly the same amount of time. The Amulet was able to detect if it was worn each hour, so we considered the Amulet to be worn if the Amulet was worn for 75% of a 15-hour day (675 minutes); the Fitbit only reported daily step count, so we considered the Fitbit to be worn if the

XSL•FO
RenderX

step count was greater than 100. We selected these parameters from an understanding of the distribution of wear time over the course of a 24-hour day and the distribution of steps expected for this population per day [21]. Comparisons between Amulet and Fitbit were limited to data from days for which both were worn. We conducted bivariate linear regression to compare the association between Amulet and video-counted steps during various activities (laboratory study) and between Amulet and Fitbit steps (2-day and 12-week field studies). We compared Amulet steps to Fitbit steps using percentage difference (difference between Amulet and Fitbit steps divided by Fitbit steps).We used Bland-Altman plots to compare the agreement between Amulet and Fitbit steps. For our analysis, we defined significance as $P<.05$.

Results

Laboratory Study

There was a strong linear association when participants walked normally (Figure 2; Table 2). For normal walking, the Amulet step-detection algorithm undermeasured the number of steps taken by an average of 6.7 steps (SD 32.6). Across all activities, the Amulet was on average 3.2 steps lower (SD 25.9) or 2.1% (SD 31.9%) lower than video-validated steps (Figure 2). The distribution was slightly positively skewed.

Figure 2. Association between Amulet-estimated steps and video-validated steps.

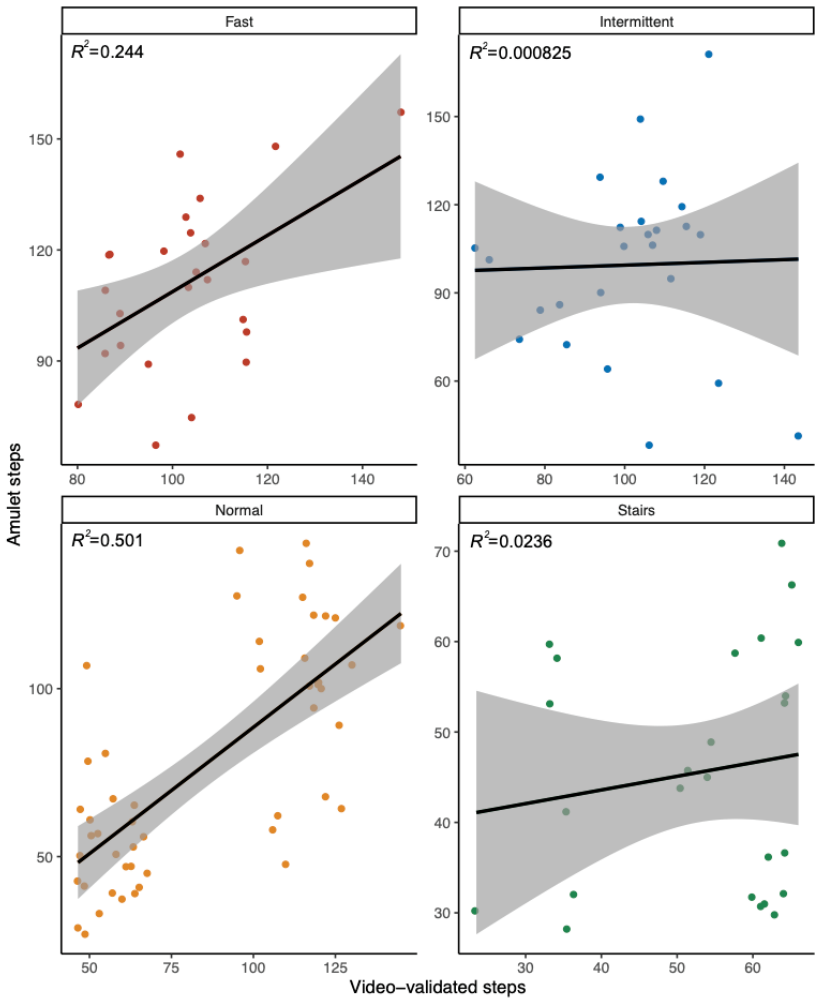


Table 2. Step count for different walking activities.

Activity	Video-validated, n (%)	Amulet, n (%)	Percentage error
Fast	102.62 (14.7)	110.72 (22.7)	8.53 (20.02)
Intermittent	101.02 (18.7)	99.48 (30.8)	1.71 (33.28)
Normal	84.9 (31.9)	77.14 (33.9)	–6.68 (32.55)
Stairs	52.76 (13.4)	45.52 (13.2)	–7.55 (36.97)

XSL•FO
RenderX

2-Day Field Study

We discarded 1 participant's data because the data indicated the devices had not been worn much (step counts were less than 350 per day). The associations between each participant's daily step counts reported by the Amulet and Fitbit were high for all cut-off thresholds (Figure 3). The third cut-off threshold had the highest association between Amulet and Fitbit steps (R^2=0.989). Cut-off threshold number 2 had the smallest mean percentage difference between Amulet and Fitbit (–3.1%, SD 25.1) of all cut-off thresholds (threshold 1: mean 15.27%, SD 33.19; threshold 3: mean –10.77, SD 23.43; threshold 4: mean 7.18, SD 29.28; threshold 5: mean –4.11, SD 24.94) (Figure 4).

Figure 3. Association between Amulet and Fitbit steps for different cut-off values: 2-day field study.

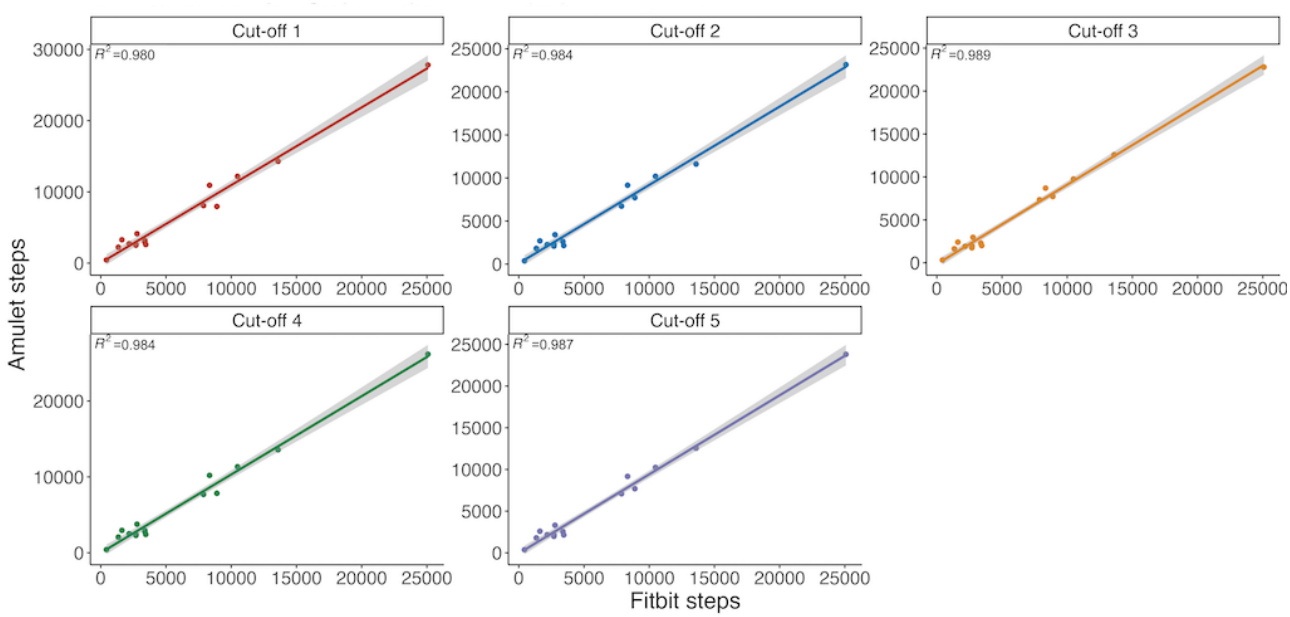
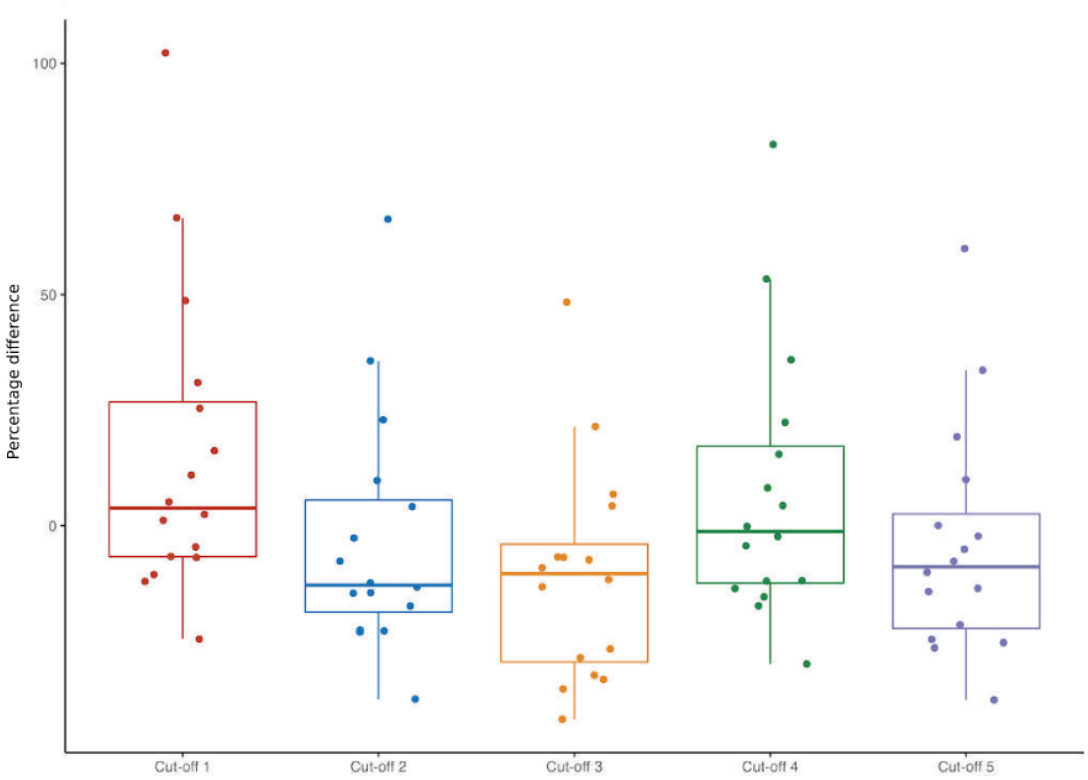


Figure 4. Distribution of percentage difference between Amulet and Fitbit steps by cut-off threshold.

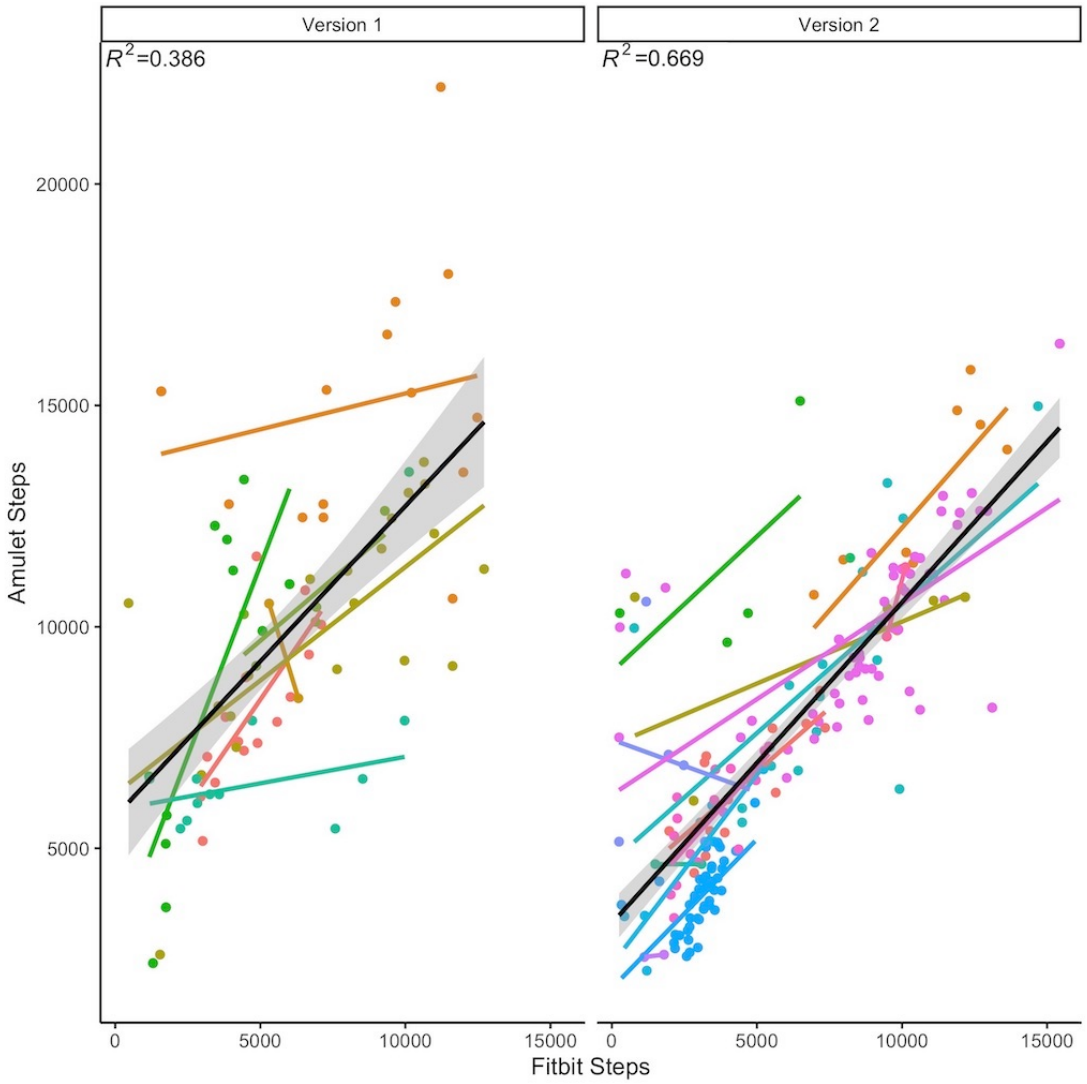


XSL•FO
RenderX

12-Week Field Study Results

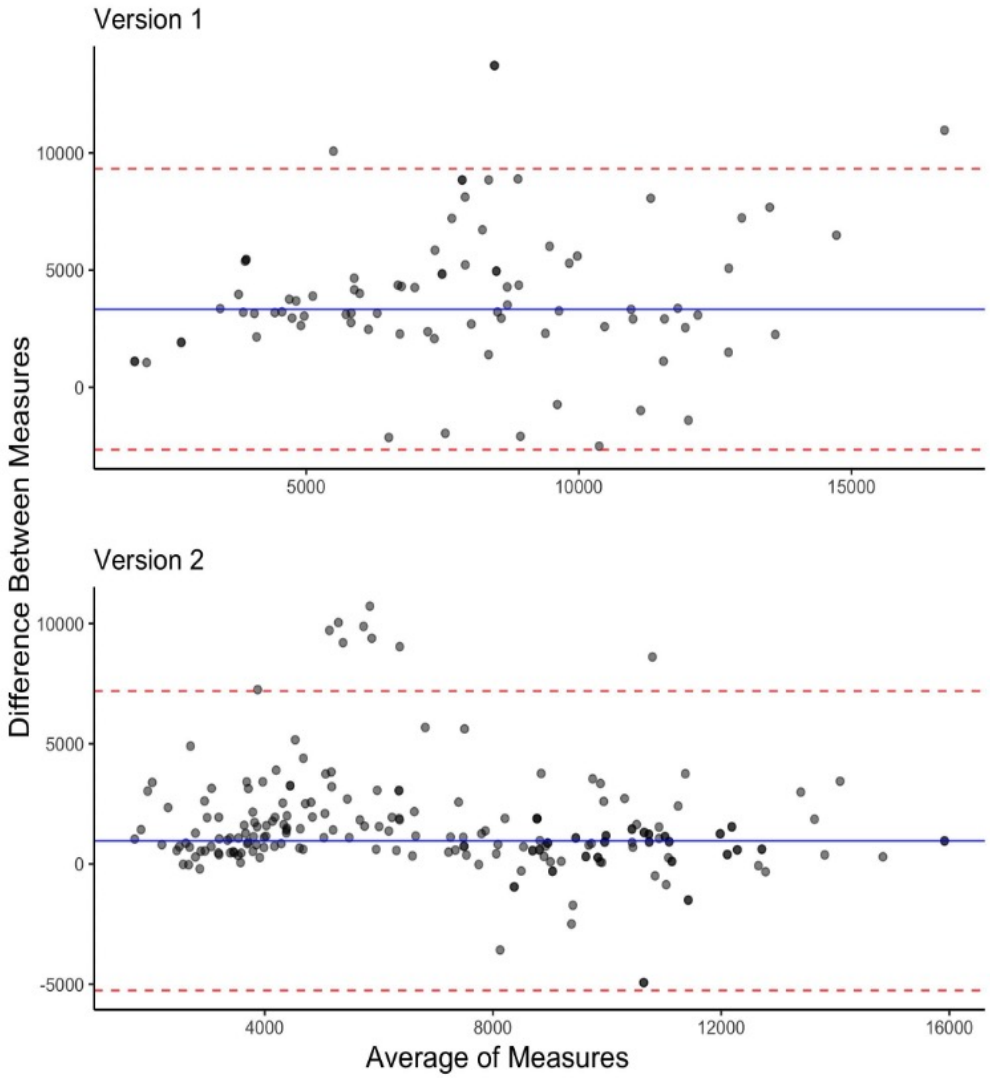
Across both 12-day field studies, there were 297 participant-days for which both the Fitbit and Amulet had been worn. Cohort 2 used the modified app. For the first cut-off threshold (version 1), we recorded 86 participant-days; the average Fitbit step count was 5797 steps (SD 3296), and the average Amulet step count was 9780 steps (SD 3719). For the second cut-off threshold (version 2), we recorded 211 participant-days; the average Fitbit step count was 6415 per day (SD 3751), and the average Amulet step count was 7956 per day (SD 3324). The association between Fitbit steps and Amulet steps improved (first cut-off threshold: R^2=0.386; second cut-off threshold: R^2=0.669) (Figure 5). There was improved agreement between both measures with the second cut-off threshold (Figure 6). There were 9 observations by 5 unique participants with differences 2 standard deviations higher than the combined difference mean. These observations had an average Fitbit and Amulet step count of 1373 (SD 1988) and 10,689 (SD 1971), which suggested that participants may have taken their Fitbits off at some point during the day before removing their Amulet.

Figure 5. Association between Amulet and Fitbit steps by algorithm version. Each line is a linear regression for each participant, colored separately, with the overall linear regression in black. Version 1 represents the app used during the first 10 weeks of cohort 1, and version 2 represents the app used during the final 2 weeks of cohort 1 and all 12 weeks of cohort 2.



XSL•FO
RenderX

Figure 6. Bland-Altman plots of Amulet and Fitbit step measures by algorithm version. The blue line represents the mean difference between Amulet and Fitbit steps per day. Each red line represents 2 SD of the difference. Version 1 represents the app used during the first 10 weeks of cohort 1, and version 2 represents the app used during the final 2 weeks of cohort 1 and all 12 weeks of cohort 2.



Discussion

We found that an open-source platform and algorithm developed for older adults can capture daily step counts with reasonable accuracy. Our findings demonstrate the importance of an iterative process in algorithm development before field deployment. First, our lab-based data provided confidence in the algorithm's step estimates, making a case for a real-world evaluation. We then tested the algorithm in a 2-day field study before full-scale deployment. The step estimates from our algorithm were highly correlated with the step counts from the Fitbit for all peak cut-off thresholds, with low error rates. These results provide evidence that the algorithm worked well in free-living conditions, albeit for a short duration. We used the cut-off threshold with the lowest error rate for the subsequent field studies conducted over longer periods.

As with any user study, there are differences between field-based conditions and those in laboratory settings or short-duration studies. The poorer results in the 12-week study (cohort 1) may have been a result of the differences between the populations used for developing and evaluating the algorithm (older adults vs older obese adults). Older adults with obesity have a higher degree of comorbid conditions [22], along with differences in stride length, cadence, and gait [23], which may impact either algorithm. Additionally, the longer time period (12 weeks vs 2 days) could have allowed the occurrence of a greater number of confounding situations, such as both devices not being worn at all times or one device being off while the other was on. These results make a strong case for developing and refining algorithms with data from the target population and evaluating algorithms in the conditions for which they were designed.

Based on our observations in the first 10 weeks of cohort 1, we modified the peak cut-off threshold to minimize the error rate and evaluated version 2 of the algorithm in a study of the same duration with different participants from the same target population (cohort 2). Version 2 exhibited better performance in terms of correlations and error rates.

XSL•FO
RenderX

One limitation of this work is that we used the Fitbit as ground truth. Given that the Fitbit device (and its proprietary algorithm) was not specifically developed for this population (older adults with obesity), it is difficult to say whether our algorithm performed better or worse than Fitbit's algorithm relative to the ground truth. The ideal ground truth would have been to use video recording, as we did in the lab study, but videorecording is not feasible for field studies due to privacy limitations. Hence, a device such as the Fitbit was the best compromise.

Although Fitbit outputs have been shown to have high correlations with those of other step-counting devices when used with older adults in free-living settings, results have varied depending on the specific Fitbit device used, device placement (wrist vs waist), and the comparison device used [24-27]. We hypothesized that the Fitbit would underestimate the steps of older adults in comparison to the Amulet, because the Fitbit was developed using data from younger adults, and older adults move more slowly [28]. Thus, we expect the overall true step count to be higher than Fitbit's estimate. Hence, we settled for the case where our algorithm overestimated the steps but was highly correlated with Fitbit's estimate. In addition, the Fitbit data were captured daily, whereas Amulet data were captured hourly. Had hourly data been available from both, we could have performed a fine-grained comparison between the algorithms. Finally, it was not possible to get a good sense of wear time from the Fitbit as we did in the Amulet. We could only use the Amulet's wear time estimate and a minimum Fitbit daily step count of 100 steps as the threshold for being worn.

Because we were evaluating predominantly intraperson variability (ie, the purpose was not to evaluate the impact of any intervention), we did not measure baseline characteristics of the participants (eg, disease regarding walking behavior, such as Parkinson disease, musculoskeletal disorders). Future studies should determine whether such characteristics could have an impact on our results.

The use of an open-source system, such as the Amulet, highlights researchers' ability to develop algorithms that are tailored and trained for a target population such as older adults. With the constant iteration of consumer devices and algorithms, it is difficult to ensure precision and accuracy for groups that need to be more physically active, such as older adults. Hence, it is important to develop and examine products that can meet their needs. Providing imprecise or inaccurate information on physical activity could undermine the motivation of this population to increase their physical activity. We recommend that future work demonstrate validity of algorithms in these populations and identify situations where data collection can be the most clinically relevant and actionable.

Our step-count algorithm performed well in comparison with Fitbit, with high correlations and low error rates. Overall, this work highlights various challenges and insights involved in developing and validating monitoring systems in real-world settings.

Acknowledgments

This work was supported in part by the National Institute on Aging and Office of Dietary Supplements of the National Institutes of Health (K23AG051681 awarded to JAB). Support was also provided by the Dartmouth Health Promotion and Disease Prevention Research Center (U48DP005018) from the Centers for Disease Control and Prevention, and the Dartmouth Clinical and Translational Science Institute (UL1TR001086) from the National Center for Advancing Translational Sciences of the National Institutes of Health. CLP was supported by the Burroughs-Wellcome Fund: Big Data in the Life Sciences at Dartmouth. DK and GB were supported by National Science Foundation awards (CNS-1314281 and CNS-1619970) and by the National Institutes of Health/National Institute on Drug Abuse (P30DA029926). KLF was supported by the National Institute of Mental Health (K01MH117496). The content is solely the responsibility of the authors and does not necessarily represent the official views of, or represent the official position of, any of the sponsors.

Conflicts of Interest

KLF provides consulting through Social Wellness.

References

1. Rubenstein LZ, Powers CM, MacLean CH. Quality indicators for the management and prevention of falls and mobility problems in vulnerable elders. Ann Intern Med 2001 Oct 16;135(8 Pt 2):686-693 [FREE Full text] [doi: 10.7326/0003-4819-135-8_part_2-200110161-00007] [Medline: 11601951]
2. Tinetti ME, Williams CS. Falls, injuries due to falls, and the risk of admission to a nursing home. N Engl J Med 1997 Oct 30;337(18):1279-1284. [doi: 10.1056/NEJM199710303371806] [Medline: 9345078]
3. Fisher KJ, Li F. A community-based walking trial to improve neighborhood quality of life in older adults: a multilevel analysis. Ann Behav Med 2004 Dec;28(3):186-194. [doi: 10.1207/s15324796abm2803_7] [Medline: 15576257]
4. Paterson DH, Warburton DE. Physical activity and functional limitations in older adults: a systematic review related to Canada's Physical Activity Guidelines. Int J Behav Nutr Phys Act 2010 May 11;7:38 [FREE Full text] [doi: 10.1186/1479-5868-7-38] [Medline: 20459782]

XSL•FO
RenderX

5. Murphy M, Nevill A, Neville C, Biddle S, Hardman A. Accumulating brisk walking for fitness, cardiovascular risk, and psychological health. Med Sci Sports Exerc 2002 Sep;34(9):1468-1474. [doi: 10.1097/00005768-200209000-00011] [Medline: 12218740]
6. Greenwood-Hickman MA, Rosenberg DE, Phelan EA, Fitzpatrick AL. Participation in older adult physical activity programs and risk for falls requiring medical care, Washington state, 2005-2011. Prev Chronic Dis 2015 Jun 11;12:E90 [FREE Full text] [doi: 10.5888/pcd12.140574] [Medline: 26068411]
7. So C, Pierluissi E. Attitudes and expectations regarding exercise in the hospital of hospitalized older adults: a qualitative study. J Am Geriatr Soc 2012 Apr;60(4):713-718. [doi: 10.1111/j.1532-5415.2012.03900.x] [Medline: 22429118]
8. Tucker JM, Welk GJ, Beyler NK. Physical activity in U.S.: adults compliance with the Physical Activity Guidelines for Americans. Am J Prev Med 2011 Apr;40(4):454-461. [doi: 10.1016/j.amepre.2010.12.016] [Medline: 21406280]
9. Social media use in 2021. Pew Research Center. URL: https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/ [accessed 2021-09-01]
10. Perrin A, Anderson M. Tech adoption climbs among older adults. Pew Research Center. 2017. URL: https://www.pewresearch.org/internet/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2021-09-01]
11. Singh AK, Farmer C, Van Den Berg ML, Killington M, Barr CJ. Accuracy of the FitBit at walking speeds and cadences relevant to clinical rehabilitation populations. Disabil Health J 2016 Apr;9(2):320-323. [doi: 10.1016/j.dhjo.2015.10.011] [Medline: 26905972]
12. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit devices: systematic review and narrative syntheses of quantitative data. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
13. Cho Y, Cho H, Kyung C. Design and implementation of practical step detection algorithm for wrist-worn devices. IEEE Sensors J 2016 Aug 25:1. [doi: 10.1109/jsen.2016.2603163]
14. Genovese V, Mannini A, Sabatini AM. A smartwatch step counter for slow and intermittent ambulation. IEEE Access 2017;5:13028-13037. [doi: 10.1109/access.2017.2702066]
15. Matthies D, Haescher M, Nanayakkara S, Bieber G. Step detection for Rollator users with smartwatches. 2018 Presented at: Symposium on Spatial User Interaction; October 13, 2018; Berlin, Germany p. 163-167. [doi: 10.1145/3267782.3267784]
16. Boateng G, Batsis JA, Proctor P, Halter R, Kotz D. GeriActive: wearable app for monitoring and encouraging physical activity among older adults. 2018 Presented at: 15th International Conference on Wearable and Implantable Body Sensor Networks; March 3, 2018; Las Vegas, Nevada p. 46-49. [doi: 10.1109/BSN.2018.8329655]
17. Hester J, Peters T, Yun T, Peterson R, Skinner J, Golla B. Amulet: an energy-efficient, multi-application wearable platform. 2016 Presented at: ACM Conference on Embedded Network Sensor Systems; November 14, 2016; Stanford, California p. 216-229.
18. Boateng G, Motti V, Mishra V, Batsis J, Hester J, Kotz D. Experience: design, development and evaluation of a wearable device for mhealth applications. 2019 Presented at: 25th Annual International Conference on Mobile Computing and Networking; October 11,2019; Los Cabos, Mexico URL: https://doi.org/10.1145/3300061.3345432 [doi: 10.1145/3300061.3345432]
19. Boateng G, Batsis JA, Halter R, Kotz D. ActivityAware: an app for real-time daily activity level monitoring on the Amulet wrist-worn device. 2017 Presented at: International Conference on Pervasive Computing and Communications Workshops; March 2017; Hawaii. [doi: 10.1109/PERCOMW.2017.7917601]
20. Batsis JA, Petersen CL, Clark MM, Cook SB, Lopez-Jimenez F, Al-Nimr RI, et al. A weight loss intervention augmented by a wearable device in rural older adults with obesity: a feasibility study. J Gerontol A Biol Sci Med Sci 2021 Jan 01;76(1):95-100 [FREE Full text] [doi: 10.1093/gerona/glaa115] [Medline: 32384144]
21. Tudor-Locke C, Hart TL, Washington TL. Expected values for pedometer-determined physical activity in older populations. Int J Behav Nutr Phys Act 2009 Aug 25;6:59 [FREE Full text] [doi: 10.1186/1479-5868-6-59] [Medline: 19706192]
22. Lynch DH, Petersen CL, Fanous MM, Spangler HB, Kahkoska AR, Jimenez D, et al. The relationship between multimorbidity, obesity and functional impairment in older adults. J Am Geriatr Soc 2022 May;70(5):1442-1449. [doi: 10.1111/jgs.17683] [Medline: 35113453]
23. Ko SU, Stenholm S, Ferrucci L. Characteristic gait patterns in older adults with obesity--results from the Baltimore Longitudinal Study of Aging. J Biomech 2010 Apr 19;43(6):1104-1110 [FREE Full text] [doi: 10.1016/j.jbiomech.2009.12.004] [Medline: 20080238]
24. McVeigh J, Ellis J, Ross C, Tang K, Wan P, Halse RE. Convergent validity of the Fitbit Charge 2 to measure sedentary behavior and physical activity in overweight and obese adults. J Meas Phys Behav 2021 Jan;4(1):39-46. [doi: 10.1123/jmpb.2020-0014]
25. Farina N, Lowry RG. The validity of consumer-level activity monitors in healthy older adults in free-living conditions. J Aging Phys Act 2018 Jan 01;26(1):128-135. [doi: 10.1123/japa.2016-0344] [Medline: 28595019]
26. Collins JE, Yang HY, Trentadue TP, Gong Y, Losina E. Validation of the Fitbit Charge 2 compared to the ActiGraph GT3X+ in older adults with knee osteoarthritis in free-living conditions. PLoS One 2019 Jan 30;14(1):e0211231 [FREE Full text] [doi: 10.1371/journal.pone.0211231] [Medline: 30699159]

XSL•FO
RenderX

27. Paul SS, Tiedemann A, Hassett LM, Ramsay E, Kirkham C, Chagpar S, et al. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. BMJ Open Sport Exerc Med 2015:e000013 [FREE Full text] [doi: 10.1136/bmjsem-2015-000013] [Medline: 27900119]
28. Mendes J, Borges N, Santos A, Padrão P, Moreira P, Afonso C, et al. Nutritional status and gait speed in a nationwide population-based sample of older adults. Sci Rep 2018 Mar 09;8(1):4227 [FREE Full text] [doi: 10.1038/s41598-018-22584-3] [Medline: 29523852]

Abbreviations

API: application programming interface

Edited by J Wang; submitted 26.09.21; peer-reviewed by U Bork, S Rostam Niakan Kalhori; comments to author 28.10.21; revised version received 13.12.21; accepted 07.02.22; published 10.08.22.

Please cite as:
Boateng G, Petersen CL, Kotz D, Fortuna KL, Masutani R, Batsis JA
A Smartwatch Step-Counting App for Older Adults: Development and Evaluation Study
JMIR Aging 2022;5(3):e33845
URL: https://aging.jmir.org/2022/3/e33845
doi:10.2196/33845
PMID:35947445

©George Boateng, Curtis L Petersen, David Kotz, Karen L Fortuna, Rebecca Masutani, John A Batsis. Originally published in JMIR Aging (https://aging.jmir.org), 10.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Validation of a Remote and Fully Automated Story Recall Task to Assess for Early Cognitive Impairment in Older Adults: Longitudinal Case-Control Observational Study

Caroline Skirrow[1], PhD; Marton Meszaros[1], MD; Udeepa Meepegama[1], MSc; Raphael Lenain[1], MSc; Kathryn V Papp[2,3], PhD; Jack Weston[1], PhD; Emil Fristed[1], MSc

[1]Novoic Ltd, London, United Kingdom

[2]Center for Alzheimer Research and Treatment, Department of Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

[3]Department of Neurology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Caroline Skirrow, PhD
Novoic Ltd
Wenlock Studios, Office G.05
50-52 Wharf Road, Islington
London, N1 7EU
United Kingdom
Phone: 44 7759 093006
Email: caroline@novoic.com

Abstract

Background: Story recall is a simple and sensitive cognitive test that is commonly used to measure changes in episodic memory function in early Alzheimer disease (AD). Recent advances in digital technology and natural language processing methods make this test a candidate for automated administration and scoring. Multiple parallel test stimuli are required for higher-frequency disease monitoring.

Objective: This study aims to develop and validate a remote and fully automated story recall task, suitable for longitudinal assessment, in a population of older adults with and without mild cognitive impairment (MCI) or mild AD.

Methods: The "Amyloid Prediction in Early Stage Alzheimer's disease" (AMYPRED) studies recruited participants in the United Kingdom (AMYPRED-UK: NCT04828122) and the United States (AMYPRED-US: NCT04928976). Participants were asked to complete optional daily self-administered assessments remotely on their smart devices over 7 to 8 days. Assessments included immediate and delayed recall of 3 stories from the Automatic Story Recall Task (ASRT), a test with multiple parallel stimuli (18 short stories and 18 long stories) balanced for key linguistic and discourse metrics. Verbal responses were recorded and securely transferred from participants' personal devices and automatically transcribed and scored using text similarity metrics between the source text and retelling to derive a generalized match score. Group differences in adherence and task performance were examined using logistic and linear mixed models, respectively. Correlational analysis examined parallel-forms reliability of ASRTs and convergent validity with cognitive tests (Logical Memory Test and Preclinical Alzheimer's Cognitive Composite with semantic processing). Acceptability and usability data were obtained using a remotely administered questionnaire.

Results: Of the 200 participants recruited in the AMYPRED studies, 151 (75.5%)—78 cognitively unimpaired (CU) and 73 MCI or mild AD—engaged in optional remote assessments. Adherence to daily assessment was moderate and did not decline over time but was higher in CU participants (ASRTs were completed each day by 73/106, 68.9% participants with MCI or mild AD and 78/94, 83% CU participants). Participants reported favorable task usability: infrequent technical problems, easy use of the app, and a broad interest in the tasks. Task performance improved modestly across the week and was better for immediate recall. The generalized match scores were lower in participants with MCI or mild AD (Cohen d=1.54). Parallel-forms reliability of ASRT stories was moderate to strong for immediate recall (mean rho 0.73, range 0.56-0.88) and delayed recall (mean rho=0.73, range=0.54-0.86). The ASRTs showed moderate convergent validity with established cognitive tests.

Conclusions: The unsupervised, self-administered ASRT task is sensitive to cognitive impairments in MCI and mild AD. The task showed good usability, high parallel-forms reliability, and high convergent validity with established cognitive tests. Remote,

XSL•FO
RenderX

low-cost, low-burden, and automatically scored speech assessments could support diagnostic screening, health care, and treatment monitoring.

(JMIR Aging 2022;5(3):e37090) doi:10.2196/37090

KEYWORDS

neurology; memory; episodic; speech; psychometrics; reliability; validity; aging; elder; older adult; Alzheimer disease; mild cognitive impairment; mobile apps; mobile health; mHealth; smartphone; cognition; cognitive decline; cognitive impairment; development; validation; recall; memory; story; stories; observational study; acceptability; usability; semantic; cognitive test; speech; linguistic; mobile phone

Introduction

With the first disease-modifying treatment for Alzheimer disease (AD) now available [1], there is an increased need for broader screening and improved monitoring of disease progression and treatment response. Cognitive assessments are currently some of the least invasive and most cost-effective measures available for screening for AD and related impairments. Furthermore, they are supported for use as endpoints of treatment efficacy early in AD by key regulatory bodies, including the US Food and Drug Administration [2] and the European Medicines Agency [3].

However, many cognitive assessments are lengthy, require trained personnel to administer and score, and offer few parallel test variants, making them susceptible to practice effects. More importantly, test performance is measurably influenced by a range of state factors such as sleep [4], exercise [5], mood [6], and stress [7]. This variation can lead to inaccurate impression of improvement or decline over time [8]. Higher-frequency sampling can generate more stable and reliable estimates of constructs of interest by controlling for state effects [9] and delineating short-term cognitive fluctuations from longer-term changes associated with treatment response and disease progression [8].

Story recall is a cognitive testing paradigm used to assess verbal episodic memory and is commonly used to track AD-related decline, often as a component of cognitive composite tests [10-14]. Story recall is impaired in Alzheimer dementia [15], shows variable differentiation of individuals with mild cognitive impairment (MCI) from those that are cognitively unimpaired (CU) [16], and predicts progression from MCI to Alzheimer dementia [17].

Most story recall tests are administered in person and scored manually, but research has shown that scoring can be fully automated using natural language processing technologies [18]. This suggests that story recall tests could be administered in clinic at a lower cost and with reduced clinician time burden. Moreover, these tests may be suitable for use in remote assessment, provided that they are properly developed and validated and that test administration can be automated.

Although remote digital assessments are not new, the COVID-19 pandemic accelerated the need to adopt remote or hybrid clinical assessment or research methods [19,20]. Alongside advances in technology and connectivity, this has led to a growing interest in the use of personal digital devices to collect clinically informative data. Beyond this, digital health technologies can enhance inclusivity, improving access for people who experience mobility problems or those with financial, geographical, or time restrictions [21]. The continued drive toward remote assessment may be particularly important in older adults who are at a substantially increased risk in the pandemic [22]. Although holding promise for improving convenience and access, there are concerns about whether digital assessment methods are particularly challenging in this population, particularly for those with dementia or milder forms of cognitive impairment [23].

This study describes the Automatic Story Recall Task (ASRT), a remote, self-administered, and automatically scored test developed for repeated cognitive assessment, opening up opportunities for more nuanced longitudinal data analysis. Test characteristics were examined in participants who were CU, and individuals with MCI or mild AD. Participants were assessed repeatedly over 1 week. This study examined (1) the acceptability of remote ASRT assessment, (2) adherence to daily ASRT assessments, (3) parallel-forms reliability, (4) convergent validity with cognitive and clinical assessments, (5) task performance characteristics, and (6) the effect of daily internal state factors.

Methods

Recruitment

Participants were recruited from November 2020 to August 2021 in the United Kingdom (London, Guildford, Plymouth, and Birmingham) and the United States (Santa Ana, California). Research participants were enrolled if they were CU or diagnosed with MCI in the previous 5 years. In the UK study, participants diagnosed with mild AD in the last 5 years were also included. MCI due to AD and mild AD diagnoses were made according to the National Institute on Aging–Alzheimer's Association core clinical criteria [24]. Participants were approached if they had undergone a prior amyloid beta positron emission tomography scan or a cerebrospinal fluid test (confirmed amyloid beta negative within 30 months or amyloid beta positive within 60 months). Eligibility was established by screening via video calls using a secure Zoom (Zoom Video Inc) link (UK study) or in-clinic assessment (US study), during which the Mini-Mental State Examination (MMSE) [25] was administered. For remote administration, no controls for potential environmental prompts to orientation questions (calendars, clocks, watches, etc) were implemented.

Inclusion criteria were as follows: age 50 to 85 years; MMSE raw score of 23 to 30 for participants with MCI or mild AD and 26 to 30 for CU; CU or clinical diagnosis of MCI or mild AD

XSL•FO
RenderX

made in the previous 5 years; English as a first language; availability of a study partner for Clinical Dementia Rating scale (CDR) [26] semistructured interview; and access and ability to use a smartphone running an operating system of Android 7 or above, or iOS 11 or above.

Exclusion criteria were as follows: current diagnosis of general anxiety disorder, recent (6 month) history of unstable psychiatric illness, history of stroke within the past 2 years, or a documented history of transient ischemic attack or unexplained loss of consciousness in the last 12 months. Participants treated with medications for symptoms related to AD were required to be stabilized on these medications for at least 8 weeks before study entry and throughout the study. Participants with a current diagnosis of major depressive disorder (United Kingdom) or those with a current or 2-year history of major depressive disorder (United States) were excluded.

Ethics Approval

This research was approved by the institutional review boards of the relevant research authorities (UK research ethics committee reference: 20/WM/0116; US Institutional Review Board reference: 8460-JGDuffy). Informed consent was obtained at the study site (United States) or electronically in accordance with the health research authority guidelines (United Kingdom). The studies are registered at ClinicalTrials.gov (NCT04828122 and NCT04928976).

Procedure

Clinical Assessments

Participants completed clinical assessments via a secure Zoom link (United Kingdom) or in clinic (United States), completed with a trained psychometrician. The tests reported in this study are described in detail below.

The Wechsler Logical Memory (LM) Test "Anna Thompson" story variant evaluated the free recall of a story according to 25 predefined information units (a metric quantifying the amount of information recalled [27]) immediately after presentation and after a 30-minute delay. Variants presented included the original Wechsler Memory Scale (WMS) text for the US sample [28] and the WMS 3rd edition text for the UK sample [29]. Paraphrased answers were accepted for both text variants, and scoring was completed manually according to the instructions and in alignment with the administration and scoring guidelines. The immediate and delayed recall scores were obtained.

Cognitive tests incorporated in the Preclinical Alzheimer's Cognitive Composite with semantic processing (PACC5) were administered. Tests were manually scored, and a mean *Z*-score was calculated as described previously [11]. The composite includes summary scores from five measures: (1) the MMSE [25], a global cognitive screening test; (2) LM Delayed Recall [28,29], a delayed story recall test; (3) Digit-Symbol Coding [30], a symbol substitution test; (4) the sum of free and total recall from the Free and Cued Selective Reminding Test [31], a multimodal associative memory test; and (5) category fluency (animals, vegetables, and fruits), a semantic memory test.

The CDR [26], a semistructured interview assessing the severity of cognitive symptoms of dementia, was completed with the participants and their study partners and scored based on the CDR–Sum of Boxes (CDR-SB) scales. The examiner was not blinded to the other assessments administered. In the US study, where participants had completed subtests of the PACC5 or CDR assessments within 1 month before the study visit, tests were not readministered, but the recent historical test results were used.

Participants completed the ASRT, a task constructed to elicit naturalistic speech within a closed domain. Prerecorded ASRTs were presented at a steady reading rate (approximately 140 words per minute) by a British male speaker. Parallel stimuli included 36 stories: 18 short stories (119 words per story, SD 4.83) and 18 long stories (224 words per story, SD 14.86). Task characteristics are presented in Table S1 in Multimedia Appendix 1, showing that stories incorporate a range of themes and are balanced for key linguistic and discourse metrics. During clinical assessments, 3 long ASRT stories were administered consecutively. After each story was presented, participants were asked to immediately retell the story in as much detail as they could remember. Recall of the same stories, in the same order, was tested again after a delay.

During clinical assessments, participants were supported with installing the Novoic mobile app ("the app") on their own smartphone device and were shown how to use it. Participants were reimbursed for their participation at the end of the study visit and before remote assessments (£65 [US $86] for United Kingdom participants and US $75 for US participants). No threshold for use during remote follow-up was required for participants to be fully remunerated.

Remote Assessments

Participants were encouraged to complete optional unsupervised self-assessments (<30 minutes in length) on the app daily for up to 8 days following the study visit. Assessments included ASRTs and other remote speech tasks not reported here (verbal and category fluency assessments, reading tasks, picture description, and procedural discourse tasks) as well as remote questionnaires. ASRTs were administered at the beginning of each assessment, with the order, inclusion, and administration of other tests varying by day.

Distinct assessment components (ASRTs [+fluency tasks as appropriate], questionnaires, and other tasks) were divided so that participants, once completing one component, were informed of their progress and given the opportunity to continue. This meant that participants could take breaks between assessment components. All ASRTs were administered within one of these assessment components, without interruption. If and where participation was interrupted because of other factors (distraction, etc), individual audio tasks administered were not repeated, but participants were able to continue with the following part of the assessments.

Remote ASRTs were administered daily, in threes (triplets) and at the beginning of each assessment session. The ASRT stories administered on the first day of remote assessment were identical to those administered in the clinical supervised assessment on the prior day, to allow for the evaluation of practice effects (not reported here). The remainder of the ASRT stories, presented

XSL•FO
RenderX

from day 2 of remote assessment onward, were novel and administered only once.

After each story was presented, participants were asked to immediately retell the story in as much detail as they could remember. Recall of the same stories, in the same order, was tested again after a delay. The schedule included delayed recall after completion of all immediate recalls or after completion of brief distractor tasks (fluency tasks: category or verbal fluency), with test administration varying by day (shown in Table S2 in Multimedia Appendix 1). Recordings of spoken responses were automatically started by the app following instructions and manually stopped by the participants. These were recorded as audio files on participants' personal smart devices and automatically uploaded to a secure server.

Due to participant-initiated feedback of high burden (that the remote assessments were too long and tiring), the assessment schedule was changed partway through the study. The new schedule implemented shorter stories and reduced the number of additional assessments following ASRTs (not reported here). However, ASRTs continued to be administered daily at the start of each assessment. Simultaneously, the number of days of remote assessment was increased from 7 days to 8 days to spread out assessments and reduce the daily burden. Details are provided in Table S2 in Multimedia Appendix 1.

Daily state effects were assessed at the end of each remote assessment via a 4-item self-report questionnaire asking how participants were feeling that day (current mood, quantity of sleep, mind-wandering, and effort), scored on a 7-point response scale from "much worse/less than usual" to "much better/more than usual." App and task usability were assessed via a self-report questionnaire on day 2 (initial assessment schedule) or day 5 (revised assessment schedule). Usability questionnaires asked participants to report technical difficulties experienced during assessments, whether technical difficulties prevented them from completing the assessments, how easy it was to use the app, and how interesting the tasks were. Questionnaires are shown in Tables S3 and S4 in Multimedia Appendix 1.

Statistical Analysis

ASRT task responses were transcribed using Google's speech-to-text [32] automatic speech recognition system, using an enhanced speech recognition model (the "video" model, suitable for recordings that may contain background noise). All task responses were also transcribed manually by following a standardized procedure, which included transcription of commentary, filled pauses, and partial words. The word error rate (WER) of the automatic transcript was calculated using the HuggingFace package [33], as the average number of errors per manual transcript word. This was calculated after removing punctuation, setting all text characters to lower case, and removing filled pauses and partial words from the transcripts before comparison.

Transcription was followed by automated textual analysis completed using a generalized match (G-match) score. G-match was computed in Python as the weighted sum of the cosine similarity between the embeddings of original ASRT text and the transcribed retellings, providing an automatic quantitative evaluation of similarity across the 2 texts. G-match provides an index of the proportional recall for each story, with potential scores ranging from 0 to 1 (hypothetically perfect performance). Mean G-match per triplet was also computed. The underlying representations of the model are based on a pretrained BERT model [34], which is a large language model pretrained on a corpus of more than 3000 million words, to produce generalized representations of language and how it is used.

Further analysis was performed using the statistical software package R v.4.0. Data were assessed for normality, followed by parametric and nonparametric analyses as appropriate. Adherence was defined as engaging with at least one ASRT story per day. Adherence patterns were examined with logistic regression models, predicting adherence at immediate and delayed recall, in relation to participant group, demographics, assessment day, and schedule. A large proportion of participants completed only 7 days of remote assessments, and longitudinal analysis of adherence was therefore limited to assessments on days 1 to 7. Participants were included as random effects. Demographics (sex, age, and years of education), assessment days (1-7), research schedule (schedule 1 or schedule 2), and participant group (CU and MCI or mild AD) were included as fixed factors.

The parallel-forms reliability of ASRTs was examined using pairwise correlational analysis. Only ASRT stories administered across both test schedules were analyzed, maintaining comparable sample sizes across comparisons and allowing for testing within the MCI or mild AD and CU subgroups. The convergent validity of these same ASRT stories was examined in relation to LM, PACC5, and CDR-SB. Analyses were repeated using the mean G-match score per triplet. Spearman rank correlation coefficients are reported throughout to maintain the consistency and comparability of reporting.

Task performance differences between groups, task administration variations, and change over time were modeled using longitudinal linear mixed-effects models. Data analyzed were restricted to remote assessment days 2 to 7, when assessments were novel and administered to all participants. The mixed model analysis included G-match as the response variable, and fixed effects of participant group, remote assessment days (2-7), order (1st, 2nd, or 3rd ASRT presented), long or short stories, and immediate or delayed recall. Demographics (age, sex, and education) were included as additional fixed effects. A random effect of participant with random slope and intercept was specified. Cohen *d* effect sizes for multilevel model objects were calculated using the lme.dscore command in the package EMAtools.

Analyses were repeated with the mean G-match per triplet, with equivalent random and fixed effects specifications, except for the story order, which was not included. The covariation of mean ASRT task performance across triplets with self-reported daily state was then examined by additionally incorporating fixed effects of self-reported mood, sleep, effort, and mind-wandering into the above model. The assumptions of all regression models were investigated by examining the distribution and patterns of residuals versus fitted values.

XSL•FO
RenderX

Group differences and effect sizes were also evaluated for traditional cognitive tests completed with a trained psychometrician via Zoom or in person during clinical assessments. Comparisons were only carried out for tests that were not directly or indirectly part of the study selection criteria (Digit-Symbol Coding, the Free and Cued Selective Reminding Test, and category fluency), thereby excluding MMSE (direct selection criterion), the LM Delayed Recall, and PACC5 (indirect). As participants were recruited from prior completed trials, in some of which performance thresholds on the LM delayed recall contributed to the MCI and mild AD group inclusion criteria, LM and PACC5 (of which the LM is a component) were not evaluated. The test distributions of traditional cognitive assessments were evaluated for normality, followed by parametric or nonparametric tests, as appropriate.

Results

Participants

A total of 200 participants, 67 from the US study and 133 from the UK study, were recruited and completed the clinical assessment protocol. In total, 75.5% (151/200) of the participants completed at least one remote ASRT. Older participants (r=−0.15; P=.03), those with lower MMSE scores (r=−0.26; P<.001), and those with MCI or mild AD (33/106, 31.1% MCI or mild AD, compared with 16/94, 17% CU; χ^2_1=5.4; P=.02) more often did not complete any remote assessments. There were no differences in sex ratio (χ^2_1=0.4; P=.50) or years of education (r=−0.01; P=.87) between participants who contributed at least one remote assessment and those who did not.

Demographic information of the participants providing remote data are presented in Table 1. In this sample, the MCI or mild AD and CU groups did not differ with respect to age, years of education, sex, or amyloid status. The US study included proportionally more participants with cognitive impairment (22/34, 65% with MCI) than the UK sample (51/117, 43.6% with a diagnosis of MCI or mild AD). The MCI or mild AD group included a minority of participants with a diagnosis of mild AD (10/73, 14%), all recruited into the UK sample as per the inclusion criteria. A detailed breakdown of the sample characteristics by US and UK studies is provided in Table S5 in Multimedia Appendix 1.

Table 1. Participant demographic characteristics of cognitively unimpaired participants and participants with MCI[a] or mild AD[b].

	Group		Statistical values	
	Cognitively unimpaired (n=78)	MCI or mild AD (n=73)	Test statistic	*P* value
Sex, n (%)			χ^2_1=0.3	.61
Female	47 (60)	41 (56)		
Male	31 (40)	32 (44)		
Country of residence, n (%)			χ^2_1=4.7	.03
United Kingdom	66 (85)	51 (70)		
United States	12 (15)	22 (30)		
Testing schedule, n (%)			χ^2_1=7.0	.008
Schedule 1	40 (51)	22 (30)		
Schedule 2	38 (49)	51 (70)		
Amyloid beta status, n (%)			χ^2_1=0.8	.36
Amyloid negative	38 (49)	41 (56)		
Amyloid positive	40 (51)	32 (44)		
Years of education, mean (SD)	15.24 (3.37)	15.06 (2.80)	r=−0.05	.57
Age (years), mean (SD)	70.37 (4.35)	69.58 (7.30)	r=−0.01	.91
MMSE[c], mean (SD)	28.92 (1.15)	27.00 (2.07)	r=0.50	<.001

[a]MCI: mild cognitive impairment.

[b]AD: Alzheimer disease.

[c]MMSE: Mini-Mental State Examination.

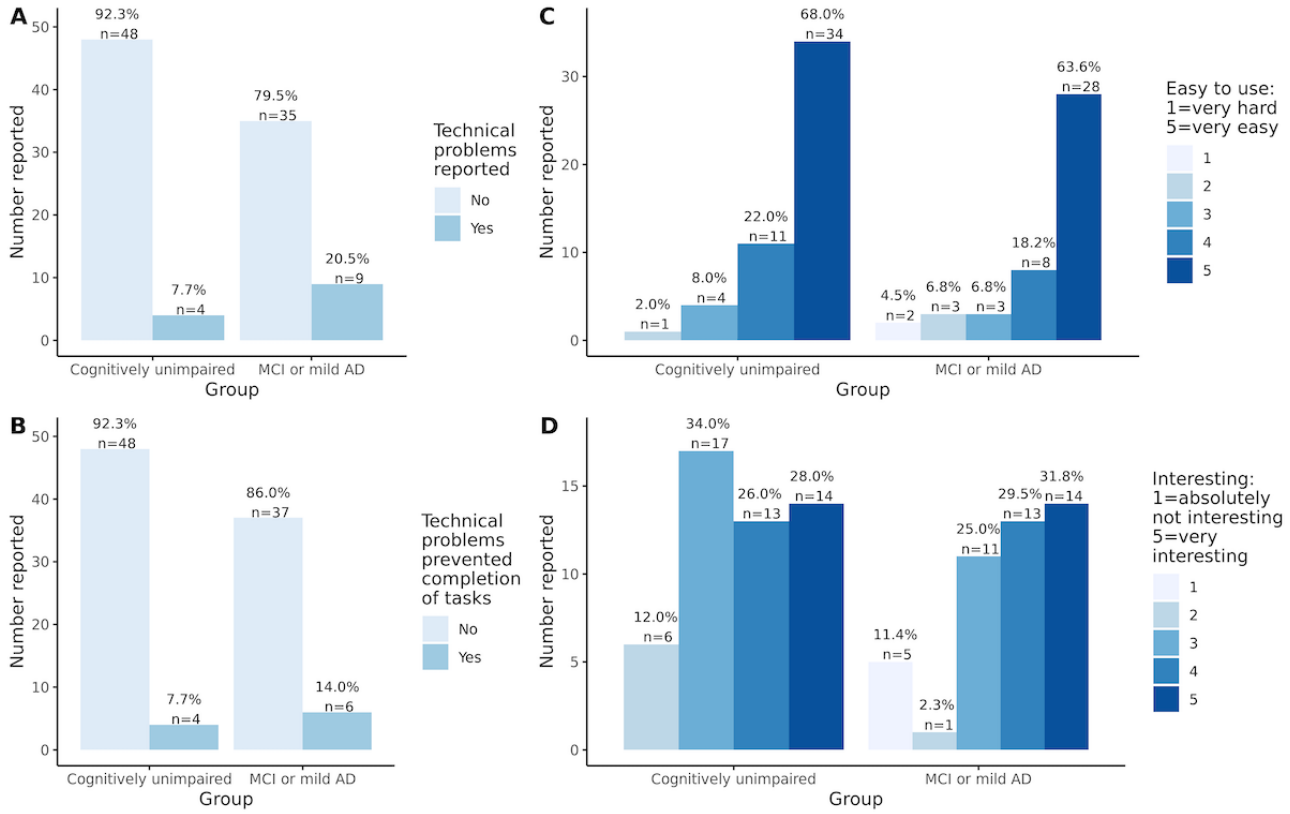
Usability

Usability questionnaires were completed by 63.6% (96/151; CU: n=52 and MCI or mild AD: n=44) of the participants who completed remote assessments (Figure 1). Those completing usability questionnaires did not differ with respect to education level (r=−0.02; P=.78), age (r=−0.12; P=.14), or MMSE scores (r=−0.08; P=.32) compared with those who engaged in remote assessments but did not complete usability questionnaires. There was also no difference in the male to female ratio (χ^2_1=0.1; P=.75) or the ratio of CU participants to participants with MCI

XSL•FO
RenderX

or mild AD (χ^2_1=0.7; *P*=.41) who did and did not complete usability questionnaires.

Figure 1. Responses to usability questionnaire: (A) technical problems reported, (B) rate at which technical problems prevented completion of tasks, (C) ease of use of app, and (D) interest in tasks completed. AD: Alzheimer disease; MCI: mild cognitive impairment.



In total, 8% (4/52) of CU participants and 20% (9/44) of participants with MCI or mild AD reported technical difficulties. Where technical difficulties were encountered, most participants reported that these did not prevent them from completing the assessments, with no group differences (χ^2_1=3.3; *P*=.07 and χ^2_1=1.0; *P*=.32, respectively for technical difficulties reported, and inability to complete assessments). Most participants responded that the app was easy to use and that the task was reasonably interesting, with no group differences (*r*=−0.08; *P*=.47 and *r*=−0.04; *P*=.70, respectively for ease of use and interest in tasks).

Adherence

Participants with MCI or mild AD completed fewer remote assessments than CU participants (adherence for immediate recall: 64.5% vs 77.5%; delayed recall: 61.5% vs 77.3%; Figure 2). Group differences were confirmed by mixed logistic regression analyses (immediate recall estimate=−0.97; *P*=.01 and delayed recall estimate=−0.84; *P*=.02). Adherence did not change over the assessment days (immediate recall estimate=−0.04; *P*=.34 and delayed recall estimate=−0.07; *P*=.11), but lower adherence to delayed recall was observed for the revised test schedule (estimate=−0.86; *P*=.03). Adherence was not associated with sex and education (all *P*>.20), but younger participants completed more immediate recall assessments (immediate recall estimate=−0.07; *P*=.02 and delayed recall estimate=−0.06; *P*=.06).

XSL•FO
RenderX

Figure 2. Adherence and task performance heat map for generalized match (G-match) in immediate recall trials. G-match is an automated measure of recall performance (refer to the Methods section). Results are plotted across individual days of remote assessment for 151 participants who completed at least one assessment. Each participant is represented by a row, missing data are shown in gray, and mean G-match across the Automatic Story Recall Task triplets is shown in color (red=low recall and yellow=high recall). AD: Alzheimer disease; MCI: mild cognitive impairment.

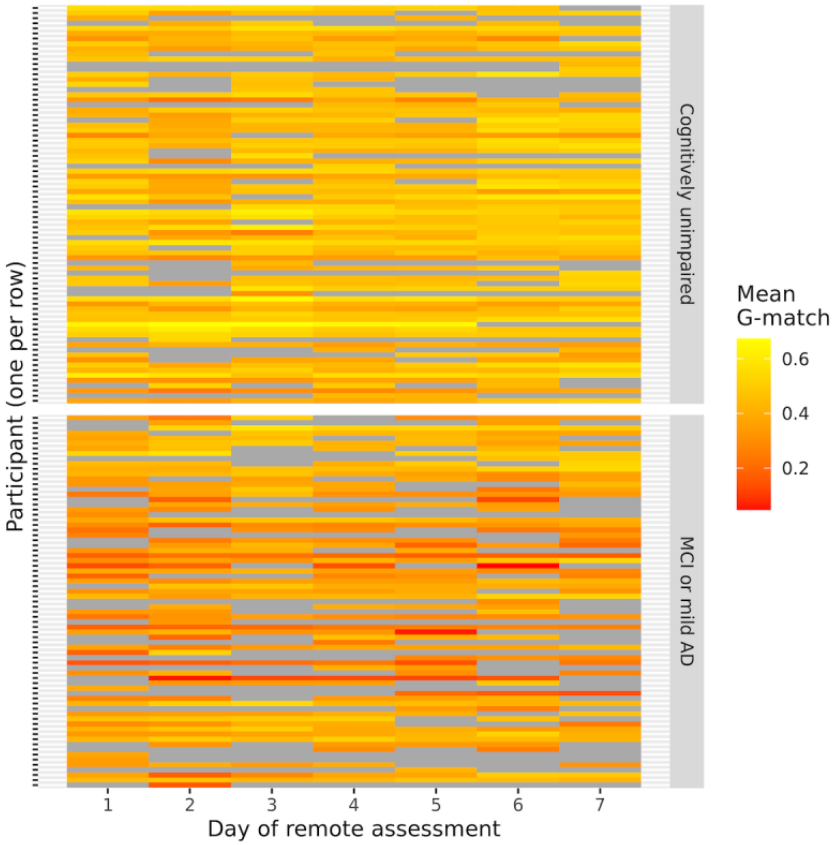


Figure 2 shows a heat map of the adherence patterns and task performance. In this figure, each participant is represented by a row, and task response and performance over the days of assessment are shown in colored blocks along the x-axis. Task performance is shown in color, with red to yellow grading representing low to high G-match scores. Missing data are shown in gray. This figure reflects the results reported above, with higher adherence in the CU group and no clear decline in adherence over the assessment period.

Transcription Accuracy

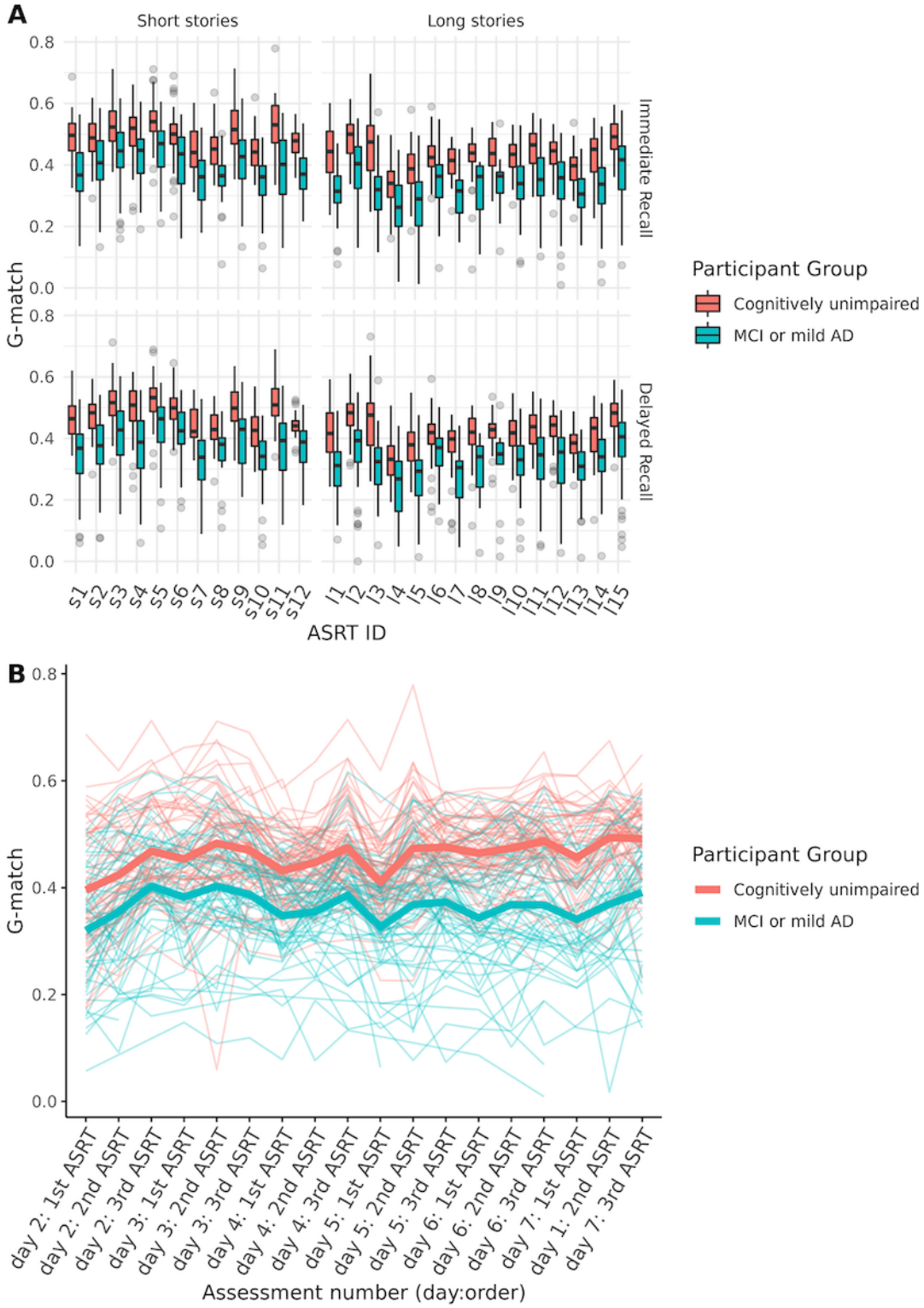
The average WER for participant recordings of automatic transcripts compared with manual transcripts was 0.11. The average WER differed across participant groups, with WER=0.09 in CU participants and WER=0.13 in participants with MCI or mild AD ($t_{108.1}$=−3.81; P<.001; Cohen d=0.63).

Task Characteristics

G-match for ASRTs and triplets showed good psychometric properties. Data generated showed no ceiling or floor effects (Figure 3A; Figures S1-S4 in Multimedia Appendix 1). Task performance characteristics are provided in Tables S6-S8 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 3. Generalized match (G-match) over repeated assessments: (A) boxplots of G-match for individual Automatic Story Recall Task stories split by short and long stories horizontally and by immediate and delayed recalls vertically and (B) average G-match (immediate recall) over individual assessment days (2-7 and immediate recall) and testing order. Group means are displayed with the thick lines, and individual participant trajectories across assessments and days are shown with paler, thinner lines. AD: Alzheimer disease; ASRT: Automatic Story Recall Task; MCI: mild cognitive impairment.

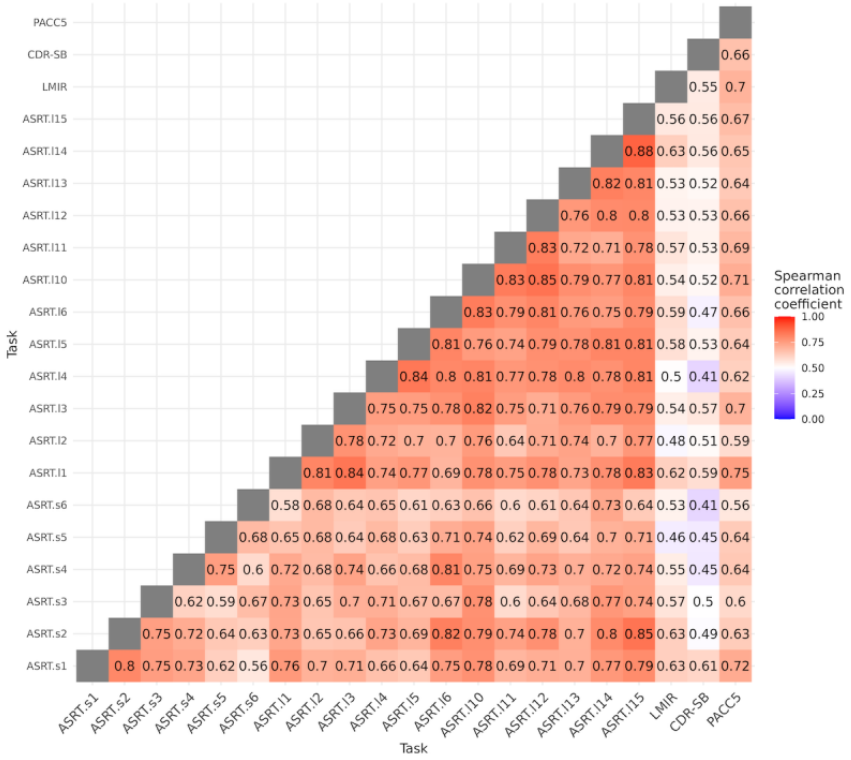


Parallel-Forms Reliability

Parallel-forms reliability for individual ASRT stories at immediate recall are presented in Figure 4. Equivalent figures for delayed recall, separated by clinical group, are presented in Figures S5-S9 in Multimedia Appendix 1. Correlation matrices for triplets separated by immediate and delayed recall, and clinical groups, are shown in Figures S10-S12 in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 4. Parallel forms reliability and convergent validity of Automatic Story Recall Task (ASRT) stories at immediate recall. ASRTs are denoted with s (short) and l (long), followed by the story number (refer to Table S1 in Multimedia Appendix 1). Correlations with other assessments are displayed (Wechsler Logical Memory Test–Immediate Recall [LMIR], Clinical Dementia Rating scale–Sum of Boxes [CDR-SB], and Preclinical Alzheimer's Cognitive Composite with semantic processing [PACC5]). The sign for the CDR-SB correlation is reversed for consistency. Correlation coefficients derived from between 75 and 116 participants, depending on adherence patterns.



Correlation coefficients in the full sample were moderate to strong for immediate recall (rho range=0.56-0.88; mean 0.73) and remained so after restricting analyses to participants with MCI or mild AD (rho range=0.31-0.87; mean 0.65) and CU participants (rho range 0.39-0.85; mean 0.65). Similarly, correlations between parallel ASRT stories were moderate to high for delayed recall (full sample: rho range=0.54-0.86; mean 0.73) and remained so when restricting analyses to participants with MCI or mild AD (rho range=0.37-0.88; mean 0.65) and CU participants (rho range=0.32-0.83; mean 0.64).

Parallel-forms reliability was higher when examined for mean scores obtained across triplets (immediate: rho range=0.77-0.88, mean 0.83; and delayed: rho range=0.76-0.89, mean 0.85), remaining consistently high in MCI or mild AD (immediate: rho range=0.57-0.88, mean 0.73; and delayed: rho range=0.60-0.89, mean 0.75) and CU subgroups (immediate: rho range=0.67-0.83, mean 0.76; and delayed: rho range=0.68-0.85, mean 0.77).

Convergent Validity

ASRT task performance correlated moderately with other cognitive and clinical measures (LM, CDR-SB, and PACC5) in the full sample across both immediate and delayed recalls (Figure 4). The mean correlation coefficients between immediate recall ASRTs with LM-immediate recall, PACC5, and CDR-SB were rho=0.56, 0.65, and 0.51, respectively. The mean correlation coefficients between ASRTs with LM-delayed recall, PACC5, and CDR-SB were rho=0.54, 0.66, and 0.50, respectively. Analysis results and figures for delayed recall and results separated by participant group are provided in Figures S5-S9 in Multimedia Appendix 1. Correlation coefficients remained in the moderate range after restricting analyses to participants with MCI or mild AD but were typically lower in CU participants. Correlations between ASRT triplets and other cognitive tests are provided in Figures S10-S12 in Multimedia Appendix 1.

Task Performance Comparison Between Groups

The longitudinal mixed models are presented in Table 2, with similar results for individual ASRTs and triplets. Task performance improved across the week, with a modest linear daily improvement in G-match by assessment day. There was an effect of group with lower scores in the MCI or mild AD group for both individual stories and triplets, with an effect size of Cohen *d*=1.54. G-match was higher for immediate recall and shorter stories and higher for the latter ASRTs administered within each triplet. Demographics were not associated with task performance. Longitudinal data are displayed in Figure 3B, showing within- and between-subject variability.

XSL•FO
RenderX

Table 2. Effects of task characteristics, participant group, and demographics on task performance metrics as estimated by longitudinal mixed models. For binary predictors (sex, ASRT[a] length, and recall type) the reference category is listed first.

	G-match individual stories		G-match triplets	
	Estimate (SE)	*P* value	Estimate (SE)	*P* value
Intercept	0.53 (0.08)	<.001	0.57 (0.08)	<.001
Group (Group 1: CU[b], Group 2: MCI[c] or mild AD[d])	−0.11 (0.01)	<.001	−0.11 (0.01)	<.001
Assessment day	0.005 (0.001)	<.001	0.005 (0.001)	<.001
Recall type (immediate and delayed)	−0.02 (0.001)	<.001	−0.02 (0.002)	<.001
ASRT length (short and long)	−0.04 (0.003)	<.001	−0.04 (0.003)	<.001
ASRT order of presentation (1,2, and 3)	0.02 (0.001)	<.001	—[e]	—
Sex (female and male)	−0.02 (0.01)	.08	−0.02 (0.01)	.07
Education (years)	0.0004 (0.002)	.83	0.0003 (0.002)	.85
Age (years)	−0.002 (0.001)	.12	−0.002 (0.001)	.13

[a]ASRT: Automatic Story Recall Task.

[b]CU: cognitively unimpaired.

[c]MCI: mild cognitive impairment.

[d]AD: Alzheimer's disease.

[e]Fixed effect not included in model.

After incorporating self-report assessments into the mixed model predicting G-match for triplets, the models revealed a significant effect of mood (estimate=0.007; SE 0.002; *P*<.001) and mind-wandering (estimate=−0.007; SE 0.002; *P*<.001), with better daily mood and lower mind-wandering associated with better daily task performance.

Comparison With Traditional Neuropsychological Tests

Traditional neuropsychological tests administered in person during in-clinic assessments were also predictive of MCI or mild AD diagnostic status, with large effect sizes identified: Digit-Symbol Coding: t_{82}=5.40, *P*<.001, Cohen *d*=1.07; category fluency total score: *t* t_{148}=7.16, *P*<.001, Cohen *d*=1.17; and the sum of free and total recall from the Free and Cued Selective Reminding Test: t_{108}=5.56, *P*<.001, Cohen *d*=1.01.

Discussion

Principal Findings

This study indicates that daily unsupervised and self-administered speech-based testing is acceptable and feasible for older participants with and without cognitive impairment. Participants engaged in daily optional assessments with moderate levels of adherence. There was no observable reduction in adherence levels over a weeklong period of assessment. The participants experienced infrequent technical problems and reported that the tests were easy to use and reasonably interesting.

Results indicate that remote automatic test administration and autoscoring of story recall can provide sensitive cognitive measurement in at-risk populations. The ASRT G-match, an automatically scored measure of proportional recall, showed consistent differences in task performance between cognitively healthy participants and those with MCI or mild AD. Separation in task performance between diagnostic groups was consistent across the assessment period and across individual ASRT stories (Figure 3), showing a strong effect size for differentiating CU participants from those with MCI or mild AD (Cohen *d*=1.54), while controlling for age, education, and sex. The equivalent area under the receiver operating characteristic curve was 0.86, based on previously published equivalence tables [35].

The effect size for ASRTs is larger than that seen for a range of traditional cognitive tests typically administered in person and under supervision. Comparisons with the LM delayed recall and PACC5 were not made, as participants in this study were recruited from prior trials in which test performance on LM delayed recall constituted part of the trial inclusion criteria for patients with MCI or mild AD, which would likely inflate effect sizes for these tests.

The ASRTs discrimination between clinical groups reported here outperforms those previously reported for differentiating CU individuals from those with MCI using other traditional cognitive tests administered in person and in the clinic, such as the MMSE (Cohen *d*=0.69), the 6-Item cognitive impairment test (Cohen *d*=0.65), and Addenbrooke's Cognitive Examination-Revised (Cohen *d*=0.73), albeit with similar results reported previously for the Montreal Cognitive Assessment battery (Cohen *d*=1.45) [36]. The test also performs well in comparison with the Cogstate brief battery, when administered remotely and unsupervised, where effect sizes for differences between MCI and CU groups in subtests range from Cohen *d*=0.22 to Cohen *d*=0.62 [37].

Although the mixed clinical group examined in this analyses limits direct comparison with previously published metrics in subjects with MCI only, the mild AD group in this study comprised only a small proportion (10/73, 14%) of those contributing to the MCI or mild AD group. After excluding

XSL•FO
RenderX

participants with mild AD from the linear mixed model analysis, this yielded an effect size of Cohen *d*=1.45 (equivalent area under the curve=0.85 [35]) for the difference between the CU participants and participants with MCI.

ASRT stimuli are carefully designed and balanced for key linguistic and discourse metrics, including the number of words, number of sentences, number of dependent clauses, mean sentence length, and ratio of dependent clauses to t-units (the number of shortest grammatically complete units into which a string of written or spoken language can be partitioned). This balancing of the stimuli is also reflected in good parallel-forms reliability between ASRT parallel stimuli, which is consistently high across immediate and delayed recall and with analyses constrained to clinical subgroups (MCI or mild AD and CU). The ASRT analysis pipeline also has significant advantages for test-retest reliability and parallel-forms reliability, as text similarity is evaluated in the same way every time, producing a standardized scoring system across the parallel test forms. A more objective quantification of text similarity, using a large language corpus for training, removes some of the more arbitrary features common to story recall task scoring, in which specific paraphrases are accepted, and the size of information units shows some variability [29].

Repeated exposure to the test stimuli may lead to unwanted practice effects that reduce the validity of the test as a measure of new learning, with research also showing differential practice effects across clinical diagnostic groups for tests such as list learning tests and LM [38]. Practice effects may be particularly important when considering where the same story recall stimuli are used repeatedly in longitudinal research or clinical monitoring or for diagnostic thresholding as cut-offs for research studies or clinical trials [16,39]. Other available story recall tests typically have a limited range of parallel forms.

The number of available parallel forms of the ASRT test allows for a higher-frequency (daily) assessment over a shorter period without test repetition, such as that carried out in this study. Alternatively, tests could be administered at larger intervals (weekly, monthly, or longer) to evaluate longer-term changes with little or no repetition of stimuli, thereby likely reducing practice effects.

Although alternate test variants can help reduce practice effects, they do not completely correct for retesting, which can be modified by repeated exposure to the task and greater familiarity with the test structure or method [40]. In this study, despite novel stimuli being presented during each assessment, test scores improved modestly during the week, indicating that increased familiarity with the app, testing procedure, and test structure resulted in a subtle improvement over time. Task performance improvements over the weeklong assessment period were modest (with an estimated daily change in G-match of only 0.4% of the initial intercept estimate value). The improvement in test performance, in combination with the absence of adherence changes over the course of the study, did not indicate any strong fatigue effects.

ASRT tests correlated moderately with a well-established test of verbal episodic memory, cognitive composites, and clinician-reported outcomes, indicating acceptable convergent validity, and with results comparable with, or better than, other studies of computerized or unsupervised remote assessments [41-43]. Correlations with LM and clinician-reported outcomes were in the moderate range, with lower correlation coefficients linked to test invariance owing to ceiling- or floor-level performance on these traditional clinical and neuropsychological assessments in CU individuals.

Task performance also varied with aspects of study design, with stories administered later in triplets delivering a more comprehensive recall. These effects appear to lead to greater variation between individual ASRT stories but are averaged out when the G-match is examined across story triplets. Analysis of story triplets showed higher parallel-forms reliability between the ASRTs administered and analyzed in threes, albeit with broadly unchanged differences in group performance. Task performance, as measured with G-match, was typically higher for shorter stories, indicating that responses more comprehensively covered the story source text where participants were asked to recall less material.

This study also showed within-subject variation in task performance, in part reflecting the measured effects of state factors on cognitive performance, in particular daily mood and effort. Variation from within-subject differences can make it challenging to differentiate clinical change from measurement error [8], and higher-frequency assessments carried out longitudinally can help generate more reliable estimates of cognitive function and change. Repeated measurements allow these state effects to be concurrently measured and included or controlled for in the longitudinal analyses.

Limitations

To meet the eligibility criteria, participants were required to be able to use and access a smartphone. This may have biased the sample by overselecting those with higher familiarity with technology. Older and more cognitively impaired participants were less likely to contribute to the remote study component, and when they engaged in remote assessments, they contributed less frequently. However, the adherence statistics presented here reflect participants' engagement in optional assessments, which may have differed had these been compulsory. Many home testing options require at least a modest level of technological fluency that some older adults may find challenging, challenges that may be compounded by cognitive impairment or comorbidities [44].

Therefore, the data presented may not reflect task performance in more impaired individuals or those with lower levels of technological familiarity. Assessments under supervision, either in the clinic or during a telemedicine visit, allowing for provision of additional support where required, could be better suited to more impaired individuals.

By collecting usability data during remote assessments, we were able to establish that most participants did not experience any technical problems and that the app was generally easy to use. However, more detailed qualitative feedback on the type of usability and any technical issues was not collected. Further evaluation of the nature of these difficulties is required. This

XSL•FO
RenderX

information can be used to improve the user interface and participant engagement with remote assessments.

In response to participants' and study centers' feedback on the high participant burden of the initial test schedule, the testing schedule was altered in the middle of the study to reduce burden, thereby limiting the data available for certain ASRT test variants.

The design of the study makes it difficult to differentiate between the effects of individual stories themselves (ie, which ASRT story was used) and the effects of study design, such as the test order or day of assessment. Future studies may benefit from adopting a randomized design, with ASRTs randomly selected and allocated to different testing instances to derive test performance metrics independent of these additional confounders. For longitudinal studies, either short or long stories should be adopted to improve the consistency of test scores over time and help better characterize cognitive change.

We found differences in WER when comparing automated and manual transcripts of CU participants and those with MCI or mild AD, indicative of the differential intelligibility of speech or recording quality in these 2 groups. Differences in the performance of automatic transcription will impact the analysis further along in the analysis pipeline, indicating that group differences in scores likely reflect not only group differences in proportional recall but may also incorporate speech intelligibility and participant's device use characteristics. However, these effects warrant further investigation.

The participants included in this study constituted a select sample. The sample was selected to exclude patients with concurrent neurological and mental health conditions. They were recruited from prior clinical trials completed in the United States and the United Kingdom and reflect a group of individuals who are actively engaged in clinical research. The participants lacked racial diversity (with most of the sample identifying as White and with only 2.6% [N=4] with Asian, Black, African, or African American background). Replication is now needed in more clinically and demographically heterogeneous samples.

Overview and Future Directions

The recent Food and Drug Administration approval for the first disease-modifying treatment for people at risk of developing AD highlights the importance of adequate screening and early detection as well as the importance of monitoring treatment response. Briefer, convenient, and lower-burden daily assessments may provide more reliable data to evaluate disease progression or treatment response than lengthy one-off assessments [9]. Brief digital assessments completed at home and repeatable over time could improve access to AD screening compared with current clinical standards, which typically require clinical visits and extensive neuropsychological assessment.

This study showed that brief, remotely administered, and automatically scored ASRTs are sensitive to early cognitive impairments commonly identified through more extensive clinical assessments. The tests showed good properties for repeated administration and convergent validity with established tests of episodic memory, cognitive composites, and clinician-reported outcomes (CDR-SB). The test showed good acceptability and usability for older adults with and without cognitive impairment. Furthermore, owing to the automatic administration and scoring of ASRTs, this test presents a minimal administrative burden, requiring no trained personnel or specialist equipment.

Speech is instrumental in daily functioning and a natural response modality for participants to use in response to current smart devices, such as smartphones. Speech responses are also a common component of cognitive tests; however, data generated in these tests, including those reported in this study, often relate simple pass or fail characteristics of response accuracy. New metrics using audio- and text-based artificial intelligence models to target other changes measurable in speech data (acoustic [45,46], semantic [47-49], and linguistic [46]) in early-stage AD could further leverage the information content of ASRTs, developing a new class of powerful, fully automated speech biomarkers.

Acknowledgments

The authors are extremely grateful to the participants who took part in the study and their families and carers who supported their participation. The authors also thank the study sites and their scientific and research teams for recruitment, study coordination, conducting interviews, and data collection efforts. The study was funded by Novoic, a clinical-stage digital medtech company developing artificial intelligence–based speech biomarkers.

Data Availability

Speech data are identifying and cannot be shared; however, all quantitative data produced in this study are available upon reasonable request to the authors.

Authors' Contributions

EF, MM, and JW designed the study protocol. MM coordinated the research study and data collection. CS, UM, and EF completed the analyses. CS and EF completed the first draft of the manuscript. All authors contributed to the revision of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

EF, JW, MM, CS, RL, and UM are current or prior employees of Novoic Ltd. KVP is an advisor to the company. EF, JW, RL, and MM are shareholders and CS, MM, UM, and KVP are option holders in the company. JW and EF are both Directors at Novoic and named inventor on multiple patents in speech processing filed by Novoic ltd. KVP has served as a paid consultant for Biogen Idec and Digital Cognition Technologies.

Multimedia Appendix 1
Supplementary information, tables, and figures for daily remote administration of the Automatic Story Recall Task.
[DOC File , 1923 KB - aging_v5i3e37090_app1.doc]

References

1. FDA Grants accelerated approval for Alzheimer's drug. U.S. Food & Drug Administration. 2021 Jun 7. URL: https://www.fda.gov/news-events/press-announcements/fda-grants-accelerated-approval-alzheimers-drug [accessed 2021-06-29]
2. Kozauer N, Katz R. Regulatory innovation and drug development for early-stage Alzheimer's disease. N Engl J Med 2013 Mar 28;368(13):1169-1171. [doi: 10.1056/NEJMp1302513] [Medline: 23484795]
3. Guideline on the clinical investigation of medicines for the treatment of Alzheimer's disease. European Medicines Agency. 2018 Feb 22. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicines-treatment-alzheimers-disease-revision-2_en.pdf [accessed 2021-04-12]
4. Lo JC, Groeger JA, Cheng GH, Dijk D, Chee MW. Self-reported sleep duration and cognitive performance in older adults: a systematic review and meta-analysis. Sleep Med 2016 Jan;17:87-98 [FREE Full text] [doi: 10.1016/j.sleep.2015.08.021] [Medline: 26847980]
5. Chang YK, Labban JD, Gapin JI, Etnier JL. The effects of acute exercise on cognitive performance: a meta-analysis. Brain Res 2012 May 09;1453:87-101. [doi: 10.1016/j.brainres.2012.02.068] [Medline: 22480735]
6. Mitchell RL, Phillips LH. The psychological, neurochemical and functional neuroanatomical mediators of the effects of positive and negative mood on executive functions. Neuropsychologia 2007 Mar 02;45(4):617-629. [doi: 10.1016/j.neuropsychologia.2006.06.030] [Medline: 16962146]
7. Angelidis A, Solis E, Lautenbach F, van der Does W, Putman P. I'm going to fail! Acute cognitive performance anxiety increases threat-interference and impairs WM performance. PLoS One 2019;14(2):e0210824 [FREE Full text] [doi: 10.1371/journal.pone.0210824] [Medline: 30730896]
8. Meier IB, Buegler M, Harms R, Seixas A, Çöltekin A, Tarnanas I. Using a Digital Neuro Signature to measure longitudinal individual-level change in Alzheimer's disease: the Altoida large cohort study. NPJ Digit Med 2021 Jun 24;4(1):101 [FREE Full text] [doi: 10.1038/s41746-021-00470-z] [Medline: 34168269]
9. Schweitzer P, Husky M, Allard M, Amieva H, Pérès K, Foubert-Samier A, et al. Feasibility and validity of mobile cognitive testing in the investigation of age-related cognitive decline. Int J Methods Psychiatr Res 2017 Sep;26(3):e1521 [FREE Full text] [doi: 10.1002/mpr.1521] [Medline: 27539327]
10. Donohue MC, Sperling RA, Salmon DP, Rentz DM, Raman R, Thomas RG, Australian Imaging, Biomarkers,Lifestyle Flagship Study of Ageing, Alzheimer's Disease Neuroimaging Initiative, Alzheimer's Disease Cooperative Study. The preclinical Alzheimer cognitive composite: measuring amyloid-related decline. JAMA Neurol 2014 Aug;71(8):961-970 [FREE Full text] [doi: 10.1001/jamaneurol.2014.803] [Medline: 24886908]
11. Papp KV, Rentz DM, Orlovsky I, Sperling RA, Mormino EC. Optimizing the preclinical Alzheimer's cognitive composite with semantic processing: The PACC5. Alzheimers Dement (N Y) 2017 Nov;3(4):668-677 [FREE Full text] [doi: 10.1016/j.trci.2017.10.004] [Medline: 29264389]
12. Insel PS, Weiner M, Mackin RS, Mormino E, Lim YY, Stomrud E, et al. Determining clinically meaningful decline in preclinical Alzheimer disease. Neurology 2019 Jul 23;93(4):e322-e333 [FREE Full text] [doi: 10.1212/WNL.0000000000007831] [Medline: 31289148]
13. Lim YY, Snyder PJ, Pietrzak RH, Ukiqi A, Villemagne VL, Ames D, et al. Sensitivity of composite scores to amyloid burden in preclinical Alzheimer's disease: introducing the Z-scores of attention, verbal fluency, and episodic memory for nondemented older adults composite score. Alzheimers Dement (Amst) 2016;2:19-26 [FREE Full text] [doi: 10.1016/j.dadm.2015.11.003] [Medline: 27239532]
14. Jonaitis EM, Koscik RL, Clark LR, Ma Y, Betthauser TJ, Berman SE, et al. Measuring longitudinal cognition: individual tests versus composites. Alzheimers Dement (Amst) 2019 Dec;11:74-84 [FREE Full text] [doi: 10.1016/j.dadm.2018.11.006] [Medline: 31673596]
15. Porto MF, Benitez-Agudelo JC, Aguirre-Acevedo DC, Barceló-Martinez E, Allegri RF. Diagnostic accuracy of the UDS 3.0 neuropsychological battery in a cohort with Alzheimer's disease in Colombia. Appl Neuropsychol Adult 2021 Mar 24:1-9. [doi: 10.1080/23279095.2021.1897007] [Medline: 33761292]
16. Chapman KR, Bing-Canar H, Alosco ML, Steinberg EG, Martin B, Chaisson C, et al. Mini Mental State Examination and Logical Memory scores for entry into Alzheimer's disease trials. Alzheimers Res Ther 2016 Feb 22;8:9 [FREE Full text] [doi: 10.1186/s13195-016-0176-z] [Medline: 26899835]

XSL•FO
RenderX

17. Belleville S, Fouquet C, Hudon C, Zomahoun HT, Croteau J, Consortium for the Early Identification of Alzheimer's disease-Quebec. Neuropsychological measures that predict progression from mild cognitive impairment to Alzheimer's type dementia in older adults: a systematic review and meta-analysis. Neuropsychol Rev 2017 Dec;27(4):328-353 [FREE Full text] [doi: 10.1007/s11065-017-9361-5] [Medline: 29019061]
18. Lehr M, Prud'hommeaux E, Shafran I, Roark B. Fully automated neuropsychological assessment for detecting mild cognitive impairment. In: Proceedings of the INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association. 2012 Presented at: INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association; Sep 9-13, 2012; Portland, Oregon, USA. [doi: 10.21437/interspeech.2012-306]
19. Ferrar J, Griffith GJ, Skirrow C, Cashdollar N, Taptiklis N, Dobson J, et al. Developing digital tools for remote clinical research: how to evaluate the validity and practicality of active assessments in field settings. J Med Internet Res 2021 Jun 18;23(6):e26004 [FREE Full text] [doi: 10.2196/26004] [Medline: 34142972]
20. Ousset PJ, Vellas B. Viewpoint: impact of the Covid-19 outbreak on the clinical and research activities of memory clinics: an Alzheimer's disease center facing the Covid-19 crisis. J Prev Alzheimers Dis 2020;7(3):197-198 [FREE Full text] [doi: 10.14283/jpad.2020.17] [Medline: 32463074]
21. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Sciences Policy, Forum on Drug Discovery, Development, and Translation. Virtual Clinical Trials Challenges and Opportunities: Proceedings of a Workshop. Washington (DC): National Academies Press (US); 2019.
22. Yanez ND, Weiss NS, Romand J, Treggiari MM. COVID-19 mortality risk for older men and women. BMC Public Health 2020 Nov 19;20(1):1742 [FREE Full text] [doi: 10.1186/s12889-020-09826-8] [Medline: 33213391]
23. Sano M, Lapid MI, Ikeda M, Mateos R, Wang H, Reichman WE. Psychogeriatrics in a world with COVID-19. Int Psychogeriatr 2020 Oct;32(10):1101-1105 [FREE Full text] [doi: 10.1017/S104161022000126X] [Medline: 32613925]
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. Alzheimers Dement 2011 May;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
25. 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]
26. 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]
27. McNeil MR, Doyle PJ, Fossett TR, Park GH, Goda AJ. Reliability and concurrent validity of the information unit scoring metric for the story retelling procedure. Aphasiology 2010 Sep 10;15(10-11):991-1006. [doi: 10.1080/02687040143000348]
28. Wechsler D. A Standardized Memory Scale for Clinical Use. The Journal of Psychology 1945 Jan;19(1):87-95. [doi: 10.1080/00223980.1945.9917223]
29. WAIS-III WMS-III Technical Manual (Wechsler Adult Intelligence Scale & Wechsler Memory Scale). San Diego, California, United States: Harcourt Brace & Company; 1997.
30. Wechsler D. WAIS-R Manual Wechsler Adult Intelligence Scale-revised. San Antonio, TX: Psychological Corporation; 1981.
31. Grober E, Ocepek-Welikson K, Teresi JA. The Free and Cued Selective Reminding Test: evidence of psychometric adequacy. Psychol Sci Q 2009;51(3):266-282 [FREE Full text]
32. Speech-to-Text. Google Cloud. URL: https://cloud.google.com/speech-to-text [accessed 2022-06-07]
33. Lhoest Q, del MA, Jernite Y, Thakur A, von PP, Patil S. Datasets: a community library for natural language processing. arXiv 2021. [doi: 10.18653/v1/2021.emnlp-demo.21]
34. Devlin J, Change M, Lee K, Toutanova K. BERT: pre-training of deep bidirectional Transformers for language understanding. arXiv 2019. [doi: 10.1090/mbk/121/79]
35. Rice ME, Harris GT. Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. Law Hum Behav 2005 Oct;29(5):615-620. [doi: 10.1007/s10979-005-6832-7] [Medline: 16254746]
36. Larner AJ. Effect size (Cohen's d) of cognitive screening instruments examined in pragmatic diagnostic accuracy studies. Dement Geriatr Cogn Dis Extra 2014 May;4(2):236-241 [FREE Full text] [doi: 10.1159/000363735] [Medline: 25177332]
37. Edgar CJ, Siemers E, Maruff P, Petersen RC, Aisen PS, Weiner MW, Alzheimer's Disease Neuroimaging Initiative. Pilot evaluation of the unsupervised, at-home cogstate brief battery in ADNI-2. J Alzheimers Dis 2021;83(2):915-925 [FREE Full text] [doi: 10.3233/JAD-210201] [Medline: 34366337]
38. Gavett BE, Gurnani AS, Saurman JL, Chapman KR, Steinberg EG, Martin B, et al. Practice effects on story memory and list learning tests in the neuropsychological assessment of older adults. PLoS One 2016;11(10):e0164492 [FREE Full text] [doi: 10.1371/journal.pone.0164492] [Medline: 27711147]
39. Duff K, Alzheimer's Disease Neuroimaging Initiative. Amnestic MCI in ADNI: maybe not enough memory impairment? Neurology 2021 Aug 02:595-596. [doi: 10.1212/WNL.0000000000012587] [Medline: 34341152]
40. Gross AL, Inouye SK, Rebok GW, Brandt J, Crane PK, Parisi JM, et al. Parallel but not equivalent: challenges and solutions for repeated assessment of cognition over time. J Clin Exp Neuropsychol 2012;34(7):758-772 [FREE Full text] [doi: 10.1080/13803395.2012.681628] [Medline: 22540849]

XSL•FO
RenderX

41. Papp KV, Samaroo A, Chou H, Buckley R, Schneider OR, Hsieh S, et al. Unsupervised mobile cognitive testing for use in preclinical Alzheimer's disease. Alzheimers Dement (Amst) 2021;13(1):e12243 [FREE Full text] [doi: 10.1002/dad2.12243] [Medline: 34621977]
42. Mackin RS, Rhodes E, Insel PS, Nosheny R, Finley S, Ashford M, et al. Reliability and validity of a home-based self-administered computerized test of learning and memory using speech recognition. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2021 Jun 17:1-15. [doi: 10.1080/13825585.2021.1927961] [Medline: 34139954]
43. Busch RM, Hogue O, Ferguson L, Parsons MW, Kubu CS, Floden DP. Validation of computerized episodic memory measures in a diverse clinical sample referred for neuropsychological assessment. Clin Neuropsychol 2019 Apr;33(3):557-570 [FREE Full text] [doi: 10.1080/13854046.2018.1488995] [Medline: 29996710]
44. Sabbagh MN, Boada M, Borson S, Chilukuri M, Doraiswamy PM, Dubois B, et al. Rationale for early diagnosis of Mild Cognitive Impairment (MCI) supported by emerging digital technologies. J Prev Alzheimers Dis 2020;7(3):158-164. [doi: 10.14283/jpad.2020.19] [Medline: 32463068]
45. Martínez-Nicolás I, Llorente TE, Martínez-Sánchez F, Meilán JJ. Ten years of research on automatic voice and speech analysis of people with Alzheimer's disease and mild cognitive impairment: a systematic review article. Front Psychol 2021;12:620251 [FREE Full text] [doi: 10.3389/fpsyg.2021.620251] [Medline: 33833713]
46. Roark B, Mitchell M, Hosom J, Hollingshead K, Kaye J. Spoken language derived measures for detecting mild cognitive impairment. IEEE Trans Audio Speech Lang Process 2011 Sep 01;19(7):2081-2090 [FREE Full text] [doi: 10.1109/TASL.2011.2112351] [Medline: 22199464]
47. Foldi NS. Getting the hang of it: preferential gist over verbatim story recall and the roles of attentional capacity and the episodic buffer in Alzheimer disease. J Int Neuropsychol Soc 2011 Jan;17(1):69-79. [doi: 10.1017/S1355617710001165] [Medline: 21122189]
48. Mueller KD, Koscik RL, Du L, Bruno D, Jonaitis EM, Koscik AZ, et al. Proper names from story recall are associated with beta-amyloid in cognitively unimpaired adults at risk for Alzheimer's disease. Cortex 2020 Oct;131:137-150 [FREE Full text] [doi: 10.1016/j.cortex.2020.07.008] [Medline: 32861209]
49. Drummond C, Coutinho G, Fonseca RP, Assunção N, Teldeschi A, de Oliveira-Souza R, et al. Deficits in narrative discourse elicited by visual stimuli are already present in patients with mild cognitive impairment. Front Aging Neurosci 2015;7:96 [FREE Full text] [doi: 10.3389/fnagi.2015.00096] [Medline: 26074814]

Abbreviations

AD: Alzheimer disease
ASRT: Automatic Story Recall Task
CDR: Clinical Dementia Rating scale
CDR-SB: Clinical Dementia Rating scale–Sum of Boxes
CU: cognitively unimpaired
G-match: generalized match
LM: Logical Memory
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
PACC5: Preclinical Alzheimer's Cognitive Composite with semantic processing
WER: word error rate

Edited by T Leung, J Wang; submitted 07.02.22; peer-reviewed by R Au, G Lui; comments to author 04.06.22; revised version received 07.07.22; accepted 13.07.22; published 30.09.22.

Please cite as:
Skirrow C, Meszaros M, Meepegama U, Lenain R, Papp KV, Weston J, Fristed E
Validation of a Remote and Fully Automated Story Recall Task to Assess for Early Cognitive Impairment in Older Adults: Longitudinal Case-Control Observational Study
JMIR Aging 2022;5(3):e37090
URL: https://aging.jmir.org/2022/3/e37090
doi:10.2196/37090
PMID:36178715

©Caroline Skirrow, Marton Meszaros, Udeepa Meepegama, Raphael Lenain, Kathryn V Papp, Jack Weston, Emil Fristed. Originally published in JMIR Aging (https://aging.jmir.org), 30.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted

XSL•FO
RenderX

use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluating Web-Based Automatic Transcription for Alzheimer Speech Data: Transcript Comparison and Machine Learning Analysis

Thomas Soroski[1], BSc; Thiago da Cunha Vasco[2]; Sally Newton-Mason[1], BSc; Saffrin Granby[2], BSc; Caitlin Lewis[1], BA; Anuj Harisinghani[2], BSc; Matteo Rizzo[2], BSc, MSc; Cristina Conati[2], MSc, PhD; Gabriel Murray[3], MSc, PhD; Giuseppe Carenini[2], MSc, PhD; Thalia S Field[1], MHSc, MD; Hyeju Jang[2], MS, PhD

[1]Vancouver Stroke Program and Division of Neurology, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

[2]Department of Computer Science, Faculty of Science, University of British Columbia, Vancouver, BC, Canada

[3]School of Computing, University of the Fraser Valley, Abbotsford, BC, Canada

Corresponding Author:
Hyeju Jang, MS, PhD
Department of Computer Science
Faculty of Science
University of British Columbia
201-2366 Main Mall
Vancouver, BC, V6T 1Z4
Canada
Phone: 1 604 822 3061
Email: hyejuj@cs.ubc.ca

Abstract

Background: Speech data for medical research can be collected noninvasively and in large volumes. Speech analysis has shown promise in diagnosing neurodegenerative disease. To effectively leverage speech data, transcription is important, as there is valuable information contained in lexical content. Manual transcription, while highly accurate, limits the potential scalability and cost savings associated with language-based screening.

Objective: To better understand the use of automatic transcription for classification of neurodegenerative disease, namely, Alzheimer disease (AD), mild cognitive impairment (MCI), or subjective memory complaints (SMC) versus healthy controls, we compared automatically generated transcripts against transcripts that went through manual correction.

Methods: We recruited individuals from a memory clinic ("patients") with a diagnosis of mild-to-moderate AD, (n=44, 30%), MCI (n=20, 13%), SMC (n=8, 5%), as well as healthy controls (n=77, 52%) living in the community. Participants were asked to describe a standardized picture, read a paragraph, and recall a pleasant life experience. We compared transcripts generated using Google speech-to-text software to manually verified transcripts by examining transcription confidence scores, transcription error rates, and machine learning classification accuracy. For the classification tasks, logistic regression, Gaussian naive Bayes, and random forests were used.

Results: The transcription software showed higher confidence scores (P<.001) and lower error rates (P>.05) for speech from healthy controls compared with patients. Classification models using human-verified transcripts significantly (P<.001) outperformed automatically generated transcript models for both spontaneous speech tasks. This comparison showed no difference in the reading task. Manually adding pauses to transcripts had no impact on classification performance. However, manually correcting both spontaneous speech tasks led to significantly higher performances in the machine learning models.

Conclusions: We found that automatically transcribed speech data could be used to distinguish patients with a diagnosis of AD, MCI, or SMC from controls. We recommend a human verification step to improve the performance of automatic transcripts, especially for spontaneous tasks. Moreover, human verification can focus on correcting errors and adding punctuation to transcripts. However, manual addition of pauses is not needed, which can simplify the human verification step to more efficiently process large volumes of speech data.

(JMIR Aging 2022;5(3):e33460) doi:10.2196/33460

XSL•FO
RenderX

KEYWORDS

Alzheimer disease; mild cognitive impairment; speech; natural language processing; speech recognition software; machine learning; neurodegenerative disease; transcription software; memory

Introduction

Identifying individuals with Alzheimer disease (AD) and mild cognitive impairment (MCI) early is beneficial for patient care, family support, and resource planning for the health care system [1]. Identification of individuals who are in the earliest stages of neurodegenerative disease, before irreversible brain changes have occurred, may also allow for the use of disease-modifying therapies when they would be most effective [2].

Analysis of speech to aid in the identification of individuals with early neurodegenerative disease can be a promising strategy, as speech recording is noninvasive, scalable, and easily repeated over time. This contrasts with the current methods for screening for AD or MCI, such as nuclear medicine scans or spinal fluid analysis, which can be both expensive and invasive [3]. Short samples of spontaneous or prompted speech can be collected remotely by telephone or videoconference. To date, speech and language have shown promising results in a significant number of studies aiming to classify AD or MCI [4].

For AD classification using speech, transcription is a key step to leverage the wealth of information contained in lexical data [5,6]. DementiaBank [7], the largest cohort of MCI and AD speech data for research, is entirely manually transcribed. Manual transcription, while highly accurate, is very low throughput (eg, requiring 4 minutes of transcriber time for each minute of audio [8]), limiting the potential scalability and cost savings associated with language-based screening for MCI and AD. As a result, there is a move toward automatically preprocessing medical speech as opposed to manual transcription.

To date, some groups have investigated AD/MCI classification using only automatically generated transcripts produced by transcription software [9,10]. While automatic transcription allows high-throughput speech transcription for a very low cost per sample, these systems can vary in their accuracy (ranging from 68% to 87% in past work [11]), which may affect the performance of downstream linguistic analysis [12]. Furthermore, the impact of automatic preprocessing on classification is not fully understood and should be investigated before continuing downstream investigations.

To better understand the use of automatic transcription for AD/MCI classification, we compared the automatically generated transcripts from Google speech-to-text [13] (“automatic transcripts”) against automatic transcripts that went through a second stage of manual correction (“manually corrected transcripts”). These manually corrected transcripts were used as ground truth.

Specifically, we first examined a confidence metric in the transcription software for transcribing speech recordings from memory clinic patients versus healthy controls. Second, we measured the word-level accuracy of the automatic transcripts against ground truth. Third, we compared classification performances of machine learning models using data from automatic versus manually-corrected transcripts. Based on these results, we discuss accuracy trade-offs associated with manual transcript verification in the context of dementia classification, and we suggest more efficient manual verification methods to improve the performance of automatically generated transcripts.

This investigation aims to highlight differences in human versus automatically processed transcripts to drive future automatic transcription–based research. Therefore, we focus here on comparing transcription methods using existing machine learning algorithms rather than building a novel model that outperforms state-of-the-art models.

This work has 4 main contributions addressing knowledge gaps in the existing literature. First, we evaluate automatic transcription and manual transcription on a data set of older adults for AD/MCI classification using 3 measures: transcription confidence, error rates, and machine learning classification accuracy. To our knowledge, this approach for evaluating transcriptions has not been used previously.

Second, our investigation is novel in that we are exploring the robustness of automatic transcription in a cohort of older adults, including those with cognitive impairment and dementia. The aging process includes changes to voice and speech (eg, presbyphonia, word-finding difficulties), which may affect automatic transcription. However, previous investigations on transcription methods have focused solely on younger or heterogeneous cohorts [12,14]. To our knowledge, this is the first investigation on the impact of transcription methods in a cohort of older adults.

Third, based on the evaluation results, we make practical suggestions about how to use automatic transcription. These suggestions will help researchers to better leverage automatic transcription for building natural language processing (NLP)–based screening methods using large data sets for AD/MCI or subjective memory complaints (SMC), which can be a prodromal state for MCI and AD [15].

Finally, while our results are generated with an AD/MCI data set, our findings could also be extrapolated to other neurological and psychiatric conditions where speech analysis is being investigated as a classification tool. This includes stroke [16], Parkinson disease [17], concussion [18], anxiety [19], bipolar disorder [20], depression, and suicidal ideation [21,22].

Methods

Overview

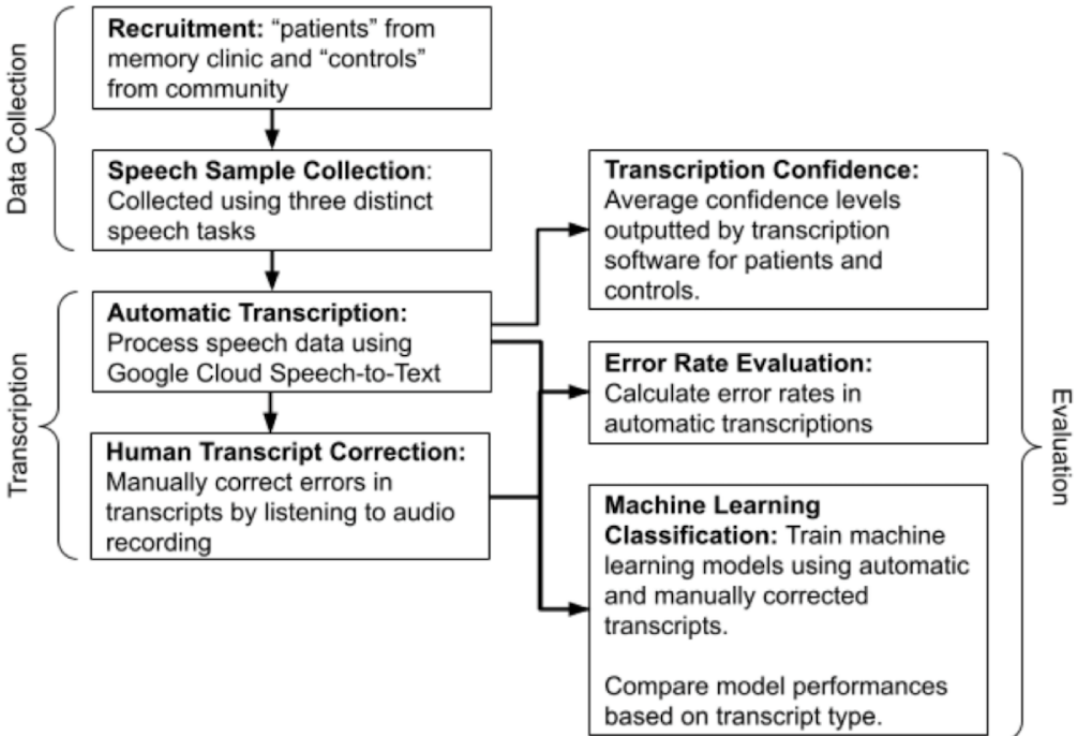
This study involved 3 main phases: (1) data collection, (2) transcription, and (3) evaluation. Our workflow is summarized in Figure 1. As part of a larger study examining machine learning algorithms for classification of memory clinic patients versus healthy controls, we recruited participants with a clinical diagnosis of mild-to-moderate AD, MCI, or SMC (“patients”)

XSL•FO
RenderX

from a subspecialty memory clinic and healthy volunteer controls from the community. Participants underwent a test battery that included describing the "Cookie Theft" picture from the Boston Diagnostic Aphasia Examination, a reading task incorporating a sixth-grade level paragraph from the International Reading Speed Texts (IReST), and recounting a pleasant past experience. Their speech was recorded, and we used Google Cloud speech-to-text (STT) to automatically transcribe speech data. We then manually corrected errors in the automatic transcripts.

For evaluation, we first aggregated transcription confidence levels provided by the software to determine whether transcription software confidence levels vary between patients and controls. Using manually corrected transcripts as the gold standard, we calculated the error rate of automatic transcripts. Then, we compared the performance of machine learning models trained with either automatic or manually corrected transcripts in classifying transcripts as belonging to "patients" versus "controls."

Figure 1. Diagram of our methods and process.



Data Collection

Recruitment

Patients were recruited from a memory clinic in British Columbia, Canada, and diagnosed with AD, MCI, or SMC. Control participants were recruited from the community, with efforts made to age- and sex-match patient participants. All participants were conversationally fluent in English, could engage in a spontaneous conversation, and were aged 50 or older (mean 68.8, SD 9.5 years). Clinic patients were excluded if they had psychiatric medication changes under 18 months ago or neurological conditions other than SMC, MCI, or AD. We report data from 72 memory clinic patients, of which 44 (30%) were diagnosed with mild-to-moderate AD, 20 (13%) with MCI, and 8 (5%) with SMC (mean age 71.9, SD 8.9 years), along with 77 (52%) healthy volunteers (mean age 65.7, SD 9.1 years).

Diagnoses were made by specialist clinicians using standard-of-care guidelines. The diagnostic process involves a combination of cognitive testing, neuroimaging, laboratory data, medical history, physical exam, and collateral information collected from individuals close to the patient.

Speech Sample Collection

Participants underwent a 10-minute computer-based battery. They were asked to complete a total of 3 speech tasks while their voice was recorded. Participants described the Cookie Theft photo [23], read a standardized paragraph from the IReST, and recalled a pleasant past experience. All tasks were carried out in English. During these spontaneous speech tasks, the audio was recorded using the Logitech C922x ProStream webcam. The Cookie Theft picture description task is a validated spontaneous speech task used extensively in prior work for AD/MCI classification [6,24-26]. This task has also been used for predicting the future risk of developing AD in cognitively normal individuals [27].

For the reading task, a single paragraph was selected from IReST, a collection of short paragraphs (<200 words) designed to be readable at a sixth-grade level [28]. To recreate a natural reading environment such as a book or newspaper, the entire paragraph was presented on the screen at the same time rather than displaying each sentence individually, as in some other investigations [29]. For the final task, participants were asked to describe a pleasant past experience ("experience description task"). Several examples were given to participants prior to

XSL•FO
RenderX

starting the task, such as their first pet, how they met their best friend, or a place they had traveled.

Automatic Transcription

Following the speech tasks, participant audio data was labeled with a unique anonymized identifier and converted to the Waveform audio file format. Next, participant audio was uploaded to the Google Cloud STT platform using US English and 16000 Hz settings, with word-level time stamps enabled, to output the automatic transcripts.

Each transcribed word was labeled as being within a specific task or as being extraneous from all tasks. Words spoken outside of tasks were removed in downstream experiments.

Human Transcript Correction

After automatic transcript files were generated, human transcribers listened to the recorded audio files and made manual corrections to the transcripts based on the recorded audio. This manual transcription involved 3 steps: fixing transcription errors, adding punctuation, and adding filled pauses and silent pause annotations.

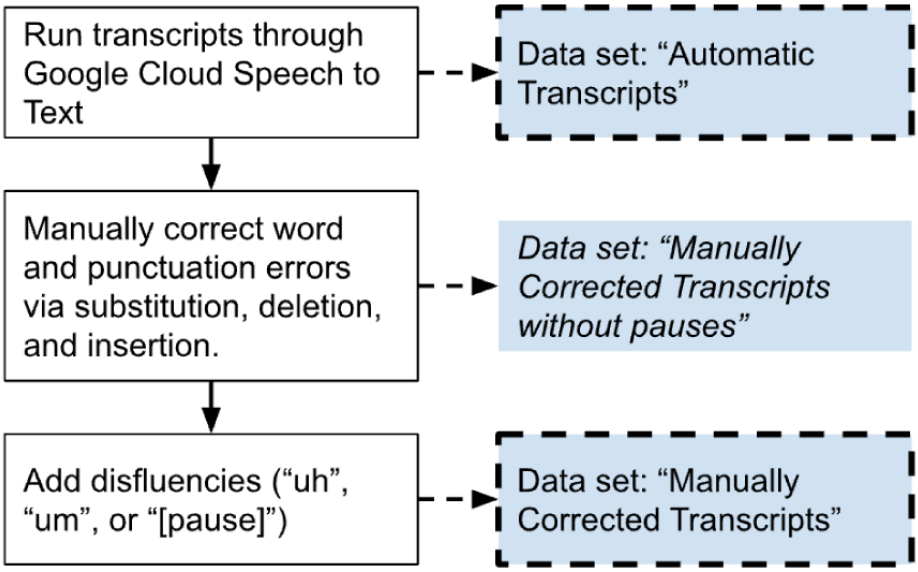
For the first step, which involved fixing transcription errors**,** human transcribers manually substituted incorrectly transcribed words (eg, change "cookie far" to "cookie jar"), inserted missed words (eg, change "cookie" to "cookie jar"), and deleted extra words (eg, change "cookie key jar" to "cookie jar").

The second step entailed adding punctuation. While Google STT adds punctuation, it is very rare, with some transcripts having as few as 0 automatically added punctuation marks. As NLP preprocessing (eg, parsing) benefits from fully formed sentences, human transcribers manually added punctuation (ie, ".", "!", and "?") to the transcripts.

For the third step, which consisted of adding filled pauses and silent pause annotations, human transcribers manually added both filled and silent pauses. A filled pause was considered to be any utterance of "uh" or "um." Filled pauses were consistently transcribed as "uh" or "um" regardless of the length of the pause. Silent pauses were specially labeled as "[pause]" to distinguish this from the word "pause." Silent pauses were considered to be any break or silence in speech for 0.25 seconds or longer, following Goldman-Eisler [30] and Park [31]. Instances where the participant was not speaking but was not silent were not labeled as a pause (eg, coughing or laughing). The duration of pauses was not differentiated.

Figure 2 summarizes the transcription process. Acoustic data were transcribed with Google Cloud STT to generate "automatic transcripts." Then, human transcribers fixed spoken words and added punctuation based on the audio recording to generate "manually corrected transcripts without pauses." Finally, human transcribers manually added both filled and silent pauses to generate the "manually corrected transcripts" data set.

Figure 2. Diagram illustrating how the 3 different transcript data sets were generated.



Ethics Approval

This study was approved by the University of British Columbia Clinical Research Ethics Board (H17-02803). All participants provided their written informed consent prior to participating in this study. Baseline demographic characteristics of the patients and controls are summarized in Multimedia Appendix 1.

Evaluation

Transcription Confidence

For a given audio clip, Google STT outputs transcribed words and a confidence level between 0 and 1. This is calculated by aggregating the likelihood values assigned to each word in the audio. A higher number indicates that the words were more likely to be transcribed accurately. We used these confidence levels to determine whether transcription software confidence levels vary between patients and controls and to determine if patient speech was more difficult to transcribe than control speech.

XSL•FO
RenderX

Error Rate Evaluation

To examine the error rate of automatic transcripts, we compared these to manually corrected transcripts without pauses. We chose not to include pauses because automatic transcripts do not transcribe pauses at all; thus, not denoting a pause would not be considered an error.

We calculated standard measures of transcription accuracy, including word error rate (WER) and match error rate (MER) [32], using a Python package, JiWER (v2.1.0, Vassen [33]). These metrics take into account the number of substitutions (eg, “cookie far” to “cookie jar”), deletions (eg, “cookie key jar” to “cookie jar”), and insertions (eg, “cookie” to “cookie jar”) in the manually corrected transcript.

WER represents the rate of errors to the number of input words. This is calculated as follows:



WER does not weigh insertions and deletions equally. For example, a 6-word transcript with 30 insertion errors has a WER value of 5, while a 36-word transcript with 30 deletion errors has a WER of 0.83.

MER represents the probability of a given word match being incorrect and is calculated as follows:



For example, a MER of 0.25 means that 1 out of 4 word matches between the manually corrected transcript and automatic transcript will be an error. MER is calculated similarly to WER. However, MER takes into account the maximum number of words between both the automatic and manually edited transcripts, as opposed to only the number of words in only the automatic transcript. MER also weighs insertions and deletions equally.

WER and MER were calculated for each individual transcript. Then, the average and standard deviation of these values were calculated for patients and controls and for each task (eg, picture description, reading, and experience description tasks).

Machine Learning Classification

To determine whether manual correction impacts machine learning classification of patients versus controls, we performed experiments using both the automatic and manually corrected transcript data sets.

Table 1 outlines the entire feature set by task. For the picture description task and the experience description task, we extracted features from transcripts following the text-based features in previous work [6,34]. These features are based on grammar rules, vocabulary, or psycholinguistics. For the experience description task, we did not include information units used for the picture description task, each of which correspond to visual features in the Cookie Theft picture, such as cookie, jar, boy, or girl.

Table 1. Features for machine learning classification models.

Task	Feature groups and number of features (n) in each group
Picture description	• Cookie Theft image information units (13) • Part-of-speech (15), context-free-grammar rules (44), syntactic complexity (24), vocabulary richness (4), psycholinguistic (5), repetitiveness (5)
Reading	• Syllable count (1), pause count (1)[a], total duration (1), total time spent speaking (1), proportion of time spent speaking (1), speech rate (1), average syllable duration (1), pauses per syllable (1)[a], pause rate (1)[a], pause duration (3)[a]
Experience description	• Part-of-speech (15), context-free-grammar rules (44), syntactic complexity (24), vocabulary richness (4), psycholinguistic (5), repetitiveness (5)

[a]These features were computed using acoustic data and transcript data and are also affected by method of pause detection (ie, acoustic vs text data).

For the reading task, we used 12 reading-task–specific features based on the work of Fraser et al [35]. Extracting text features from reading task data may be counterintuitive because each participant reads an identical prompt. However, transcripts may contain repeated words, incorrectly read words, or filled pauses, making transcribed text features potentially informative. Since automatic transcripts do not contain pause information, we first compared automatic transcripts and manually corrected transcripts by using acoustic data to detect unfilled pauses. As an additional comparison for the reading task, we compared using unfilled pauses detected from audio and using unfilled pauses annotated in manually corrected transcripts to determine whether manually adding pauses to transcripts is useful for the reading task or not.

To parse text data and tag parts of speech, we used Stanford CoreNLP [36]. Psycholinguistic features were generated using the MRC database [37], which provides concreteness, familiarity, and imageability scores of English words. Pauses in the reading task were detected using pydub (v0.25.1 [38]), a Python audio processing package. Syllables were detected using Syllables (v1.0.3 [39]), a Python package.

Based on these features, we performed binary classification to distinguish patients from controls. We chose to perform binary classification due to the data size. The number of data samples for finer classes (MCI and SMC) was too small for multiclass classification. To investigate the usefulness of manual correction, we first compared the performance of automatic to manually corrected transcripts. To determine the importance of pause annotation we compared the performance of manually corrected transcripts with and without pauses.

XSL•FO
RenderX

We tested with 3 classification algorithms that have shown best performances in previous work on dementia classification [40]: logistic regression (LR), random forest (RF), and Gaussian naive Bayes (GNB). In addition, we tested with an end-to-end fine-tuned pretrained model using Bidirectional Encoder Representations from Transformers (BERT) [41] for the picture description and experience description tasks. Note that we did not try BERT models for the reading task because participants read the same text. We used the Python package scikit-learn (v0.19.1 [42]) to perform classification. We used a stratified 10-fold cross-validation approach and repeated this process 10 times in total with differently stratified splits, each generated with a unique random seed. We report the classification performance in terms of area under the receiver operating characteristic curve (AUROC). AUROC is an evaluation metric for classification at various threshold settings and is commonly used for evaluating diagnostic accuracy [43]. The performance metric was averaged over the 10 folds and 10 runs. To remove highly pairwise correlated features and features poorly correlated with the label, we performed correlation feature selection [44]. Highly correlated features were defined as having a Pearson correlation coefficient greater than 0.85, while poorly correlated features had a Pearson correlation coefficient less than 0.20.

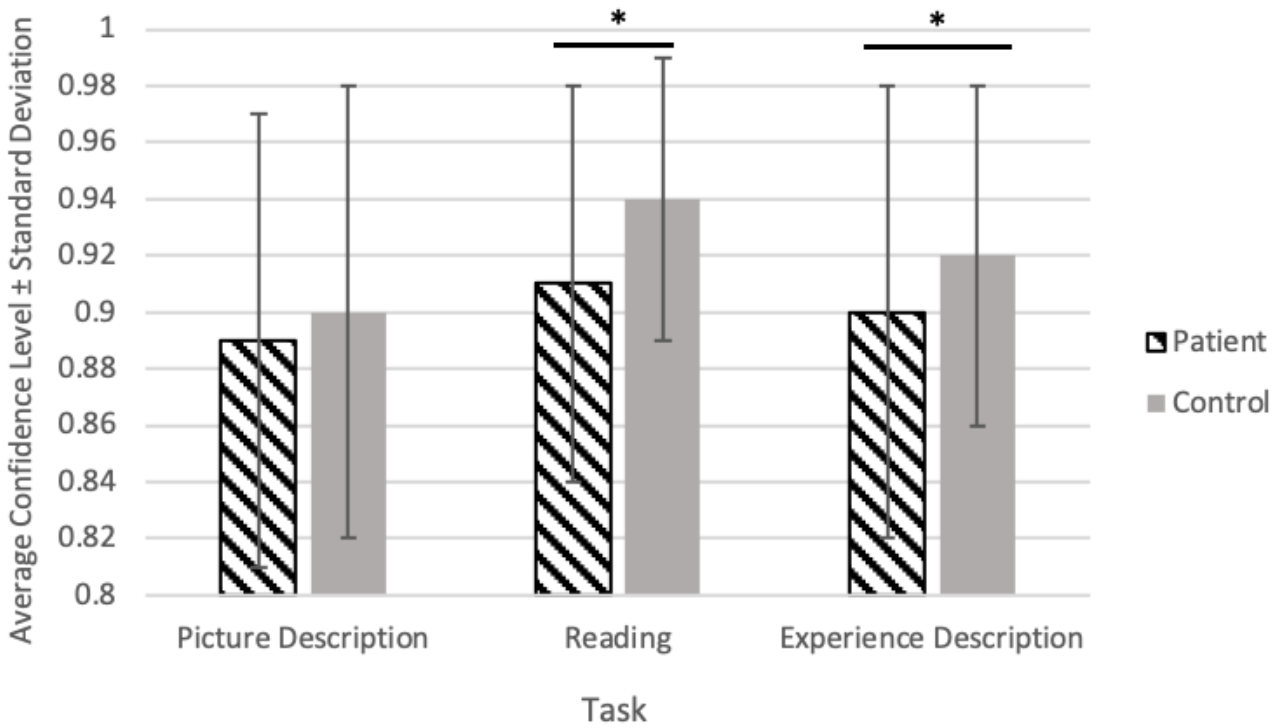
We performed a statistical analysis on the model results to determine if the different transcript data sets led to significant changes in model performance. For each classification algorithm for a given task, we ran a double-sided *t*-test using the null hypothesis that the mean AUROC was no different for automatic and manually corrected transcripts.

Results

Transcription Confidence Results

Google confidence level results are shown in Figure 3. Generally, Google STT produced a higher confidence level when transcribing audio from controls. In the reading task, for example, the average confidence level was 0.94 (SD 0.05) for controls, compared to 0.91 (SD 0.07) for patients. Both the reading and experience description tasks showed a significantly higher confidence level for controls than patients.

Figure 3. Google speech-to-text confidence results. Error bars represent the standard deviation. * represents $P<.001$, calculated by *t*-test.



Error Rate Evaluation Results

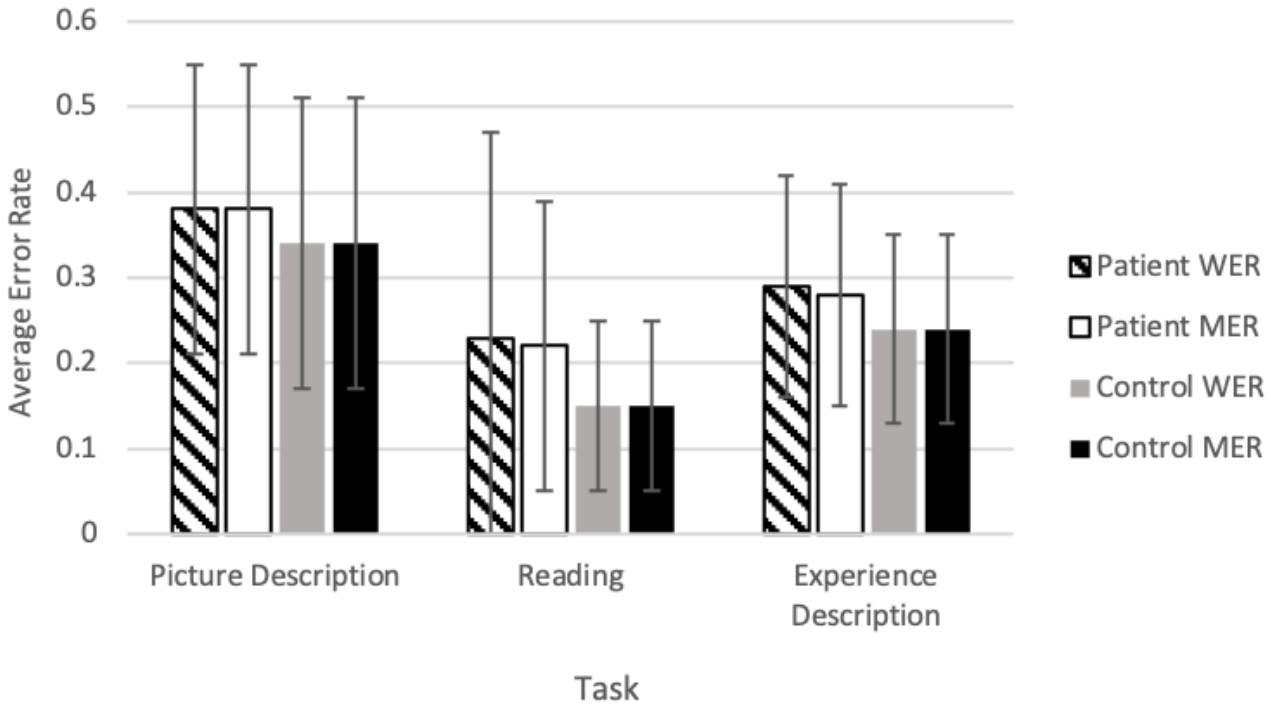
Figure 4 shows the error rate results. In general, automatic transcription had a lower error rate when transcribing control speech compared to patient speech, as shown by the lower average WER and MER.

The reading task was the most accurate overall, showing an average MER of 0.15 (SD 0.10) for controls and 0.22 (SD 0.19) for patients. This could be because people tend to enunciate when they are asked to read a text aloud. WER and MER results were largely similar overall, suggesting that there were not disproportionately high rates of insertion errors. In other words, manual correction did not involve more word addition as opposed to word deletion or word substitution.

The picture description task was found to have the highest error rate overall when compared to the reading and experience description tasks. This indicates more manual corrections or poorer accuracy of automatic transcription, but it is not clear why this is the case.

XSL•FO
RenderX

Figure 4. Average error rates by task and participant type. Error bars represent the standard deviation. There were no significant differences in error rates between or within tasks. MER: match error rate; WER: word error rate.



Machine Learning Model Results

Models trained on manually corrected transcripts from the picture description and experience description tasks significantly outperformed models trained on automatic transcripts (Table 2). However, there was no significant difference in model performance trained using either transcription method from the reading task. This finding was true regardless of whether pause-related features were included or not (Table 3).

Table 4 shows results of the models using manually corrected transcripts, with and without pauses for the picture description and experience description tasks. There was no clear trend or significant change in any AUROC result when comparing transcripts with and without pauses.

XSL•FO
RenderX

Table 2. Machine learning classification results of models trained on automatic transcripts compared to results of models trained on manually corrected transcripts.

Task and model type	Automatic transcripts AUROC[a]	Manually corrected transcripts AUROC	Change in AUROC[b]
Picture description			
RF[c]	0.617	0.687	0.070[d]
GNB[e]	0.662	0.725	0.063[d]
LR[f]	0.671	0.743	0.072[d]
BERT[g]	0.618	0.686	0.068[d]
Experience description			
RF	0.503	0.636	0.133[d]
GNB	0.549	0.677	0.128[d]
LR	0.543	0.674	0.131[d]
BERT	0.630	0.650	0.020[d]

[a]AUROC: area under the receiver operating characteristic curve.

[b]Positive change in AUROC indicates that the manually corrected transcript model outperformed the automatic transcript model.

[c]RF: random forest.

[d]Indicates $P<.001$.

[e]GNB: Gaussian naive Bayes.

[f]LR: logistic regression.

[g]BERT: Bidirectional Encoder Representations from Transformers.

Table 3. Machine learning classification results of models trained on reading task data with pause features computed using acoustic data or computed using text data.

Reading task	(1) Automatic transcripts AUROC[a,b]	(2) Manually corrected transcripts AUROC[b]	(3) Manually corrected transcripts AUROC[c]	Change in AUROC (3)–(1)
RF[d]	0.638	0.655	0.662	0.024
GNB[e]	0.677	0.677	0.693	0.016
LR[f]	0.589	0.587	0.568	−0.021

[a]AUROC: area under the receiver operating characteristic curve.

[b]Pauses detected from acoustic data.

[c]Pauses detected from text data.

[d]RF: random forest.

[e]GNB: Gaussian naive Bayes.

[f]LR: logistic regression.

XSL•FO
RenderX

Table 4. Machine learning classification results of models trained on manually corrected transcripts without pauses compared to results of models trained on manually corrected transcripts (with pauses).

Task and model type	Transcripts without pauses AUROC[a]	Transcripts with pauses AUROC	Change in AUROC[b]
Picture description			
RF[c]	0.666	0.687	0.021
GNB[d]	0.730	0.725	−0.005
LR[e]	0.755	0.743	−0.012
BERT[f]	0.686	0.691	0.005
Experience description			
RF	0.631	0.636	0.005
GNB	0.676	0.677	0.001
LR	0.692	0.674	−0.018
BERT	0.622	0.649	0.027

[a]AUROC: area under the receiver operating characteristic curve.

[b]Positive change in AUROC indicates that the pause model outperformed the no-pause model.

[c]RF: random forest.

[d]GNB: Gaussian naive Bayes.

[e]LR: logistic regression.

[f]BERT: Bidirectional Encoder Representations from Transformers.

Discussion

Transcription Confidence

The transcription confidence results showed that the automatic transcription software was consistently more confident in transcribing the speech of controls compared to patients. This may indicate that patient speech differs from the speech used to train the automatic transcription software (which was likely trained using speech from a more general population, including younger or cognitively unimpaired individuals). This may be attributed to the fact that people with Alzheimer disease often have impaired speech production [45], such as distortions (eg, “ook” instead of “cookie”) and phonological paraphasias (eg, “tid” instead of “kid”) [46]. It is especially interesting that the confidence difference between the 2 groups was highest and most significant for the reading task. This confirms that reading task speech is effective for distinguishing AD/MCI patients from controls, as also shown in prior work [35,47,48].

Error Rate Evaluation

Automatic transcriptions were more accurate for healthy controls compared to patients with AD or MCI, as shown by higher error rate and information loss in patient transcripts. This result is logical in the context of the confidence result, as patient transcripts had significantly lower confidence, meaning that the transcription software was more unsure about its output.

Our results are markedly different from Google’s reports on the error rates of their own software (Google Cloud STT has not disclosed the composition of their training data set). According to Google, their transcription program achieved a WER of 6.7% using 12,500 hours of voice search data and a WER of 4.1% in a dictation task [49]. In contrast, for spontaneous speech tasks, we found a WER range of 24% to 34% for controls and 29% to 38% for patients. The reading task showed a lower WER of 15% for controls and 23% for patients.

While our error rate results differ from Google’s reported results, they are comparable to the results of other investigations using Google STT derived from simulated medical encounters. Kim et al [50] used audio data from 12 simulated patient and medical student interactions. In this investigation, Google STT showed an average WER of 34%, similar to our WER result of 34% for controls and 38% for patients completing the picture description task. Miner et al [14] recorded audio from 100 patients aged 18-52 (mean age 23) during therapy sessions and found that Google STT had an average WER of 25%. This result is comparable to the WER for our experience description task, which was 29% for patients and 24% for controls. Both therapy-related discussion and the experience description tasks typically involve spontaneous speech with minimal prompting.

Surprisingly, the experience description task showed lower error rates than the picture description task. This might be because Google STT repeatedly transcribed certain phrases or words related to the picture incorrectly across participants, leading to a higher average error rate in this task. It is also possible that the experience description is easier for automatic transcription because it is more conversational, like the material Google STT may have been trained on. Further investigation into to the discrepancy in performance between different spontaneous speech tasks is warranted.

The reading task WER for our cohort was notably higher than previous research. Kepuska et al [51] used Google STT to transcribe audio from 630 speakers reading 10 sentences each and found an average WER of 9%. This is markedly lower than the results of our investigation, in which we found that the 9-sentence reading task produced a WER of 23% for patients

XSL•FO
RenderX

and 15% for controls. One possible reason for this large disparity is that Google STT is not specific to a particular population (eg, older adults may experience normal age-related changes to the larynx and vocal cords over time, known as presbyphonia) and may produce more accurate transcriptions in a more generalized sample.

Machine Learning Models

Despite our previous results showing that automatic transcription for our data set is more inaccurate than values reported by Google, our machine learning model results still show that automatic transcripts are discriminative for AD/MCI. Other studies using automatic transcription for classification experiments have noted that inaccuracies or errors in the audio-to-text transcription do not necessarily affect classification results [52].

However, manually correcting the picture description and experience description task transcriptions led to significantly higher performances in the machine learning models. By comparison, both automatic and manually corrected reading task transcripts showed similar performance, likely due to the majority of reading features being computed from audio data. To address this concern, we examined text versus audio-based silent pause detection and again found no significant changes in performance. This indicates that using either audio or text to detect pauses will produce similar results and that manually correcting transcripts does not significantly change model performance.

Surprisingly, the addition of filled and silent pauses did not significantly change performance for any of the tasks and algorithms. Moreover, using the pauses from the transcripts showed similar classification results to using pauses detected from audio data for the reading task. Previous studies have shown that people with Alzheimer disease demonstrate a multitude of disfluencies in their speech, including pauses [53-55]. However, manually adding pauses as either words (“um” or “uh”) or tokens (“[pause]”) to transcripts did not seem to have any effect on classification models. This could be because older adults also experience age-related changes in their speech, such as an increase in silent pauses [56], potentially weakening the association of pauses to either the patient or control category. Alternatively, this result may be due to the fact that there are no features that “directly” model pauses for the description tasks, weakening the association of the tasks with pauses.

Limitations

Some limitations with our cohort include varying language ability and variations between transcribers. In our cohort, English was not the first language of 13% of the patients and 21% of the control group, which could potentially contribute to transcription errors. Additionally, our use of 3 different transcriptionists may have introduced interrater variability, especially for more subjective correction steps such as adding punctuation, although variation in manual transcription was controlled via inter-transcriptionist review and protocol development for standardized transcription. Another limitation of our investigation is the size of the data set (N=149), which is quite small for machine learning experiments. However, this is an issue facing most work on using machine learning for dementia classification, especially with newly built data sets (N=55-82) [5,29,35]. While the DementiaBank and ADReSS data sets are larger (N=287 with 687 samples and N=156, respectively), they were originally created in the mid-1980s and are limited by the diagnostic practices of that time. The work described herein aims to mitigate this challenge. Our best practice suggestions for automatic transcription will facilitate data collection at a much faster rate in the future.

It is also important to note that this investigation was completed using Google speech-to-text software in an English-speaking cohort. Competitor speech-to-text software may produce different results, so readers should be wary when applying our conclusions to other software. Applying a similar method to a non-English data set may also produce different results, especially because automatic transcription in other languages might not be as advanced as English. Finally, speech-to-text software is continually being refined and improved. In the future, automatically generated transcripts may be indistinguishable from human-generated transcripts. In the meantime, it is still valuable to understand the impacts of automatic transcription, especially for medical speech data sets.

Conclusion

Our results showed that automatically transcribed speech data from a web-based speech recognition platform can be effectively used to distinguish patients from controls. According to our results, to improve the classification performance of automatically generated transcripts, especially those generated from spontaneous speech tasks, a human verification step is recommended. Our analyses indicate that human verification should focus on correcting errors and adding punctuation to transcripts and that manual addition of pauses is not needed, which can simplify the human verification step to more efficiently process large volumes of speech data.

Acknowledgments

Vancouver Coastal Health Research Institute, Centre for Aging + Brain Health Innovation, Alzheimer’s Society, and the Canadian Consortium on Neurodegeneration in Aging funded research personnel and equipment necessary to recruit participants, collect data, perform analyses, and synthesize results.

XSL•FO
RenderX

Authors' Contributions

TS, SNM, CC, GM, GC, TSF, and HJ contributed to the conception and design of the study. TS, SNM, and CL recruited study participants and administered study assessments. TS organized the database. TdCV, SG, AH, and HJ designed machine features and performed all machine learning analyses. TdCV, AH, and HJ performed the statistical analyses. TS performed other analyses. TS and HJ wrote the first draft of the manuscript. TdCV and AH contributed to sections of the manuscript. All authors contributed to manuscript revision and read and approved the submitted version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary baseline characteristics of patients and controls.
[DOCX File , 13 KB - aging_v5i3e33460_app1.docx]

References

1. Rasmussen J, Langerman H. Alzheimer's disease - why we need early diagnosis. Degener Neurol Neuromuscul Dis 2019;9:123-130 [FREE Full text] [doi: 10.2147/DNND.S228939] [Medline: 31920420]
2. Mahase E. FDA approves controversial Alzheimer's drug despite uncertainty over effectiveness. BMJ 2021 Jun 08;373:n1462. [doi: 10.1136/bmj.n1462] [Medline: 34103308]
3. Watson JL, Ryan L, Silverberg N, Cahan V, Bernard MA. Obstacles and opportunities in Alzheimer's clinical trial recruitment. Health Aff (Millwood) 2014 Apr;33(4):574-579 [FREE Full text] [doi: 10.1377/hlthaff.2013.1314] [Medline: 24711317]
4. Petti U, Baker S, Korhonen A. A systematic literature review of automatic Alzheimer's disease detection from speech and language. J Am Med Inform Assoc 2020 Nov 01;27(11):1784-1797 [FREE Full text] [doi: 10.1093/jamia/ocaa174] [Medline: 32929494]
5. Gosztolya G, Vincze V, Tóth L, Pákáski M, Kálmán J, Hoffmann I. Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features. Comput Speech Lang 2019 Jan;53:181-197. [doi: 10.1016/j.csl.2018.07.007]
6. Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in narrative speech. J Alzheimers Dis 2016;49(2):407-422. [doi: 10.3233/JAD-150520] [Medline: 26484921]
7. Becker JT, Boller F, Lopez OL, Saxton J, McGonigle KL. The natural history of Alzheimer's disease. Description of study cohort and accuracy of diagnosis. Arch Neurol 1994 Jun;51(6):585-594. [doi: 10.1001/archneur.1994.00540180063015] [Medline: 8198470]
8. Bazillon T, Esteve Y, Luzzati D. Manual vs assisted transcription of prepared and spontaneous speech. In: Proceedings of the Sixth International Conference on Language Resources and Evaluation.: European Language Resources Association (ELRA); 2008 May Presented at: LREC'08; May 2008; Marrakech, Morocco URL: https://aclanthology.org/L08-1522/
9. Mirheidari B, Blackburn D, Walker T, Venneri A, Reuber M, Christensen H. Detecting signs of dementia using word vector representations. 2018 Sep Presented at: Interspeech 2018; September 2018; Hyderabad, India p. 1893-1897 URL: https://www.isca-speech.org/archive/interspeech_2018/mirheidari18_interspeech.html [doi: 10.21437/Interspeech.2018-1764]
10. Luz S, Haider F, de la Fuente S, Fromm D, MacWhinney B. Detecting cognitive decline using speech only: the ADReSSo challenge. 2021 Aug Presented at: Interspeech 2021; August 2021; Brno, Czechia p. 3780-3784 URL: https://www.isca-speech.org/archive/interspeech_2021/luz21_interspeech.html [doi: 10.21437/Interspeech.2021-1220]
11. Chlebek P, Shriberg E, Lu Y, Rutowski T, Harati A, Oliveira R. Comparing speech recognition services for HCI applications in behavioral health. In: Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers. USA: Association for Computing Machinery; 2020 Presented at: 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12; Virtual Event p. 483-487 URL: https://dl.acm.org/doi/10.1145/3410530.3414372 [doi: 10.1145/3410530.3414372]
12. Pentland S, Spitzley L, Fuller C, Twitchell D. Data quality relevance in linguistic analysis: the impact of transcription errors on multiple methods of linguistic analysis. 2019 Presented at: Americas' Conference on Information Systems 2019; August 15; Cancun, Mexico.
13. Google Cloud Speech-to-Text. URL: https://cloud.google.com/speech-to-text [accessed 2019-11-19]
14. Miner AS, Haque A, Fries JA, Fleming SL, Wilfley DE, Terence Wilson G, et al. Assessing the accuracy of automatic speech recognition for psychotherapy. NPJ Digit Med 2020;3:82 [FREE Full text] [doi: 10.1038/s41746-020-0285-8] [Medline: 32550644]
15. Kaup AR, Nettiksimmons J, LeBlanc ES, Yaffe K. Memory complaints and risk of cognitive impairment after nearly 2 decades among older women. Neurology 2015 Nov 24;85(21):1852-1858 [FREE Full text] [doi: 10.1212/WNL.0000000000002153] [Medline: 26511452]

XSL•FO
RenderX

16. Rigas E, Papoutsoglou M, Tsakpounidou K, Serdari A, Mouza E, Proios H. Analysis of spontaneous speech using natural language processing techniques to identify stroke symptoms. Encephalos 2020 Jan;57(1):1-12 [FREE Full text]
17. Moro-Velazquez L, Gomez-Garcia J, Arias-Londoño J, Dehak N, Godino-Llorente J. Advances in Parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects. Biomed Signal Process Control 2021 Apr;66:102418. [doi: 10.1016/j.bspc.2021.102418]
18. Falcone M, Yadav N, Poellabauer C, Flynn P. Using isolated vowel sounds for classification of mild traumatic brain injury. : IEEE; 2013 Presented at: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; May 26-31; Vancouver, BC p. 7577-7581 URL: https://ieeexplore.ieee.org/document/6639136 [doi: 10.1109/ICASSP.2013.6639136]
19. McGinnis EW, Anderau SP, Hruschak J, Gurchiek RD, Lopez-Duran NL, Fitzgerald K, et al. Giving voice to vulnerable children: machine learning analysis of speech detects anxiety and depression in early childhood. IEEE J Biomed Health Inform 2019 Nov;23(6):2294-2301 [FREE Full text] [doi: 10.1109/JBHI.2019.2913590] [Medline: 31034426]
20. Karam Z, Provost E, Singh S, Montgomery J, Archer C, Harrington G, et al. Ecologically valid long-term mood monitoring of individuals with bipolar disorder using speech. 2014 Presented at: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2014; May 4; Florence, Italy p. 4858-4862 URL: https://ieeexplore.ieee.org/document/6854525 [doi: 10.1109/icassp.2014.6854525]
21. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. A review of depression and suicide risk assessment using speech analysis. Speech Communication 2015 Jul;71:10-49. [doi: 10.1016/j.specom.2015.03.004]
22. Rude S, Gortner E, Pennebaker J. Language use of depressed and depression-vulnerable college students. Cogn Emot 2004 Dec;18(8):1121-1133. [doi: 10.1080/02699930441000030]
23. Goodglass H, Kaplan E. BDAE: the Boston Diagnostic Aphasia Examination. proedinc.com. Philadelphia, PA: Lippincott Williams & Wilkins; 1972. URL: https://www.proedinc.com/Products/11850/bdae3-boston-diagnostic-aphasia-examinationthird-edition.aspx?bCategory=TBI!APXDY [accessed 2022-09-06]
24. Karlekar S, Niu T, Bansal M. Detecting linguistic characteristics of Alzheimer’s dementia by interpreting neural models. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2018 Presented at: 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; June 1; New Orleans, LA p. 701-707. [doi: 10.18653/v1/N18-2110]
25. Kong W, Jang H, Carenini G, Field T. A neural model for predicting dementia from language. In: Proceedings of Machine Learning Research. 2019 Presented at: 4th Machine Learning for Healthcare Conference; August 8; Ann Arbor, MI p. 270-286.
26. Cummings L. Describing the Cookie Theft picture. Pragmat Soc 2019 Jul 5;10(2):153-176. [doi: 10.1075/ps.17011.cum]
27. Eyigoz E, Mathur S, Santamaria M, Cecchi G, Naylor M. Linguistic markers predict onset of Alzheimer's disease. EClinicalMedicine 2020 Nov;28:100583 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100583] [Medline: 33294808]
28. Trauzettel-Klosinski S, Dietz K, IReST Study Group. Standardized assessment of reading performance: the New International Reading Speed Texts IReST. Invest Ophthalmol Vis Sci 2012 Aug 13;53(9):5452-5461. [doi: 10.1167/iovs.11-8284] [Medline: 22661485]
29. Biondi J, Fernandez G, Castro S, Agamennoni O. Eye movement behavior identification for Alzheimer's disease diagnosis. J. Integr. Neurosci 2018 Nov 15;17(4):1702-00837. [doi: 10.31083/j.jin.2018.04.0416]
30. Goldman-Eisler F. Speech production and the predictability of words in context. Q J Exp Psychol (Hove) 2018 Jan 01;10(2):96-106. [doi: 10.1080/17470215808416261]
31. Park S. Measuring fluency: Temporal variables and pausing patterns in L2 English speech. Dissertation, Perdue University. 2016 Apr. URL: https://docs.lib.purdue.edu/open_access_dissertations/692/ [accessed 2021-01-01]
32. Errattahi R, El Hannani A, Ouahmane H. Automatic speech recognition errors detection and correction: a review. Procedia Comput Sci 2018;128:32-37. [doi: 10.1016/j.procs.2018.03.005]
33. Vaessen N. JiWER: Similarity measures for automatic speech recognition evaluation. Python Package Index. 2018 Jun 19. URL: https://pypi.org/project/jiwer/ [accessed 2021-01-01]
34. Barral O, Jang H, Newton-Mason S, Shajan S, Soroski T, Carenini G, et al. Non-invasive classification of Alzheimer’s disease using eye tracking and language. In: Proceedings of the 5th Machine Learning for Healthcare Conference. 2020 Presented at: 5th Machine Learning for Healthcare Conference; August 7; Virtual p. 813-841.
35. Fraser KC, Lundholm Fors K, Eckerström M, Öhman F, Kokkinakis D. Predicting MCI status from multimodal language data using cascaded classifiers. Front Aging Neurosci 2019;11:205. [doi: 10.3389/fnagi.2019.00205] [Medline: 31427959]
36. Manning C, Surdeanu M, Bauer J, Finkel J, Bethard S, McClosky D. The Stanford CoreNLP natural language processing toolkit. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations. 2014 Presented at: 52nd Annual Meeting of the Association for Computational Linguistics; June; Baltimore, MD p. 55-60 URL: https://stanfordnlp.github.io/CoreNLP/ [doi: 10.3115/v1/P14-5010]
37. MRC Psycholinguistic Database: machine usable dictionary. University of Western Australia. 1987 Apr 01. URL: https://websites.psychology.uwa.edu.au/school/MRCDatabase/mrc2.html [accessed 2020-01-01]
38. Robert J. Pydub. Python Package Index. 2011 May 03. URL: https://github.com/jiaaro/pydub [accessed 2019-12-31]

XSL•FO
RenderX

39. Day D. Syllables: A fast syllable estimator for Python. Python Package Index. 2018 Nov 25. URL: https://github.com/prosegrinder/python-syllables [accessed 2020-01-01]
40. Masrani V. Detecting dementia from written and spoken language. Doctoral Dissertations, The University of British Columbia. Vancouver, BC; 2018 Jan 08. URL: https://doi.org/10.14288/1.0362923 [accessed 2020-12-31]
41. Devlin J, Chang M, Lee K, Toutanova K. BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2019 Presented at: 2019 Conference of the North American Chapter of the Association for Computational Linguistics; June; Minneapolis, MN p. 4171-4186 URL: https://aclanthology.org/N19-1423/ [doi: 10.18653/v1/N19-1423]
42. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res 2011;12:2825-2830 [FREE Full text]
43. Šimundić AM. Measures of diagnostic accuracy: basic definitions. EJIFCC 2009 Jan;19(4):203-211 [FREE Full text] [Medline: 27683318]
44. Hall MA. Correlation-based feature selection for machine learning. Doctoral Dissertation, The University of Waikato. Waikato, New Zealand; 1999. URL: https://researchcommons.waikato.ac.nz/handle/10289/15043 [accessed 2021-01-01]
45. Ahmed S, Haigh AF, de Jager CA, Garrard P. Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease. Brain 2013 Dec;136(Pt 12):3727-3737 [FREE Full text] [doi: 10.1093/brain/awt269] [Medline: 24142144]
46. Kohn S. Phonological production deficits in aphasia. In: Phonological Processes and Brain Mechanisms. New York, NY: Springer; 1988:93-117.
47. Mirzaei S, El Yacoubi M, Garcia-Salicetti S, Boudy J, Kahindo C, Cristancho-Lacroix V, et al. Two-stage feature selection of voice parameters for early Alzheimer's disease prediction. IRBM 2018 Dec;39(6):430-435. [doi: 10.1016/j.irbm.2018.10.016]
48. Martínez-Sánchez F, Meilán JJG, Vera-Ferrandiz JA, Carro J, Pujante-Valverde IM, Ivanova O, et al. Speech rhythm alterations in Spanish-speaking individuals with Alzheimer's disease. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2017 Jul;24(4):418-434. [doi: 10.1080/13825585.2016.1220487] [Medline: 27684109]
49. Chiu C, Sainath T, Wu Y, Prabhavalkar R, Nguyen P, Chen Z, et al. State-of-the-art speech recognition with sequence-to-sequence models. 2018 Presented at: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing; April 15; Calgary, AB p. 4774-4778. [doi: 10.1109/ICASSP.2018.8462105]
50. Kim J, Liu C, Calvo R, McCabe K, Taylor S, Schuller B, et al. A Comparison of Online Automatic Speech Recognition Systems and the Nonverbal Responses to Unintelligible Speech. In: 12th International Workshop on Spoken Dialog System Technology Proceedings. 2021 Nov 15 Presented at: 12th International Workshop on Spoken Dialog System Technology; 15 Novmeber 2021; Singapore. [doi: 10.1007/978-3-030-14596-5_8]
51. Këpuska V. Comparing speech recognition systems (Microsoft API, Google API And CMU Sphinx). Int J Eng Res Appl 2017 Mar;07(03):20-24. [doi: 10.9790/9622-0703022024]
52. Murray G, Carenini G. Summarizing spoken and written conversations. In: Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing.: Association for Computational Linguistics; 2008 Presented at: 2008 Conference on Empirical Methods in Natural Language Processing; Honolulu, HI p. 773-782. [doi: 10.3115/1613715.1613813]
53. Pistono A, Jucla M, Barbeau EJ, Saint-Aubert L, Lemesle B, Calvet B, et al. Pauses during autobiographical discourse reflect episodic memory processes in early Alzheimer's disease. J Alzheimers Dis 2016;50(3):687-698 [FREE Full text] [doi: 10.3233/JAD-150408] [Medline: 26757034]
54. Pistono A, Pariente J, Bézy C, Lemesle B, Le Men J, Jucla M. What happens when nothing happens? An investigation of pauses as a compensatory mechanism in early Alzheimer's disease. Neuropsychologia 2019 Feb 18;124:133-143 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2018.12.018] [Medline: 30593773]
55. Sluis RA, Angus D, Wiles J, Back A, Gibson T, Liddle J, et al. An automated approach to examining pausing in the speech of people with dementia. Am J Alzheimers Dis Other Demen 2020 Jul 10;35:1533317520939773 [FREE Full text] [doi: 10.1177/1533317520939773] [Medline: 32648470]
56. Bóna J. Temporal characteristics of speech: the effect of age and speech style. J Acoust Soc Am 2014 Aug;136(2):EL116-EL121. [doi: 10.1121/1.4885482] [Medline: 25096134]

Abbreviations

AD: Alzheimer disease
AUROC: area under the receiver operating characteristic curve
BERT: Bidirectional Encoder Representations from Transformers
GNB: Gaussian naive Bayes
IReST: International Reading Speed Texts
LR: logistic regression
MCI: mild cognitive impairment

XSL•FO
RenderX

MER: match error rate
NLP: natural language processing
RF: random forest
SMC: subjective memory complaints
STT: speech-to-text
WER: word error rate

Edited by T Leung, J Wang; submitted 08.09.21; peer-reviewed by S Kim, M Burns, X Zhou, Y Liu; comments to author 14.03.22; revised version received 11.07.22; accepted 23.07.22; published 21.09.22.

Please cite as:
Soroski T, da Cunha Vasco T, Newton-Mason S, Granby S, Lewis C, Harisinghani A, Rizzo M, Conati C, Murray G, Carenini G, Field TS, Jang H
Evaluating Web-Based Automatic Transcription for Alzheimer Speech Data: Transcript Comparison and Machine Learning Analysis
JMIR Aging 2022;5(3):e33460
URL: https://aging.jmir.org/2022/3/e33460
doi:10.2196/33460
PMID:36129754

©Thomas Soroski, Thiago da Cunha Vasco, Sally Newton-Mason, Saffrin Granby, Caitlin Lewis, Anuj Harisinghani, Matteo Rizzo, Cristina Conati, Gabriel Murray, Giuseppe Carenini, Thalia S Field, Hyeju Jang. Originally published in JMIR Aging (https://aging.jmir.org), 21.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

mHealth Interventions to Support Caregivers of Older Adults: Equity-Focused Systematic Review

Anna Garnett[1], RN, PhD; Melissa Northwood[2], RN, PhD; Justine Ting[1], RN, MSc; Ruheena Sangrar[3], OT Reg (Ont), PhD

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

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

[3]Department of Occupational Science and Occupational Therapy, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Anna Garnett, RN, PhD
Arthur Labatt Family School of Nursing
Western University
FIMS & Nursing Building, Rm 2306
London, ON, N6A 5B9
Canada
Phone: 1 519 661 2111
Email: agarnet6@uwo.ca

Abstract

Background: Informal caregivers, hereafter referred to as caregivers, provide support to older adults so that they can age safely at home. The decision to become a caregiver can be influenced by individual factors, such as personal choice, or societal factors such as social determinants of health, including household income, employment status, and culture-specific gender roles. Over time, caregivers' health can be negatively affected by their caregiving roles. Although programs exist to support caregivers, the availability and appropriateness of services do not match caregivers' expressed needs. Research suggests that supportive interventions offered through mobile health (mHealth) technologies have the potential to increase caregivers' access to supportive services. However, a knowledge gap remains regarding the extent to which social determinants of health are considered in the design, implementation, and evaluation of mHealth interventions intended to support the caregivers of older adults.

Objective: This study aimed to conduct a systematic review to determine how health equity is considered in the design, implementation, and evaluation of mHealth interventions for caregivers of older adults using Cochrane Equity's PROGRESS-Plus (place of residence, race, ethnicity, culture, language, occupation, gender, religion, education, social capital, socioeconomic status–plus age, disability, and sexual orientation) framework and synthesize evidence of the impacts of the identified caregiver-focused mHealth interventions.

Methods: A systematic review was conducted using 5 databases. Articles published between January 2010 and June 2021 were included if they evaluated or explored the impact of mHealth interventions on the health and well-being of informal caregivers of older adults. mHealth interventions were defined as supportive services, for example, education, that caregivers of older adults accessed via mobile or wireless devices.

Results: In total, 28 articles met the inclusion criteria and were included in the review. The interventions evaluated sought to connect caregivers with services, facilitate caregiving, and promote caregivers' health and well-being. The PROGRESS-Plus framework factors were mainly considered in the results, discussion, and limitations sections of the included studies. Some PROGRESS-Plus factors such as sexual orientation, religion, and occupation, received little to no consideration across any phase of the intervention design, implementation, or evaluation. Overall, the findings of this review suggest that mHealth interventions were positively received by study participants. Such interventions have the potential to reduce caregiver burden and positively affect caregivers' physical and mental health while supporting them as caregivers. The study findings highlight the importance of making support available to help facilitate caregivers' use of mHealth interventions, as well as in the use of appropriate language and text.

Conclusions: The successful uptake and spread of mHealth interventions to support caregivers of older adults will depend on creating opportunities for the inclusive involvement of a broad range of stakeholders at all stages of design, implementation, and evaluation.

XSL•FO
RenderX

(JMIR Aging 2022;5(3):e33085) doi:10.2196/33085

KEYWORDS

caregivers; older adults; mobile health; social determinants of health; intervention; mobile phone

Introduction

Background

Globally, it is estimated that 101 million older adults require care from a friend or family caregiver (informal caregiver; hereafter referred to as caregiver), with women providing most of the support [1]. The support provided by these informal caregivers is often crucial for enabling older adults to safely remain in their home environment [2-4]. Caregiving support, such as assistance with activities of daily living, attending appointments, and health management, is associated with positive outcomes for both caregivers [5,6] and care recipients [7]. Although caregivers often willingly engage in caring, their role can negatively affect their psychological well-being, particularly when care is provided over a prolonged period [8-10].

The Social Determinants of Health and Inequities Among Caregivers

The social determinants of health can influence entry into the caregiving role and the subsequent experience of being a caregiver. For example, factors such as being a woman, lower educational attainment, and living in a rural setting can bias caregiving toward individuals who may perceive that they have little agency in their choice to become a caregiver [11]. Furthermore, a greater intensity of caregiving has been identified among caregivers who are female, people of color, and of lower socioeconomic status [12]. These inequities highlight the need for interventions with both scope and accessibility to support caregivers with varied demographic characteristics.

Although some programs and community initiatives are available to support caregivers, the literature suggests that caregivers struggle to access these supportive services [13-15]. Challenges in system navigation, accessing support, geographic location, and scheduling factors can impede the successful use of services [16,17]. Recent research indicates that supportive services provided or augmented through mobile health (mHealth) technologies have the potential to make services more accessible to caregivers [18-20].

mHealth Interventions as a Potential Solution for Caregiver Support

The term *mobile health* (mHealth) was first coined in 2003 in response to the rapid development and expansion of mobile communication technologies being used within the health care industry [21]. The World Health Organization defines mHealth as a "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices" [22]. The use of health information technology (computer, internet, and email) to access health records or locate health information on the web has become commonplace among caregivers as a means of informing their caregiving role [23]. Research suggests that mobile apps have the potential to have a greater positive impact on caregivers by providing support, communication, and facilitation of care, thereby reducing the burden and positively affecting caregiver health outcomes [24]. However, to the best of our knowledge, a systematic review of the impact of mHealth support on caregivers of older adults does not currently exist. Furthermore, to date, reviews on standard caregiver interventions suggest that limited work has been conducted to determine the suitability of these interventions for caregivers from backgrounds representing diverse social determinants of health characteristics [25]. Individual characteristics, such as sociodemographic characteristics and the ability to engage with technology, should be considered in the design of mHealth interventions [26].

Therefore, the objectives of this systematic review were to (1) determine how health equity is considered in the design, implementation, and evaluation of mHealth interventions aimed at caregivers of older adults using the Cochrane Equity PROGRESS-Plus (place of residence, race, occupation, gender, religion, education, social capital, socioeconomic status–plus age, disability, and sexual orientation) framework [27] and (2) synthesize the evidence on the impacts of caregiver-focused mHealth interventions, subsequently discussed through a health equity lens.

Methods

A systematic review was conducted in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement guidelines [28]. The protocol for this systematic review was registered on PROSPERO (International Prospective Register of Systematic Reviews; CRD42021239584) and is available for electronic access [29].

Research Questions

The research questions guiding this systematic review were as follows:

- To what extent is health and social equity considered in the design, implementation, and evaluation of mHealth interventions for caregivers of older adults?
- What are the impacts of the examined mHealth interventions on caregivers of older adults based on the following outcomes: caregiver mental and physical health, caregivers' ability to provide care, usability or feasibility of the mHealth intervention for caregivers, and caregivers' experiences and perspectives of engaging in an mHealth intervention intended to support them?

Eligibility Criteria

Eligible articles were available in full text in the English language and were published from 2010 onward to reflect the recent surge in mHealth interventions, concurrent with the rapid increase in mobile device ownership within the past decade [30,31]. This review included both quantitative (experimental,

XSL•FO
RenderX

quasi-experimental, and observational studies with or without control or comparison groups) and qualitative study designs, which evaluated or explored the impacts of mHealth interventions aimed at improving the health of, or providing support to, informal caregivers of older adults. Mixed methods studies were also included. mHealth interventions were defined as those that the caregivers of older adults accessed via mobile or wireless devices (including mobile phones, tablets, handheld computers, and PDAs). Interventions not accessed by mobile or wireless devices (eg, interventions applied or accessed by landline telephone as opposed to mobile phone) were excluded, as were mHealth interventions that targeted the recipient of care only or only assessed outcomes focused on the recipient of care. Studies that exclusively included formal caregivers of older adults (eg, nurses and personal support workers) or caregivers of individuals who were not identified as older adults (eg, children, adolescents, young and middle-aged adults, or adults aged <65 years) were also excluded.

Eligible studies were also required to report at least one caregiver-specific outcome or finding, including those relating to (1) caregiver mental and physical health, (2) caregivers' ability to provide care, (3) usability or feasibility of the mHealth intervention by caregivers, and (4) caregivers' experiences and perspectives of engaging in mHealth interventions intended to support them. Research protocols, dissertations, reviews, commentaries, and abstracts were also excluded.

Search Strategy and Study Selection

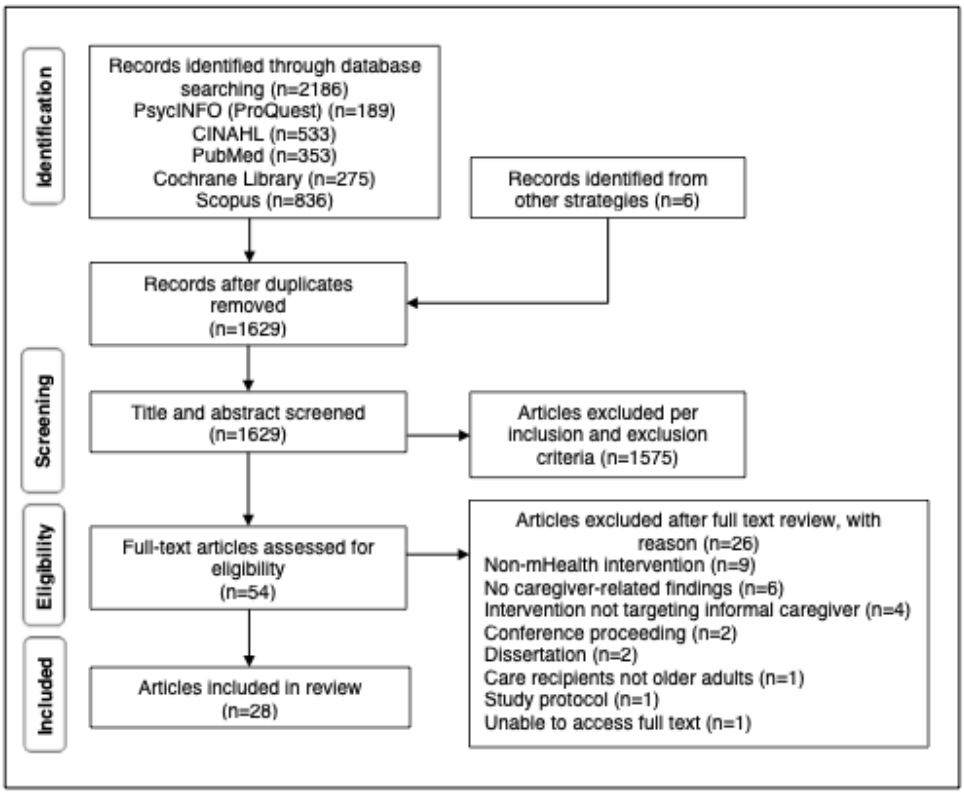
A systematic search was conducted on five databases: PubMed, PsycINFO (ProQuest), CINAHL, Scopus, and Cochrane Library. An academic librarian was consulted during database search strategy development. Database searches combined a comprehensive suite of similar and related terms for the key domains of *caregivers*, *older adults*, and *mHealth interventions*. Detailed search strategies for each database are provided in Multimedia Appendix 1. The search results were limited by the year of publication from 2010 to February 2021, when the search was initially conducted. The search strategy was repeated in June 2021 to capture newly published articles. Ancestry searches were also conducted using the reference lists of eligible studies, as well as related reviews [19,32-34], to search for additional potential articles for inclusion.

Eligible studies identified from the database and ancestry searches were independently assessed by a group of 4 reviewers (AG, MN, RS, and JT). Each document was reviewed by 2 reviewers (AG, MN, RS, or JT) based on the title and abstract. The full texts of relevant studies were then obtained, and 2 reviewers (AG, MN, RS, or JT) independently examined the full texts of the selected studies to determine the final included articles in accordance with the eligibility criteria outlined previously. Covidence systematic review software (Veritas Health Innovation) was used to organize the search results and facilitate communication between the reviewers. Disagreements were resolved by consensus. In cases where consensus could not be reached, a third reviewer resolved the disagreement.

The search strategy yielded an initial 1629 articles for screening of titles and abstracts. On the basis of the initial screening, the full texts of the 3.31% (54/1629) of articles were assessed. Of the 54 articles, 26 (48%) were subsequently excluded after a full-text review. The literature search and study selection processes are shown in Figure 1. A total of 28 articles met the inclusion criteria and were included in the review.

XSL•FO
RenderX

Figure 1. Literature search and study selection process. mHealth: mobile health.



Data Extraction

The data were extracted using reviewer-designed data extraction forms in Covidence. A total of 2 reviewers independently performed the data extraction. Disagreements were resolved by consensus. In cases where consensus could not be achieved, a third reviewer was consulted.

Data extracted from full-text articles included (1) country of investigation; (2) study design and methods; (3) participant recruitment, demographics, and baseline characteristics; (4) description of the mHealth intervention; and (5) caregiver-specific outcomes or findings. In addition, the review team identified which (if any) social determinants of health and factors contributing to health inequities were addressed by study investigators, as described by the PROGRESS-Plus framework [27,35].

PROGRESS-Plus is a framework developed with evidence from working groups from the Campbell and Cochrane Collaborations, which can be applied to determine whether an equity lens has been used throughout the stages of study design, implementation, and reporting of research [27]. The framework includes the following equity factors: place of residence, race, ethnicity, language, culture, occupation (eg, full-time employment or retirement), gender or sex, religion, education, socioeconomic status, and social capital, as well as age, disability, sexual orientation, features of relationships, and time-dependent relationships (*Plus* factors) [27]. The manner in which investigators addressed these factors within the intervention itself and the study of the intervention was considered in their report of these factors within the following sections: (1) mHealth intervention design, (2) study participant recruitment, (3) study results or findings, and (4) discussion or limitations of the investigation.

Risk of Bias Assessment

Risk of bias (quality) assessments were performed for each study using standardized critical appraisal tools from the Joanna Briggs Institute Manual for Evidence Synthesis [36]. The Joanna Briggs Institute provides distinct critical appraisal checklists for experimental, quasi-experimental, observational, and qualitative study design. One of the reviewers performed the risk of bias assessments for each study, which was then checked by a second reviewer. Disagreements were resolved by discussion and consensus. No studies were excluded from the review based on quality assessments to achieve a comprehensive understanding of the quality of the available literature exploring the impacts of mHealth interventions for caregivers of older adults. The findings of the quality assessments and the limitations of the included articles are summarized in the results, and the summary scores of the quality assessments are presented in the *Results* section.

Data Synthesis

A narrative synthesis of findings was pursued because of the range of included mHealth interventions, caregiver characteristics, and caregiver-related outcome measures, as well as the inclusion of both quantitative and qualitative study designs. The narrative synthesis was organized under the following categories: (1) study characteristics; (2) mHealth intervention characteristics; (3) consideration of social determinants and factors contributing to health inequities in

XSL•FO
RenderX

mHealth intervention design, participant recruitment, study results or findings, and discussion or limitations; (4) quantitative caregiver-related outcomes; and (5) qualitative caregiver-related findings.

Results

A total of 28 articles were included in this review. A summary of the included articles is presented in Multimedia Appendix 2 [37-62].

Characteristics of Included Studies

Among the 28 included studies, 14 (50%) were quantitative [37-48,63,64], 7 (25%) were qualitative [49-55], and 7 (25%) used mixed methods [56-62]. Studies were most frequently conducted in the United States [38,41,45-48,50,53,58,60], the Netherlands [37,55,57,59], the United Kingdom [54,62], and Australia [52,56]. Most studies targeted nonspecific informal caregivers of older adults; however, 25% (7/28) targeted family or spousal caregivers specifically [38,44,51,52,54,60,64]. Approximately 7% (2/28) of studies targeted caregivers who reported being isolated [56] or experiencing caregiving strain [38]. Caregivers most commonly provided care to older adults with dementia or other forms of cognitive impairment [37-39,41-44,46,47,50-56,58-60,62,64]. Other studies recruited caregivers who provided care to older adults with urinary incontinence [63], older veterans who were medically complex [45], and older adults with functional loss or struggling to remain independent at home [49,57,61].

Risk of Bias Within Included Studies

The full risk of bias assessments for the 28 included studies are presented in Multimedia Appendix 3 [37-62]. The potential for bias within the 11% (3/28) included randomized controlled trials [37,45,46] most commonly stemmed from a lack of blinding of participants and outcome assessors. Potential sources of bias within other quantitative studies include a lack of control groups [60,62,63] and limited consideration of potential confounders [41,56]. Most of the included quantitative studies recruited small convenience samples of caregivers or caregiver–care recipient dyads; for example, recruiting from single clinics [38,39] or from attendees of an Alzheimer's Association chapter event [50]. Included qualitative studies were most often limited by a lack of clear alignment between philosophical underpinnings, methodology, and research questions or objectives [49-53,55-58,60,62]. Although most studies provided sufficient information to demonstrate a logical flow from the analysis and interpretation of data to the overall conclusions, few studies addressed the potential influence of the researcher on the research (eg, positionality, trustworthiness, and rigor) [54,62]. In addition, only 7% (2/28) of qualitative studies provided information on the location of the researcher's theoretical approach [53,54]. Although other studies may also have used a theoretical lens or framework to guide their intervention and analysis, they did not report this information.

mHealth Intervention Characteristics

The included studies' interventions were web-based or non–web-based applications, interventions, or videoconferencing software, which were delivered via mobile phones, tablets, and handheld computers. The intervention details, including intervention description, hardware, stakeholder input, and comparison groups, are outlined in Table 1.

The aims of these interventions fell under three interrelated categories: making connections, facilitating caregiving, and promoting caregiver health and well-being (Figure 2). The included mHealth interventions facilitated various linkages and connections between caregivers and supportive services, such as (1) connecting the care recipient's circle of care, including caregivers and health professionals [44,45,48,51,53,55,57,58,61]; (2) connecting the caregiver to existing social support or facilitating new connections to peer support [40,43,46,56,59]; and (3) connecting the caregiver to services and resources for both themselves and the recipient of care [37,43,47,50,51,53,58].

XSL•FO
RenderX

Table 1. Details of mobile health interventions of included studies.

Study	Intervention description	Hardware provided	Stakeholder input described	Comparator intervention (as applicable)	Study quality appraisal scores[a]
Quantitative studies—randomized controlled trials					
Beentjes et al [37]	FindMyApps, a web-based selection tool and learning training program to help caregivers find user-friendly apps	Yes; tablet	No	Caregiver controls received a tablet but no FindMyApps training or access; received a list of links to websites with apps for people with dementia or mild cognitive impairment	8/13
Hastings et al [45]	Video-enhanced care management: a 14-week care management intervention that included 3 monthly video calls with nurses via a secure internet-based web-based meeting room	Yes; tablet	No	One group received the intervention (video); the comparator group received telephone-based care management	5/13
Kales et al [46]	WeCareAdvisor, a web-based tool for family caregivers, which guided them through a clinical reasoning process to identify, monitor, and manage behaviors while addressing their motivation, self-efficacy, and problem-solving skills	Yes; tablet	No	Waitlist for the tool; this group received the tool 1 month later	8/13
Quantitative studies—quasi-experimental					
Davis et al [63]	TelePrompt, a tablet-based, prompted voiding and educational intervention to support caregivers of older adults with urinary incontinence	Yes; tablet	No	No comparison group; the study was described by authors as a quasi-experimental, single-group pre-post design	6/9
Lai et al [44]	Telehealth delivered via videoconferencing platforms (apps) aimed at minimizing the possible negative impact of social distancing measures made necessary by the COVID-19 pandemic	No	No	Received a weekly care service via telephone covering information relevant to caregiving; did not receive the intervention of weekly health services delivered through video communication apps	7/9
Park et al [64]	Comprehensive Mobile Application Program, a tool providing real-time support to families caring for patients with dementia by helping family caregivers manage behavior and psychological symptoms	No	No	Comparator intervention was a handbook that contained the same information as the mobile app	5/9
Watcharasarnsap et al [42]	A mobile app system based on the reminiscence therapy framework; the app was developed to promote the relationship between caregivers and people with dementia and better the mental well-being of both parties	No	No	Control group did not use the intervention (no intervention)	9/9
Quantitative s tudies — other (ie, noncomparative)					
Callan et al [38]	A self-administered cognitive training intervention using an adaptive, paced serial attention task, targeting the dorsolateral prefrontal cortex, which is implicated in regulating emotions, anxiety, and stress	Yes; handheld computer	No	N/A[b]	6/10

XSL•FO
RenderX

Study	Intervention description	Hardware provided	Stakeholder input described	Comparator intervention (as applicable)	Study quality appraisal scores[a]
Davis et al [43]	An e-mobile multimedia app for community-based dementia caregiver support, designed to offer reassurance, information, and services to caregivers and facilitate the implementation of other interventions by nurses and therapists	Yes; mobile phone	No	N/A	1/10
Ptomey et al [47]	A remotely delivered exercise intervention to increase moderate physical activity in caregivers	Yes; tablet	No	N/A	4/10
Quinn et al [48]	A mobile app designed to improve engagement of the patient-informal caregiver team; the mobile web-based app allowed older adult users to record social and health information and share this information with their caregivers	No	No	N/A	4/10
Lai et al [39]	A simple smartphone app for people with mild cognitive impairment and their family caregivers living in the community; the app supported communication with friends and family, navigation, and serving as a memory prompt and emergency alert system	Yes; mobile phone	No	N/A	6/10
Salin and Laaksonen [40]	A multicomponent intervention, including live broadcasts related to caregiver self-care exercises, informational videos, and videoconferencing web-based meetings to connect informal caregivers	Yes; tablet	Yes	N/A	2/10
Sourbeer et al [41]	A preliminary tablet app developed for the Behavioral and Environmental Sensing and Intervention for Dementia Caregiver Empowerment; the goal of this app is to support the early detection of signs of agitation, allowing caregivers to intervene early	Yes; tablet	No	N/A	2/11
Mixed methods studies					
Banbury et al [56]	A telehealth peer-support program for isolated caregivers of people with dementia via group videoconferencing	Yes; not specified	No	N/A	3/8 and 3/10
Breebaart and van Groenou [57]	A groupware app for digital network communication to promote collaboration among informal and formal caregivers in a mixed care network of home-dwelling older adults	Yes; not specified	No	N/A	1/10 and 3/10
Brown et al [58]	CareHeroes, an app providing caregivers with a platform for bidirectional sharing of observations and knowledge with providers about care recipients and, in so doing, provide them with information and support for caregiving activities	No	Yes	N/A	4/10 and 3/10
Dam et al [59]	Inlife, a web-based social support platform for caregivers of individuals with dementia aiming to enhance positive interaction, involvement, and social support	No	No	Control group did not receive the intervention (waiting list)	4/10 and 7/10
Sikder et al [60]	A mobile app intervention delivering mentalizing imagery therapy (a guided imagery and mindfulness intervention) for family caregivers	No	No	N/A	5/9 and 3/10

XSL•FO
RenderX

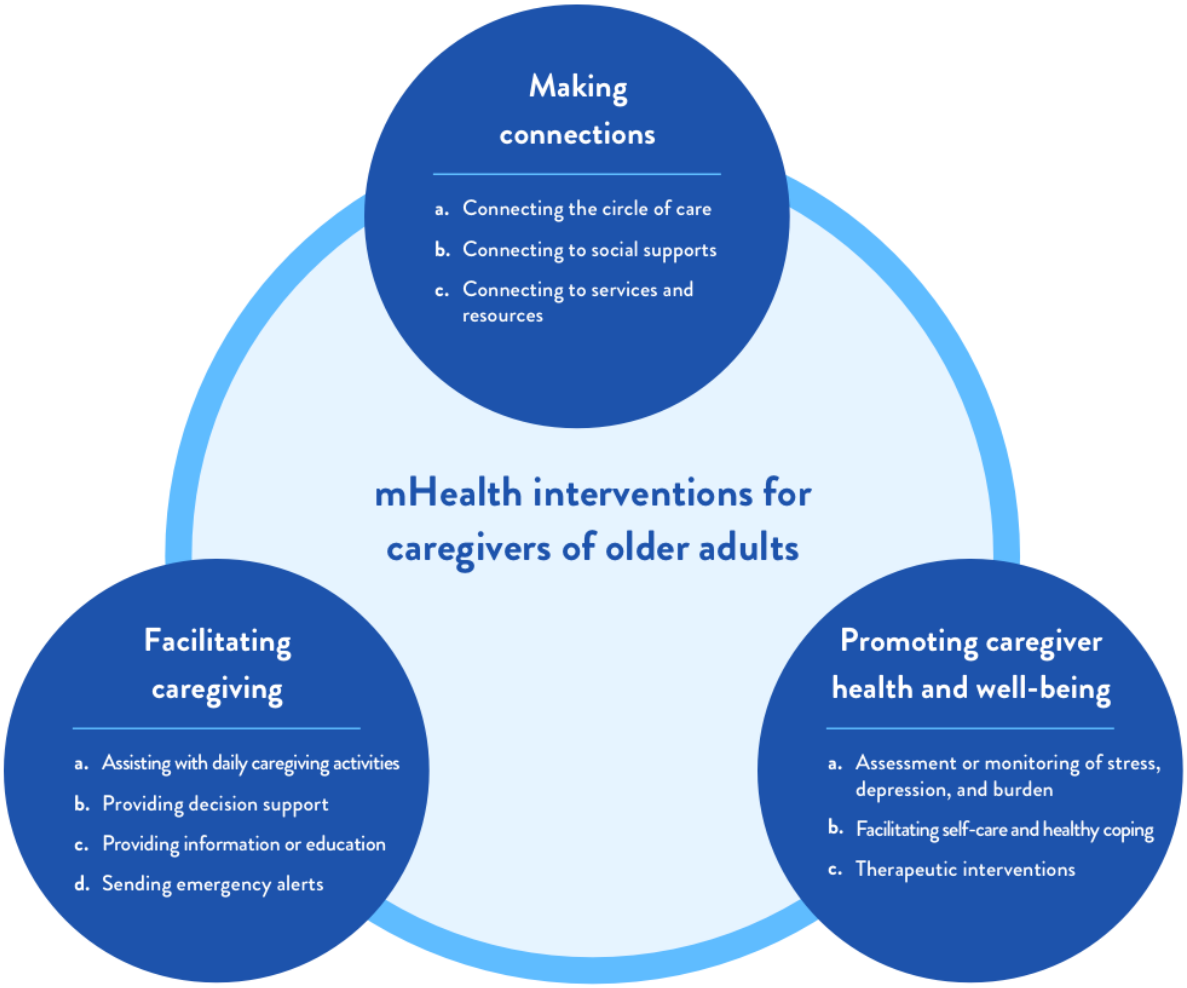
Study	Intervention description	Hardware provided	Stakeholder input described	Comparator intervention (as applicable)	Study quality appraisal scores[a]
Stutzel et al [61]	A mobile phone app, The Mobile System for Elderly Monitoring, which aimed to support caregivers in monitoring care recipients with functional loss and to improve support for caregivers' communication with the health team	Yes; mobile phone	Yes	N/A	5/10 and 7/10
Tyack et al [62]	An art-based app intervention delivered via a touch screen tablet displaying art images aiming to stimulate and benefit the well-being of caregivers and care recipients with dementia	Yes; tablet	Yes	N/A	6/9 and 8/10
Qualitative studies					
Garvelink et al [49]	A decision support website to inform caregivers about ways of staying independent at home for as long as possible, called Supporting Seniors and Caregivers to Stay Mobile at Home	No	No	N/A	3/10
Hughes et al [50]	A tablet app with multiple components, including games and a stress questionnaire for caregivers	No	Yes	N/A	5/10
Killin et al [51]	The Digital Support Platform, an internet-based, postdiagnostic support tool for families of individuals who had recently received a diagnosis of dementia	Yes; tablet	No	N/A	6/10
Rathnayake et al [52]	Mobile health apps used for health information seeking	No	No	N/A	7/10
Ruggiano et al [53]	CareIT, a multifunctional smartphone and web-based app designed to meet the education and support needs of caregivers; the app allowed caregivers to self-assess for depression and burden and linked caregivers to resources	Yes; mobile phone	Yes	N/A	5/10
Ryan et al [54]	InspireD—Individual Specific Reminiscence in Dementia, a personalized reminiscence program for family carers and people living with dementia	Yes; tablet	Yes	N/A	10/10
Span et al [55]	The DecideGuide, an interactive web tool that helps informal caregivers, people with dementia, and case managers make shared decisions	Yes; tablet	Yes	N/A	5/10

[a]Complete quality appraisal tools and scores are presented in Multimedia Appendix 3.

[b]N/A: not applicable.

XSL•FO
RenderX

Figure 2. Mobile health (mHealth) interventions for caregivers of older adults.



mHealth interventions included in the review also facilitated caregiving by (1) assisting with daily caregiving activities (eg, digital calendars to organize appointments, providing reminders for medication administration, helping caregivers manage care recipient behaviors, and tracking information related to the care recipient) [39,41,46,48,51,53,57,59,61,63,64], (2) providing support for decisions related to care [46,49,55,58], (3) providing information or education (eg, regarding the care recipient's condition) [40,43-46,48,49,51-53,56,58,63,64], and (4) sending emergency alerts to the caregiver or to the care team if needed [39,61].

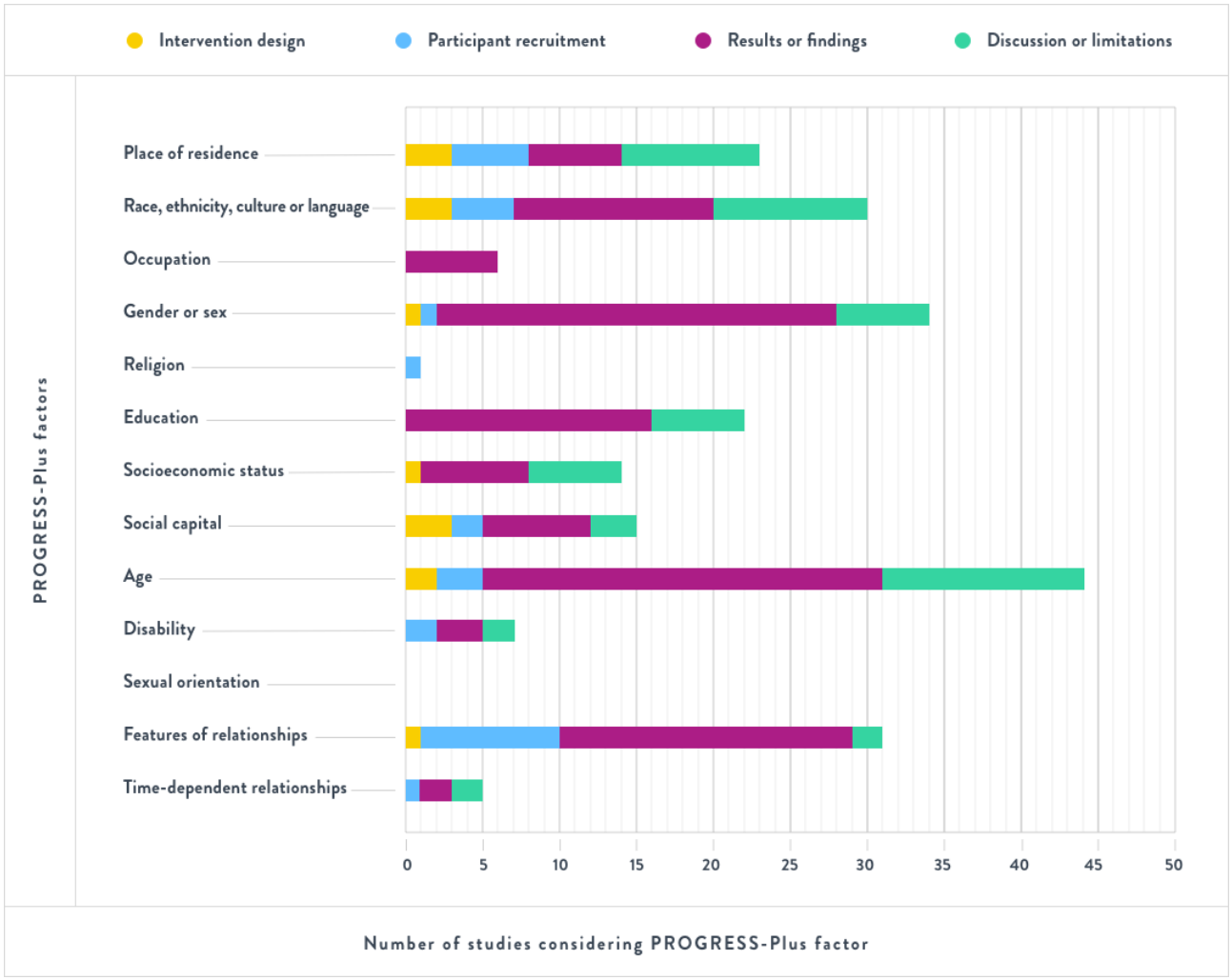
Finally, the mHealth interventions represented in the review promoted caregiver health and well-being by (1) monitoring or assessing caregiver stress, depression, and burden to facilitate early detection and intervention before reaching crisis levels [41,50,53,58,61]; (2) promoting self-care and healthy coping behaviors (eg, encouraging physical activity or suggesting evidence-based coping strategies for care recipient behaviors) [40,43,47,63,64]; and (3) providing therapeutic interventions (eg, art-based interventions [62], reminiscence therapy [42,54], cognitive training therapy [38], and mentalizing imagery therapy [60]).

Consideration of Factors That Influence Health Inequities

Figure 3 provides a visual summary of the number of studies that included or considered the factors listed in the PROGRESS-Plus framework in their report on (1) the design of their mHealth intervention, (2) participant recruitment, (3) study results or findings, and (4) study discussion or limitations.

XSL•FO
RenderX

Figure 3. Consideration of place of residence, race, occupation, gender, religion, education, social capital, socioeconomic status–plus age, disability, and sexual orientation (PROGRESS-Plus) factors in included studies.



Reporting of PROGRESS-Plus Factors in Intervention Design

When describing the design of their interventions, 36% (10/28) of studies provided considerations for ≥1 PROGRESS-Plus factor [37,40,41,46,48,49,56,59,61,63]. Approximately 11% (3/28) of studies considered the place of residence in their recruitment approaches as their interventions were designed specifically for geographically isolated caregivers [40,56,61]. Approximately 11% (3/28) considered languages through the provision of alternative language options in the mobile app, readability (ie, lay language), and accessibility options such as larger font or less text [37,46,49]. Approximately 11% (3/28) described social capital as an element of the intervention itself (eg, intervention aimed at providing a platform to organize and access social support) [56,59,63]. Approximately 7% (2/28) described considerations for caregivers' age in the design of their interventions by improving readability, comprehensibility, and clarity of the language used in the intervention; providing caregivers with assistance in completing web-based forms; and integrating opportunities for regular check-ins to support mHealth tool use [41,46]. One of the studies considered gender or sex, as the intervention was tailored to address the unique needs of caregivers of different genders [46]. Another study considered socioeconomic status by deliberately selecting inexpensive mobile apps and devices [61]. Features of relationships between caregivers and care recipients were considered in the study design such that the mHealth intervention was a collaborative tool whereby older adults and their caregivers worked together on their health management [48]. None of the studies mentioned considering participants' occupation, religion, education, disability, sexual orientation, or time-dependent relationships when describing the design of their mHealth interventions.

Reporting of PROGRESS-Plus Factors in Participant Recruitment

At the participant recruitment stage, 57% (16/28) of studies considered ≥1 PROGRESS-Plus factor [38,40,42,44-46,49,51-53,56-60,64]. Approximately 32% (9/28) considered features of relationships (eg, living situation) [38,40,44,46,51,52,58,60,64]. Approximately 18% (5/28) of studies considered place of residence in participant recruitment (eg, recruiting participants dwelling in rural areas) [40,42,53,56,57]. Approximately 14% (4/28) of studies reported that they used specific recruitment strategies to help ensure that various races, ethnicities, cultures, and languages were represented in their study samples (eg, recruiting from minority populations) [46,49,53,58]. Approximately 11% (3/28) of studies considered age (eg, recruiting older caregivers) [38,45,60].

XSL•FO
RenderX

Approximately 7% (2/28) of studies considered social capital (eg, recruiting caregivers with an existing social support network) [56,59] and 7% (2/28) considered disability (eg, excluding caregivers with sensory impairment) [38,46]. One of the studies considered time-dependent relationships (eg, excluding dyads where the care recipient was awaiting imminent institutional placement) [46], and another considered gender or sex [59] at the stage of participant recruitment. No studies mentioned occupation, religion, education, socioeconomic status, or sexual orientation during participant recruitment.

Reporting of PROGRESS-Plus Factors in Results or Findings

All but 1 study [54] described ≥1 PROGRESS-Plus factor within their results or findings. These factors were typically reported as part of the sample demographics. The key demographic characteristics of the caregivers in the included studies are presented in Table 2. The most commonly reported PROGRESS-Plus factors within the included articles' results or findings were age and gender or sex [37-50,52,53,55-64]; features of relationships [37,39,40,42,43,45,46,48,49,51-53,55-59,61,63]; education [37-39,44,46-50,52,55,56,58,61,63,64]; and race, ethnicity, culture, and language [38,41,43,45-49,53,58,60,62,63]. Other factors reported in the results or findings included socioeconomic status [38,44,48,53,61,63,64], social capital [48,55-57,59,61,64], place of residence [40,49,53,56,62,64], and occupation [50,52,56,61,63,64]. A small number of studies reported on caregivers' disabilities [49,61,63], time-dependent relationships (eg, participants' housing situation) [49,58], and religion [64]. No studies reported on sexual orientation in their results or findings.

XSL•FO
RenderX

Table 2. Demographic characteristics of caregiver participants of included studies.

Study and country	Sample size	Age (years)	Sex, n (%)	Education, n (%)	Ethnicity, n (%)
Banbury et al [56], Australia	69	Mean 62.6 (SD 13.54)	• 50 (72.5) female • 19 (27.5) male	• 6 (8.7) did not complete high school • 6 (8.7) completed high school • 17 (24.6) had technical and further education or trade • 24 (34.8) attended university • 16 (23.2) had postgraduate qualifications	• Not reported
Beentjes et al [37], Netherlands	59	Experimental group mean 65.61 (SD 10.196); control group mean 68.03 (SD 11.675)	• 38 (64.4) female • 21 (35.6) male	• 12 (20.3) had secondary education (vocational) • 8 (13.6) had secondary education (academic) • 11 (18.6) had further education (vocational) • 20 (33.9) had higher education (vocational) • 8 (13.6) had higher education (academic)	• Not reported
Breebaart and van Groenou [57], Netherlands	7	1 (14.3%) middle-aged, 1 (14.3%) aged between 60 and 65, and 5 (71.4%) aged ≥70	• 3 (42.9) female • 3 (42.9) male • 1 (14.3) not specified	• 4 (57.1) had low education • 2 (28.6) had average education • 1 (14.3) did not specify	• Not reported
Brown et al [58], United States	11	Mean 56.6 (SD 13.6)	• 9 (81.8) female • 2 (18.2) male	• Not reported	• 3 (27.3%) White • 7 (63.6%) African American • 1 (9.1%) Hispanic • 1 (9.1) other
Callan et al [38], United States	27	Mean 74.61 (SD 6.52)	• 22 (81.5) female • 5 (18.5) male	• 11 (40.7) had middle school to technical school education • 14 (51.9) had some college to college graduate education • 2 (7.4) had some postgraduate to postgraduate degree	• 26 (96.3) White
Dam et al [59], Netherlands	10	Range 49-71	• 6 (60) female • 4 (40 male)	• Not reported	• Not reported
Davis et al [43], United States	4	Mean 52	• 4 (100) female	• Not reported	• Not reported
Davis et al [63], United States	3	Range 54-85	• 3 (100) female	• 2 (66.7) attended college • 1 (33.3) had a master's degree	• 3 (100) White
Garvelink et al [49], Canada and France	10	Mean 56.9 (SD 14)	• 6 (60) female • 4 (40) male	• 10 (100) had a university degree	• Not reported
Hastings et al [45], United States	40	Mean 64.7 (SD 10.8)	• 40 (100) female	• Not reported	• 11 (27.5) Black
Hughes et al [50], United States	10	Mean 60 (range 48-76)	• 10 (100) female	• 10 (100) had high school education • 9 (90) had higher education	• Not reported
Kales et al [46], United States	57	Mean 65.9 (SD 14.0)	• 43 (75.4) female • 14 (24.6) male	• 48 (84.2) had greater than high school education • 9 (15.8) had high school or GEDa	• 36 (63.2) White • 18 (31.6) African American • 3 (5.3) other

XSL•FO
RenderX

Study and country	Sample size	Age (years)	Sex, n (%)	Education, n (%)	Ethnicity, n (%)
Killin et al [51] [51], Scotland	10	Not reported	• Not reported	• Not reported	• Not reported
Lai et al [44], Hong Kong, China	60	Experimental group mean 72.43 (SD 0.80, range 66-82); control group mean 71.83 (SD 0.80, range 66-82)	• 35 (58.3) female • 25 (41.7) male	• Experimental group: 7.90 (SD 0.25, range 5-11) years of education • Control group: 7.04 (SD 0.31, range 5-9) years of education	• Not reported
Lai et al [39], Germany	24	Mean 62.4 y (SD 16.0, range 31-83)	• 9 (37.5) female • 15 (62.5) male	• 11 (45.8) had >12 years of education	• Not reported
Park et al [64], South Korea	24	Experimental group mean 54.50 (SD 3.71); control group mean 61.00 (SD 6.42)	• 14 (58.3) female • 10 (41.7) male	• 15 (62.5) were high school graduates or below • 9 (37.5) were college graduates or above	• Not reported
Ptomey et al [47], United States	9	Mean 67	• 3 (33.3) female • 6 (66.7) male	• 3 (33.3) had high school diploma or GED • 6 (67.6) attended postgraduate classes	• 8 (88.9) White • 1 (11.1) Black
Quinn et al [48], United States	12	Mean 54.8 (SD 13.3)	• 11 (91.7) female • 1 (8.3) male	• 6 (50) had a business or some college degree or graduate degree • 6 (50) graduated school	• 6 (50) Black • 6 (50) White
Rathnayake et al [52], Australia	10	8 (80%) aged <65; 2 (20%) aged ≥65	• 9 (90) female • 1 (10) male	• 5 (50) had high school education and below • 5 (50) had above high school education	• Not reported
Ruggiano et al [53], United States	36	Mean 65.7 (range 42-89)	• 26 (72.2) female • 10 (27.8) male	• Not reported	• 13 (36.1) non-Hispanic White • 23 (63.9) African American
Ryan et al [54], United Kingdom	17	Mean 69.1 (SD 15.1, range 31-91)	• 13 (76.5) female • 4 (23.5) male	• Not reported	• Not reported
Salin and Laaksonen [40], Finland	20	Range 61-88	• 15 (75) female • 5 (25) male	• Not reported	• Not reported
Sikder et al [60], United States	17	Mean 66.52 (SD 9.61)	• 12 (70.6) female • 5 (29.4) male	• Not reported	• 17 (100) White
Sourbeer et al [41], United States	46	42 (91.3%) aged >60; 4 (8.7%) aged <60	• 38 (82.6) female • 8 (17.4) male	• Not reported	• 39 (84.8) White • 6 (13.0) African American • 1 (2.2) Hispanic
Span et al [55], Netherlands	12	Mean 54.3 (range 19-86)	• 7 (58.3) female • 5 (41.7) male	• 1 (8.3) had low education • 4 (33.3) had medium education • 6 (50) had high education • 1 (8.2) did not specify	• Not reported
Stutzel et al [61], Brazil	38	Mean 61 (SD 10.75)	• 32 (84.2) female • 6 (15.8) male	• 21 (55.3) had ≤12 years of education • 17 (44.7) had >12 years of education	• Not reported
Tyack et al [62], United Kingdom	12	Mean 66 (range 48-77)	• 10 (83.3) female • 2 (16.7) male	• Not reported	• 12 (100) White

XSL•FO
RenderX

Study and country	Sample size	Age (years)	Sex, n (%)	Education, n (%)	Ethnicity, n (%)
Watcharasarnsap et al [42], Thailand	60	8 (13.3%) aged between 18 and 27, 19 (31.7%) aged between 28 and 37, 15 (25%) aged between 38 and 47, 10 (16.7%) aged between 48 and 57, and 8 (13.3%) aged ≥58	• 31 (51.7) female • 29 (48.3) male	• Not reported	• Not reported

[a]GED: General Educational Development.

Reporting of PROGRESS-Plus Factors in Discussion or Limitations

Approximately 79% (22/28) of studies considered ≥1 PROGRESS-Plus factor in the discussion or limitations sections of their studies [37,38,40,41,43,44,47-56,58-63]. The most frequently discussed PROGRESS-Plus factors in the included articles' discussion or limitations were age [37,38,40,41,43,48,50,51,53,54,60,62,63], such as challenges faced by older caregivers in using mobile devices; race, ethnicity, culture, and language [40,41,43,47,49,52,53,55,58,60], such as a lack of diversity of the study sample; and place of residence [40,44,47,49,51,53,55,56,61], such as challenges related to the lack of access to stable internet in rural locations. Other PROGRESS-Plus factors described in the study discussions or limitations were gender or sex [38,41,52,54,55,63], education [37,38,49,52,56,63], and socioeconomic status [44,47,48,52,53,63]. To a lesser extent, caregivers' social capital [48,56,59], disabilities [38,49], features of relationships (eg, nature of relationship between caregiver and care recipient) [54,55], and time-dependent relationships (eg, the impact of COVID-19 on the amount of time caregivers could spend visiting the care recipient) [37,49] were also discussed. No studies considered occupation, religion, or sexual orientation in their discussion or limitations sections.

Quantitative Caregiver Outcomes

Outcomes Relating to Caregiving

Approximately 21% (6/28) of studies assessed the impact of mHealth interventions on outcomes related to caregivers' capabilities or experiences in providing care. These outcomes included caregivers' self-efficacy [44,63], sense of competence [37] and confidence [46] in their caregiving role, knowledge related to the care recipient's condition [63], positive care experience [37], and caregiver burden [43,44,46,63,64]. Although some studies found that caregiving self-efficacy and knowledge improved after the implementation of an mHealth intervention [44,63], other studies observed no difference after the intervention in caregivers' sense of competence [37], confidence [46], or positive caregiving experience [37].

Approximately 14% (4/28) of studies using the Zarit Burden Inventory [65] found that mHealth interventions led to improvements in caregiver burden [43,44,46,64]. However, one of the studies, which specifically assessed caregiver burden related to the management of urinary incontinence, found that burden was similar before and after the mHealth intervention [63]; however, study investigators noted that the intervention did not worsen caregiver burden [63].

Outcomes Relating to Caregivers' Health and Well-being

Approximately 39% (11/28) of studies assessed the impact of mHealth interventions on various aspects of caregivers' health and well-being [37,40,42-44,46,47,60-62,64]. Impacts on caregivers' mental and psychological health status were assessed in 25% (7/28) of studies [42,44,46,60-62,64], with generally positive results. Specifically, mental health status [44], psychological well-being [42], depression [46,60], mood [60], distress [46], and fatigue [64] were each noted to have improved after the implementation of an mHealth intervention. For example, the implementation of the WeCareAdvisor tool, designed to provide caregivers with peer navigation, information, and daily messaging, led to significant improvement in self-reported distress (−6.08, SD 6.31 points; P<.001) [46]. In this study, those in the control group demonstrated a significant decrease in their confidence in caregiving (−6.40, SD 10.30; P=.002) [46]. Conversely, a study that assessed caregiver stress by testing cortisol levels in saliva in a pretest-posttest design found no differences after the use of an mHealth intervention designed to manage the behavioral and psychological symptoms of dementia [64]. Caregivers' self-appraised happiness was also unchanged after the intervention in one of the studies [62].

Approximately 11% (3/28) of studies assessed outcomes related to caregivers' physical health and well-being [44,47,64]. Caregivers self-reported improvements in their general physical health status following the use of an mHealth intervention to support the well-being and community living of older adults and their spousal caregiver dyads [44]. Ptomey et al [47], who implemented an mHealth app to encourage exercise, observed that caregivers' weekly moderate physical activity increased by 49 minutes (30% increase) per week over the 12-week intervention period, whereas light physical activity increased by 11.6 minutes (3% increase) per week. However, Park et al [64] found no difference in caregivers' sleep quality after the implementation of a supportive mHealth app.

Approximately 14% (4/28) of studies used caregivers' quality of life as an outcome measure for their respective interventions, with mixed findings. Ptomey et al [47] found nonsignificant trends toward improvement in quality of life after the implementation of an mHealth intervention. Beentjes et al [37] and Tyack et al [62] found no significant changes in quality of life following their interventions aimed at supporting caregivers in finding user-friendly apps and viewing art to encourage

XSL•FO
RenderX

therapeutic reminiscence, respectively. Salin and Laaksonen [40] observed that some aspects of quality of life, in fact, worsened, albeit mildly (breathing, sexual activity, vitality, depression, and usual activities). One of the studies assessed the impact of an mHealth intervention on caregivers' social engagement and found high positive responses using the Kaye Gain Through Involvement Scale [66], suggesting that the gains in well-being experienced while using the mHealth intervention may be applicable when tested in a larger sample [43]. However, the study investigators noted that their sample was meant only for determining intervention efficacy and warranted testing with a larger sample [43].

Outcomes Related to Usability, Feasibility, and Acceptability of mHealth Interventions

Half of the reviewed studies assessed outcomes related to the usability, acceptability, or feasibility of mHealth interventions for caregivers of older adults [38-41,45,47,48,57-63].

Approximately 32% (9/28) of studies measured the usability or ease of use of mHealth interventions by caregivers [40,41,45,47,48,58,59,61,63]. Approximately 14% (4/28) of articles used the System Usability Scale [67] to do so; usability scores varied across studies, ranging from marginally acceptable [45], moderate [48], and good to excellent [61]. Only 4% (1/28) of studies compared the system usability scores across 2 phases of their mHealth app intervention. Sourbeer et al [41] found that usability did not significantly improve in a subsequent version of their mHealth app updated in response to participant feedback. The remaining 18% (5/28) of studies assessed caregivers' ease of use or perceived user-friendliness of the mHealth intervention using descriptive statistics or averaged Likert scale scores. These studies generally reported positive results, suggesting that caregivers believed the interventions were easy or very easy to use [40,47,58,59,63].

Approximately 21% (6/28) of studies examined caregivers' satisfaction or positive feelings toward the intervention [39,40,47,48,58-61]. Most reported that caregivers were generally satisfied with the mHealth intervention, perceived the intervention as relevant and useful to their caregiving activities, and felt positive about their experiences with the intervention [39,40,47,48,58-61]. However, greater technical difficulties were reported in a study of participants who lived rurally and reported lower levels of satisfaction [40].

Approximately 29% (8/28) of studies explored the feasibility of an mHealth intervention by measuring the regularity, frequency, and extent of its use by caregivers over the intervention period [38,57-60,62]. Use varied across the included studies, and investigators did not consistently establish expectations of use for their participants nor defined what constituted adequate use of the intervention. Tyack et al [62] reported that the participants used their app at least five times during the intervention period, as suggested by the study investigators. Callan et al [38] found that 22 out of 27 (81.5%) caregivers used the mHealth intervention regularly (as defined by the study investigators as at least 3 weeks out of the 4-week intervention period). Baseline caregiver stress, worry, and sleep quality did not adversely affect the use of the mHealth intervention, and caregivers with the highest self-reported stress and worry reported the highest levels of mHealth intervention use [38]. Sikder et al [60] reported that over half of their 17 study participants accessed ≥75% of the informational documents in their mHealth app. The remaining 11% (3/28) of studies reported varying frequencies or hours of use per week during the intervention period [57-59]; however, these studies did not comment on whether these frequencies constituted low, medium, or high use of their mHealth interventions.

Approximately 11% (3/28) of studies assessed feasibility by measuring the intervention attendance and retention of caregivers during the intervention period [40,45,47]. The attendance rates for caregivers varied from 72% (13.7/19) [40] to 97.1% (34/35) [45]. Ptomey et al [47] and Hastings et al [45] reported similar figures (7/9, 78% dyads, and 31/40, 78% dyads, respectively) for the caregiver–care recipient dyads completing their interventions.

Other feasibility measures used by the reviewed studies included the extent to which caregivers followed or adhered to the mHealth intervention [38,63]. Callan et al [38] reported that caregivers' continued engagement in a cognitive training mHealth intervention program was evidenced by improvements in their ability to perform cognitive training tasks. Davis et al [63] reported that caregivers were capable of learning and implementing the prompted toileting strategies to support care recipients with the help of an mHealth intervention, as evidenced by a reduction in care recipient wetness in 2 out of 3 participant dyads.

Qualitative Caregiver Findings

Overview

Of the 28 studies, 7 (25%) qualitative studies and 7 (25%) mixed methods studies presented findings relating to caregivers' experiences of engaging in mHealth interventions [49-62]. These qualitative findings included (1) positive impacts of caregivers' experiences with mHealth interventions, (2) challenging aspects of caregivers' experiences with mHealth interventions, (3) barriers to caregivers' engagement with mHealth interventions, and (4) caregivers' suggestions to improve mHealth interventions.

Positive Experiences With mHealth Interventions

Most studies highlighted promising findings related to the positive impacts of caregivers' experiences with mHealth interventions. Participants across the included studies found mHealth interventions to be helpful, user-friendly, and easy to understand [49,50,54,55]. mHealth interventions were perceived to help caregivers connect with the care team and provide care for their loved ones [53,55,57,60,61]. The information provided through mHealth interventions was described as relevant to addressing participants' educational needs [49,52]. Caregivers also valued the role of mHealth interventions in detecting their stress levels [50] and facilitating timely connections to a diverse range of professional services and social support [49,52,54,56,62]. Participants in the included studies reported benefits to their emotional and cognitive well-being [60,62] and described reappraising and feeling closer to the care recipient [54,62]. The mobile delivery of the interventions also contributed to feelings of safety and security, as caregivers could

XSL•FO
RenderX

participate from their homes [54,56]. Although some participants initially felt a lack of confidence in using technology, caregivers in 7% (2/28) of studies reported becoming more engaged and comfortable over time by integrating the mHealth intervention into their lives [54,57].

Challenging Experiences With mHealth Interventions

Several studies described the negative aspects of caregivers' experiences with using mHealth interventions, although these were often reported as being applicable to only a minority of participants. Approximately 11% (3/28) of studies indicated that some participants felt that the mHealth intervention was too complex or difficult to understand [49,51,60]. In another study, participants felt that the intervention included questions that were overly obtrusive or confronting; for example, participants were not always comfortable answering questions they perceived as challenging [55]. Some studies highlighted caregivers' concerns regarding the potentially detrimental impacts of mHealth interventions; for example, interventions that facilitated reminiscence could trigger painful memories and lower mood [54,62]. Hughes et al [50] further described caregivers' concerns regarding the diversion of their time and attention toward the mHealth intervention and away from the care recipient. One of the studies highlighted the preference of some participants for in-person interventions, citing physical contact as an important element of care (eg, hugging), which was not possible in a digital environment [56].

Barriers to Caregivers' Engagement With mHealth Interventions

Caregivers relayed frustration with the usability of mHealth interventions, including difficulties navigating the intervention on their mobile devices [49,50,62]. Challenges included print that was too small [49,50], screens that were overly sensitive or had too much glare [62], and language that was too complex [49]. Several studies highlighted a lack of familiarity or experience with technology as a key barrier to the use of mHealth interventions, particularly for older caregivers [51-53,55]. The busy schedules of caregivers for older adults were also identified as a barrier to regular mHealth intervention use, particularly if caregivers were often pulled away from their devices by care recipients or if they were experiencing health issues themselves [50,52,58,60].

In other cases, participants felt that the intervention's content was not relevant to their immediate needs [49,51] or lacked realism (eg, lack of ethnic diversity among actors portraying caregivers in the mHealth intervention) and up-to-date links to relevant resources [49]. Other barriers included the prohibitive cost of mobile devices and internet or data plans [52] and the availability of a stable internet connection in rural regions [56].

Caregivers' Perspectives Regarding Next Steps

Qualitative findings frequently incorporated participants' suggestions to make mHealth interventions more user-friendly and accessible to caregivers. Suggestions included simplifying the intervention's interface or instructions, enlarging text and images, and including subtitles on video resources for individuals with hearing impairment [49,52,61,62]. Participants voiced the need for ongoing technical support, particularly for caregivers who were unfamiliar with using mobile devices [51,56].

The participants also made suggestions to develop more relevant and up-to-date content for mHealth interventions. Several studies highlighted the need to embed local and national services for caregiver support, including interventionists and respite care [58-60]. For interventions that targeted the caregiver–care recipient dyad, participants highlighted the need for more information specifically related to their own health, such as healthy coping [49,52,58,61]. Participants also called for greater emphasis on topics that caregivers often find difficult, including information about deciding to move to a care home, managing activities of daily living and aggressive behaviors, and resources for individuals experiencing abuse [49,52,58].

Other findings suggested to improve mHealth engagement among caregivers included greater ethnic diversity portrayed within the mHealth intervention [49], establishing a reward system to encourage regular use [50], and creating a component for the care recipient to be included when the caregiver uses the mHealth intervention [50].

Discussion

Principal Findings

This systematic review examined how health and social equity are considered in the design, implementation, and evaluation of mHealth interventions developed for caregivers of older adults using the PROGRESS-Plus framework. The interventions described in the included studies were designed to create linkages between caregivers and external supports, streamline and optimize caregiving activities, and encourage a focus on caregiver health and well-being. As such, evidence on the impacts of caregiver-focused mHealth interventions was synthesized across a range of outcomes.

The findings indicate that health and social factors are not consistently taken into consideration when designing research studies (ie, used to develop and guide recruitment and intervention design). Furthermore, participant characteristics are most often only reported within study results when summarizing participant characteristics or when identifying limits to the generalizability of the findings. However, this review highlights how mHealth interventions are well-positioned to improve caregivers' self-efficacy and knowledge, their perceived mental and physical wellness, and their relationships with care recipients. The usability and acceptability of mHealth interventions were characterized by ease of use, ease of navigating technical challenges, and relevance of intervention content to the caregivers' individual roles and context.

Consideration of PROGRESS-Plus Factors in Studies on mHealth Interventions for Caregivers of Older Adults

Overview

Most studies in this systematic review on mHealth interventions for caregivers of older adults considered some PROGRESS-Plus factors, particularly when describing their study samples. However, such demographic reporting reflects standardized

XSL•FO
RenderX

reporting practices of participant composition rather than deliberate and targeted approaches to recruiting caregivers across sociodemographic characteristics to determine whether an intervention is suitable for a diversity of participants. The factors described in the following sections were considered critical in the intervention design.

Gender Sex or Sexual Orientation

Importantly, few studies considered actively recruiting caregivers of different self-reported genders or considered the relevance of gender in intervention design or implementation. Research suggests that biological and gender differences affect health across a range of parameters such as risk, disease incidence, and the need for health services [27]. Furthermore, sexual orientation was, in fact, eclipsed across all studies, particularly when many studies focused on caregiver health and well-being, which includes the relationship they have with care recipients. Recent evidence indicates that sexual and gender minority caregivers, such as those identifying as queer and transgender, report higher depressive symptoms (78%) than the overall population of caregivers of people with dementia (34%) [68]. This finding highlights the importance of diversifying samples across genders and sexual orientations to reliably assess and address caregivers' mental health. The importance of considering the intersections among gender, sexual orientation, and other sociodemographic factors was also highlighted in the survey of a cross-sectional sample of members of the National Alliance for Caregiving. Caregivers who identified as lesbian, gay, bisexual, and transgender were more likely to be racially and ethnically diverse and represent lower socioeconomic classes than those who did not [69].

Education

Education, although frequently reported in demographics, was also rarely considered as an important factor in informing intervention design and recruitment. Women with lower education are more likely to assume caregiving roles than those who have had additional educational opportunities [11]. Lower literacy levels among caregivers can affect their ability to navigate the health system and locate appropriate support for themselves and their care recipient [70], factors that can directly influence the design and usability of mHealth interventions. For example, lower literacy can affect comprehension of text-based content in mHealth apps, the ability to correctly enter spelled words in search functions, and the ability to navigate app menus [71]. The importance of designing mHealth interventions that account for varying levels of educational background is underscored by the association of literacy with health and digital literacy [72].

The findings of the included studies suggest that experience with technology can be a key barrier to the use of mHealth interventions, particularly among older caregivers [51-53,55]. A survey of a broad age range of caregivers suggests that younger caregivers (aged <50 years) are more than twice as likely than older caregivers to be receptive to using mHealth apps to support them in their caregiving roles [24]. For older adults, trust in technology as it relates to privacy and access to information can be an important factor in the use of mHealth interventions, especially given the heterogeneity of this population [24,73]. These findings suggest that exploring barriers and facilitators, as identified by the included qualitative studies, aimed at educating older adults on how to use mHealth interventions is essential to facilitate perceived trust, comfort, and usability of technology. Thus, beyond education as a social determinant of health, wide disparities exist across caregivers in comfort with using various technologies, such as tablets, iPads, and mobile phones [73].

Socioeconomic Status

Socioeconomic status was minimally considered in the intervention design and was most often addressed when describing sample characteristics. Multiple studies reported providing participants with devices to support the use of mHealth apps [37-41,43,45-47,51,53-57,61-63]. In some cases, participants were allowed to keep the devices; however, especially in those instances in which they were not, the feasibility of such interventions for caregivers across income levels needs to be explored.

Some interventions were designed to facilitate communication access to health professionals and other individuals (eg, support groups), highlighting the need for access to a reliable internet connection. This lack of access may be due in part to financial constraints, as a survey of caregivers in the United States found that cost was a commonly reported barrier to the use of technology [74]. Furthermore, older adults living on fixed incomes may be reticent to spend money on devices they do not value or find overly complicated [75]. Importantly, older caregivers tend to have fewer technological devices than their younger peers, and these technologies are often used for communication purposes rather than health management purposes [18]. Although most caregivers report valuing technology, those that use it for health-related activities tend to use it for targeted caregiving activities such as medication tracking or safety [18]. Therefore, additional support or education may be required to increase caregiver uptake of mHealth interventions as a tool for addressing broader caregiver needs such as communication with health teams or liaising with other caregivers. Computers and smartphones are increasingly being owned by people with higher income and education, and the provision of caregiver support through mHealth apps could increase inequalities if economic resources are not considered in the design and implementation of these interventions [71].

Culture, Language, and Race

The nature of caregiver–care recipient relationships can be an important factor in the design of mHealth tools, particularly when it comes to cultural expectations of family members, gender roles, and other caregiver demographics. The included studies had samples primarily made up of women, validating the literature that suggests women are most likely to provide caregiving support, corroborating cultural norms across a range of identities [76]. However, these studies did not address how intersecting identities (eg, culture, gender, race and ethnicity, and socioeconomic status) might shape expectations and responsibilities within a caregiving role [11,12,68]. Research suggests that culture strongly affects caregiving but that cultural influences on the caregiver role must be understood within the context of race and gender socialization [77]. For example,

XSL•FO
RenderX

individualistic or Western notions of strategies to address caregiver burdens, such as spending time alone or sharing caregiving responsibilities with friends or family, might not resonate with caregivers from other cultures, particularly those with a strong sense of filial responsibility or immigrant caregivers without local support [78]. Furthermore, mHealth apps not provided in caregivers' first languages decrease accessibility and would require careful translation and cultural adaptation to remain meaningful [79]. The impact of these factors on caregiver-specific outcomes, such as caregiving self-efficacy, health and well-being, and technology usability, is yet to be explored. Intersecting identities are increasingly important to consider when tailoring web-based caregiver interventions to participants' individual needs [19].

mHealth Interventions Developed for Caregivers of Older Adults

Impacts of mHealth Interventions on Caregiver Health and Wellness

Studies evaluating mobile technology interventions aiming to promote caregivers' perceived mental and psychological health reported benefits to their emotional and cognitive well-being [60,62]. Some of these interventions, such as the videoconferencing platform developed by Lai et al [44], were designed in lieu of in-person community services, following shelter-in-place orders during the COVID-19 pandemic. Connecting caregivers to professional and peer support using web-based technologies has been shown to improve mental health outcomes and can help caregivers overcome common access-related barriers related to PROGRESS-Plus factors, such as geographical and time constraints or community mobility limitations related to physical or mental health [18-20]. However, findings from the included studies suggest that caregivers still require opportunities for in-person interaction (eg, hands-on training from a health care provider to successfully use external support systems), suggesting that the impact of hybrid models of interventions to improve caregiver health and wellness is not well understood [20]. Furthermore, a review of these interventions using the PROGRESS-Plus factors suggests that, although caregivers stand to benefit from mHealth interventions and many older adults report being comfortable with smartphone use, uptake may continue to be constrained if support is not provided to help caregivers learn and familiarize themselves with mHealth apps at the outset [80]. Hybrid approaches have the potential to increase caregiver self-efficacy, as opposed to overwhelming caregivers with new tools and technology, which warrants further research.

Supporting the Caregiver Role Through mHealth Interventions

Caregivers' ability to perform their roles was a key focus of the examined mHealth interventions and outcomes of interest within the included studies. Although some interventions focused on creating external structures that facilitated responsibilities of providing care (eg, medication alarms, and checklists), the use of these tools had the potential to complicate caregiving responsibilities. For example, in one case, caregivers described that the increased *screen time* to engage in the intervention was taking away from the time they had to complete other caregiving tasks [58]. The impact of such detrimental experiences, as they relate to, for example, PROGRESS-Plus factors of gender-informed cultural caregiving roles, features of relationships, or caregiver disability, is not well understood. Wasilewski et al [34] found that caregivers' decline in web-based intervention use may be attributed to a malalignment with their specific needs and capabilities across the caregiving trajectory. In such cases, it is important for those recommending mHealth interventions to caregivers to consider whether a particular intervention itself might increase the caregiver burden [81]. Furthermore, research suggests that if older adults perceive an mHealth app to be beneficial to their health and well-being, their likelihood of ongoing and increased engagement with other apps increases [82]. Individualized tailoring of mHealth apps and providing the necessary access and universal design can foster equitable uptake and increase the potential benefits of mHealth interventions.

Usability, Feasibility, and Acceptability of mHealth Interventions

Overall, caregivers in the included studies were generally comfortable using mHealth interventions and reported positive impacts on their caregiving role [49,50,54,55]. However, findings such as the prohibitive costs associated with mobile devices and internet and data plans, in combination with the quality of internet provision to those living in rural settings, highlight the importance of equitable service provision across the PROGRESS-Plus factors [52,56]. The findings of this review also showed that 64% (18/28) of studies [37-41,43,45-47,51,53-57,61-63] provided participants with the devices required to engage in the interventions, suggesting that the economic feasibility of these interventions needs to be better understood.

Technical features such as app use data may provide valuable insights into the frequency and applicability of interventions to caregiver needs and their unique lifestyles. Furthermore, researchers have been urged to include older adults and their caregivers in the design and development of mobile app technologies [48]. However, a minority of the studies included in this review described stakeholder input as a component of their intervention design or implementation [40,50,53-55,58,61]. Co-design approaches present important opportunities for engaging diverse populations to help ensure that mHealth interventions are inclusive and accessible.

Implications

Moving forward, an important reminder is that social determinants of health should be consciously considered in all aspects of mHealth intervention design and implementation to avoid perpetuating inequities experienced by historically and currently systemically disadvantaged caregivers of older adults living with chronic conditions [25,83]. Purposeful efforts to include a diverse range of participants in research, such as evidence-based recruitment strategies, can help redress these potential inequities and inform the development of more inclusive interventions [84,85]. The PROGRESS-Plus framework is an appropriate tool to help ensure that a health

XSL•FO
RenderX

and social equity lens is applied in research design and reporting, the use of which should be widely endorsed [27,86].

This review highlights the need for high-quality mHealth studies. Particular attention must be paid to improving the design of mHealth interventions and ensuring equality in access and adoption of mHealth interventions [71]. Participatory action approaches to research, such as co-design, are ideal for ensuring that mHealth interventions meet the needs of diverse caregivers. Furthermore, inclusive design principles can be used in more traditional research methodologies to ensure that mHealth interventions do not amplify health disparities. This could be achieved by accommodating low literacy by including audio narration and visual depictions or by directing funding to increase access to human resource infrastructures (eg, technical support) that promote mHealth interventions in remote or low-income regions [71].

Strengths and Limitations

The studies included in this systematic review represent the diversity of mHealth interventions that have been conceptualized and created to address caregiver needs. Unfortunately, many studies were found to be poorly designed and executed. Although half of the included studies assessed usability, feasibility, and acceptability of mHealth interventions, which are all important aspects of technology use, many of these used qualitative approaches and lacked overall methodological rigor. Given the variety of mHealth apps, technological devices, and implementation protocols, equivalent comparisons could not be made across studies. A small number of studies were identified evaluating the impact of caregiver-focused interventions on caregiver-specific outcomes, limiting the ability to make conclusive recommendations to guide practice. Encouragingly, some of the included quantitative studies that used valid and reliable standardized tools thoroughly described their approach to statistical analysis and generally addressed fidelity of intervention delivery.

In this review, multiple steps were taken to achieve methodological rigor. The review was conceptualized and designed using an equity framework and the best evidence on interventions for caregivers of older adults. The search strategy was developed in consultation with a health research librarian, and database searches, screening, data extraction, and risk of bias evaluations were conducted in duplicate, with a strong agreement between reviewers. The review protocol was also made publicly available a priori and was adhered to without any deviations. In addition, the PRISMA and PRISMA-Equity guidelines guided each phase of this study [28,87].

Inevitably, this study has some limitations. Although these searches were conducted by health and rehabilitation investigators across 3 large academic institutions in the Global North, these institutions use similar health research databases and search algorithms, which can affect future reproducibility (ie, replicating searches in different institutions with different journal accesses). The identification of potentially eligible literature from the Global South, other disciplines beyond health research (eg, technology literature databases), or those that are categorized in other ways (eg, gray literature) is another limitation of this review. However, this study highlights that research on mHealth interventions for caregivers of older adults primarily occurs within applied health settings. As such, future reviews should examine non–peer-reviewed evidence such as reports and program evaluations produced by the government and health authorities that trial mHealth interventions.

This study could have been further strengthened by involving additional team members, such as administrators of clinical settings who would implement mHealth interventions and, most importantly, caregivers of older adults themselves. By selecting the PROGRESS-Plus framework as a theoretical guide, this study did not examine the included interventions and investigations in light of compounding factors that disadvantaged caregivers (eg, impact of the intervention on older women living in rural settings) or capture other health and social factors beyond the framework (eg, access to health insurance). However, using the framework as an approach to name and identify how key individual factors have been considered in intervention design and evaluation, this study has set the stage for future investigations that examine the confluence of multiple social determinants of health.

Conclusions

mHealth supports are well-positioned to support caregivers of older adults by providing them with information, communication, and assistance in their caregiving role. However, access, uptake, and the ability to benefit from this technology can be affected by the social determinants of health and inequities among caregivers. This systematic review of mHealth interventions to support caregivers of older adults suggests that these tools are well-received by caregivers and have the potential to support caregivers across a variety of parameters by facilitating education, communication, and a sense of security for caregivers. The social determinants of health and equity factors are not widely considered in the design and implementation of mHealth interventions, although these parameters are frequently collected for demographic reporting. Recognizing that there are many challenges in designing and implementing mHealth interventions that are equitable, going forward, it will be important to strive for greater inclusion of the social determinants of health at all stages of mHealth development and implementation if there is to be widespread successful uptake of this supportive technology.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategies for all databases.

XSL•FO
RenderX

[PNG File , 393 KB - aging_v5i3e33085_app1.png]

Multimedia Appendix 2
Summary table of included studies.
[PNG File , 3092 KB - aging_v5i3e33085_app2.png]

Multimedia Appendix 3
Risk of bias assessments.
[PNG File , 739 KB - aging_v5i3e33085_app3.png]

References

1. Evidence profile: caregiver support. World Health Organization. 2017. URL: https://www.who.int/ageing/health-systems/icope/evidence-centre/ICOPE-evidence-profile-caregiver.pdf?ua=1 [accessed 2021-08-15]
2. Liu Z, Heffernan C, Tan J. Caregiver burden: a concept analysis. Int J Nurs Sci 2020 Oct 10;7(4):438-445 [FREE Full text] [doi: 10.1016/j.ijnss.2020.07.012] [Medline: 33195757]
3. Roth DL, Fredman L, Haley WE. Informal caregiving and its impact on health: a reappraisal from population-based studies. Gerontologist 2015 Apr;55(2):309-319 [FREE Full text] [doi: 10.1093/geront/gnu177] [Medline: 26035608]
4. Wolff JL, Mulcahy J, Huang J, Roth DL, Covinsky K, Kasper JD. Family caregivers of older adults, 1999-2015: trends in characteristics, circumstances, and role-related appraisal. Gerontologist 2018 Nov 03;58(6):1021-1032 [FREE Full text] [doi: 10.1093/geront/gnx093] [Medline: 28637266]
5. Ploeg J, Garnett A, Fraser KD, Baird LG, Kaasalainen S, McAiney C, et al. The complexity of caregiving for community-living older adults with multiple chronic conditions: a qualitative study. J Comorb 2020;10:2235042X20981190 [FREE Full text] [doi: 10.1177/2235042X20981190] [Medline: 33403202]
6. Peacock S, Sethi B, Williams A, Duggleby W, Bayly M, Swindle J, et al. Older adult spouses with multiple chronic conditions: challenges, rewards, and coping strategies. Can J Aging 2017 Jun;36(2):209-222. [doi: 10.1017/S0714980817000095] [Medline: 28349856]
7. Wolff JL, Spillman B. Older adults receiving assistance with physician visits and prescribed medications and their family caregivers: prevalence, characteristics, and hours of care. J Gerontol B Psychol Sci Soc Sci 2014 Nov;69 Suppl 1:S65-S72 [FREE Full text] [doi: 10.1093/geronb/gbu119] [Medline: 25342825]
8. Vaingankar JA, Chong SA, Abdin E, Picco L, Jeyagurunathan A, Zhang Y, et al. Care participation and burden among informal caregivers of older adults with care needs and associations with dementia. Int Psychogeriatr 2016 Feb;28(2):221-231 [FREE Full text] [doi: 10.1017/S104161021500160X] [Medline: 26478530]
9. Bom J, Bakx P, Schut F, van Doorslaer E. The impact of informal caregiving for older adults on the health of various types of caregivers: a systematic review. Gerontologist 2019 Sep 17;59(5):e629-e642 [FREE Full text] [doi: 10.1093/geront/gny137] [Medline: 30395200]
10. Musich S, Wang SS, Kraemer S, Hawkins K, Wicker E. Caregivers for older adults: prevalence, characteristics, and health care utilization and expenditures. Geriatr Nurs 2017;38(1):9-16 [FREE Full text] [doi: 10.1016/j.gerinurse.2016.06.017] [Medline: 27477086]
11. Hengelaar AH, Wittenberg Y, Kwekkeboom R, Van Hartingsveldt M, Verdonk P. Intersectionality in informal care research: a scoping review. Scand J Public Health 2021 Jul 07:14034948211027816 [FREE Full text] [doi: 10.1177/14034948211027816] [Medline: 34232094]
12. Cook SK, Cohen SA. Sociodemographic disparities in adult child informal caregiving intensity in the United States: results from the new national study of caregiving. J Gerontol Nurs 2018 Sep 01;44(9):15-20. [doi: 10.3928/00989134-20180808-05] [Medline: 30148528]
13. Garnett A, Ploeg J, Markle-Reid M, Strachan PH. Formal health and social services that directly and indirectly benefit stroke caregivers: a scoping review of access and use. Can J Nurs Res 2022 Jun;54(2):211-233 [FREE Full text] [doi: 10.1177/08445621211019261] [Medline: 35130749]
14. King RB, Semik PE. Stroke caregiving: difficult times, resource use, and needs during the first 2 years. J Gerontol Nurs 2006 Apr;32(4):37-44. [doi: 10.3928/00989134-20060401-07] [Medline: 16615711]
15. Simon C, Kumar S, Kendrick T. Formal support of stroke survivors and their informal carers in the community: a cohort study. Health Soc Care Community 2008 Dec;16(6):582-592. [doi: 10.1111/j.1365-2524.2008.00782.x] [Medline: 18371168]
16. Measuring the family/friend caregiver experience in Ontario: pilot study results. Health System Performance Research Network (HSPRN). URL: https://www.huntingtonsociety.ca/wp-content/uploads/2016/12/Caregiver-Study-Pilot-Results_HSPRN-Final-Report.pdf [accessed 2021-08-15]
17. Taylor MG, Quesnel-Vallée A. The structural burden of caregiving: shared challenges in the United States and Canada. Gerontologist 2017 Feb;57(1):19-25. [doi: 10.1093/geront/gnw102] [Medline: 27521577]
18. Mo GY, Biss RK, Poole L, Stern B, Waite K, Murphy KJ. Technology use among family caregivers of people with dementia. Can J Aging 2021 Jun;40(2):331-343. [doi: 10.1017/S0714980820000094] [Medline: 32515318]

XSL•FO
RenderX

19. Ploeg J, Markle-Reid M, Valaitis R, McAiney C, Duggleby W, Bartholomew A, et al. Web-based interventions to improve mental health, general caregiving outcomes, and general health for informal caregivers of adults with chronic conditions living in the community: rapid evidence review. J Med Internet Res 2017 Jul 28;19(7):e263 [FREE Full text] [doi: 10.2196/jmir.7564] [Medline: 28754652]
20. Sherifali D, Ali MU, Ploeg J, Markle-Reid M, Valaitis R, Bartholomew A, et al. Impact of internet-based interventions on caregiver mental health: systematic review and meta-analysis. J Med Internet Res 2018 Jul 03;20(7):e10668 [FREE Full text] [doi: 10.2196/10668] [Medline: 29970358]
21. Bashshur R, Shannon G, Krupinski E, Grigsby J. The taxonomy of telemedicine. Telemed J E Health 2011;17(6):484-494. [doi: 10.1089/tmj.2011.0103] [Medline: 21718114]
22. mHealth: new horizons for health through mobile technologies. World Health Organization. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2021-08-15]
23. Zulman DM, Piette JD, Jenchura EC, Asch SM, Rosland A. Facilitating out-of-home caregiving through health information technology: survey of informal caregivers' current practices, interests, and perceived barriers. J Med Internet Res 2013 Jul 10;15(7):e123 [FREE Full text] [doi: 10.2196/jmir.2472] [Medline: 23841987]
24. Grossman MR, Zak DK, Zelinski EM. Mobile apps for caregivers of older adults: quantitative content analysis. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e162 [FREE Full text] [doi: 10.2196/mhealth.9345] [Medline: 30061093]
25. Young HM, Bell JF, Whitney RL, Ridberg RA, Reed SC, Vitaliano PP. Social determinants of health: underreported heterogeneity in systematic reviews of caregiver interventions. Gerontologist 2020 Feb 14;60(Suppl 1):S14-S28 [FREE Full text] [doi: 10.1093/geront/gnz148] [Medline: 32057083]
26. Sangrar R, Docherty-Skippen SM, Beattie K. Blended face-to-face and online/computer-based education approaches in chronic disease self-management: a critical interpretive synthesis. Patient Educ Couns 2019 Oct;102(10):1822-1832. [doi: 10.1016/j.pec.2019.05.009] [Medline: 31101427]
27. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. J Clin Epidemiol 2014 Jan;67(1):56-64. [doi: 10.1016/j.jclinepi.2013.08.005] [Medline: 24189091]
28. 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. Int J Surg 2021 Apr;88:105906. [doi: 10.1016/j.ijsu.2021.105906] [Medline: 33789826]
29. Garnett A, Northwood M, Sangrar R, Ting J. Interventions to support caregivers of older adults: an equity-focused systematic review. PROSPERO. URL: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021239584 [accessed 2021-08-15]
30. Dugas M, Gao GG, Agarwal R. Unpacking mHealth interventions: a systematic review of behavior change techniques used in randomized controlled trials assessing mHealth effectiveness. Digit Health 2020;6:2055207620905411 [FREE Full text] [doi: 10.1177/2055207620905411] [Medline: 32128233]
31. Mobile fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2021-08-15]
32. Petrovic M, Gaggioli A. Digital mental health tools for caregivers of older adults-a scoping review. Front Public Health 2020;8:128 [FREE Full text] [doi: 10.3389/fpubh.2020.00128] [Medline: 32411643]
33. Rathnayake S, Moyle W, Jones C, Calleja P. mHealth applications as an educational and supportive resource for family carers of people with dementia: an integrative review. Dementia (London) 2019;18(7-8):3091-3112. [doi: 10.1177/1471301218768903] [Medline: 29631492]
34. Wasilewski MB, Stinson JN, Cameron JI. Web-based health interventions for family caregivers of elderly individuals: a scoping review. Int J Med Inform 2017 Jul;103:109-138. [doi: 10.1016/j.ijmedinf.2017.04.009] [Medline: 28550996]
35. Health Promotion, Inequalities and Young People's Health A Systematic Review of Research. London: University of London; 2008.
36. Critical appraisal tools. JBI. URL: https://jbi.global/critical-appraisal-tools [accessed 2021-06-06]
37. Beentjes KM, Neal DP, Kerkhof YJF, Broeder C, Moeridjan ZD, Ettema TP, et al. Impact of the FindMyApps program on people with mild cognitive impairment or dementia and their caregivers; an exploratory pilot randomised controlled trial. Disabil Rehabil Assist Technol 2020 Nov 27:1-13. [doi: 10.1080/17483107.2020.1842918] [Medline: 33245000]
38. Callan JA, Siegle GJ, Abebe K, Black B, Martire L, Schulz R, et al. Feasibility of a pocket-PC based cognitive control intervention in dementia spousal caregivers. Aging Ment Health 2016 May 06;20(6):575-582 [FREE Full text] [doi: 10.1080/13607863.2015.1031635] [Medline: 25945597]
39. Lai R, Tensil M, Kurz A, Lautenschlager NT, Diehl-Schmid J. Perceived need and acceptability of an app to support activities of daily living in people with cognitive impairment and their carers: pilot survey study. JMIR Mhealth Uhealth 2020 Jul 31;8(7):e16928 [FREE Full text] [doi: 10.2196/16928] [Medline: 32735223]
40. Salin S, Laaksonen H. Bringing older informal caregivers and recipients to the digital age: experiences of using digital services. Technol Disability 2019 Oct 08;31(3):115-127. [doi: 10.3233/tad-190230]
41. Sourbeer K, Bankole A, Anderson M, Newbold T, Nama R, Belay M, et al. Assessing BESI mobile application usability for caregivers of persons with dementia. Gerontechnology 2018 Aug 10;17(2):102-112. [doi: 10.4017/gt.2018.17.2.004.00]

XSL•FO
RenderX

42. Watcharasarnsap T, Tanlamai U, Lipikorn R. Innovative communication mechanism for the psychological well-being of caregivers and elderly with dementia in the context of Thailand. Int J Innov Creativity Change 2020;13(6):458-474 [FREE Full text]
43. Davis B, Nies M, Shehab M, Shenk D. Developing a pilot e-mobile app for dementia caregiver support: lessons learned. Online J Nurs Informatics 2014;18(1) [FREE Full text]
44. Lai FH, Yan EW, Yu KK, Tsui W, Chan DT, Yee BK. The protective impact of telemedicine on persons with dementia and their caregivers during the COVID-19 pandemic. Am J Geriatr Psychiatry 2020 Nov;28(11):1175-1184 [FREE Full text] [doi: 10.1016/j.jagp.2020.07.019] [Medline: 32873496]
45. Hastings SN, Mahanna EP, Berkowitz TS, Smith VA, Choate AL, Hughes JM, et al. Video-enhanced care management for medically complex older adults with cognitive impairment. J Am Geriatr Soc 2021 Jan;69(1):77-84 [FREE Full text] [doi: 10.1111/jgs.16819] [Medline: 32966603]
46. Kales HC, Gitlin LN, Stanislawski B, Myra Kim H, Marx K, Turnwald M, et al. Effect of the WeCareAdvisor™ on family caregiver outcomes in dementia: a pilot randomized controlled trial. BMC Geriatr 2018 May 10;18(1):113 [FREE Full text] [doi: 10.1186/s12877-018-0801-8] [Medline: 29747583]
47. Ptomey LT, Vidoni ED, Montenegro-Montenegro E, Thompson MA, Sherman JR, Gorczyca AM, et al. The feasibility of remotely delivered exercise session in adults with Alzheimer's disease and their caregivers. J Aging Phys Act 2019 Sep 01;27(5):670-677 [FREE Full text] [doi: 10.1123/japa.2018-0298] [Medline: 30747564]
48. Quinn CC, Staub S, Barr E, Gruber-Baldini A. Mobile support for older adults and their caregivers: dyad usability study. JMIR Aging 2019 May 23;2(1):e12276 [FREE Full text] [doi: 10.2196/12276] [Medline: 31518271]
49. Garvelink MM, Agbadjé TT, Freitas A, Bergeron L, Petitjean T, Dugas M, et al. Improving a web-based tool to support older adults to stay independent at home: qualitative study. JMIR Mhealth Uhealth 2020 Jul 22;8(7):e16979 [FREE Full text] [doi: 10.2196/16979] [Medline: 32412908]
50. Hughes JC, Banerjee T, Goodman G, Lawhorne L. A preliminary qualitative analysis on the feasibility of using gaming technology in caregiver assessment. J Technol Human Serv 2017 Jul 25;35(3):183-198. [doi: 10.1080/15228835.2017.1347553]
51. Killin LO, Russ TC, Surdhar SK, Yoon Y, McKinstry B, Gibson G, et al. Digital Support Platform: a qualitative research study investigating the feasibility of an internet-based, postdiagnostic support platform for families living with dementia. BMJ Open 2018 Apr 12;8(4):e020281 [FREE Full text] [doi: 10.1136/bmjopen-2017-020281] [Medline: 29654028]
52. Rathnayake S, Jones C, Calleja P, Moyle W. Family carers' perspectives of managing activities of daily living and use of mHealth applications in dementia care: a qualitative study. J Clin Nurs 2019 Dec;28(23-24):4460-4470. [doi: 10.1111/jocn.15030] [Medline: 31408554]
53. Ruggiano N, Brown EL, Shaw S, Geldmacher D, Clarke P, Hristidis V, et al. The potential of information technology to navigate caregiving systems: perspectives from dementia caregivers. J Gerontol Soc Work 2019;62(4):432-450. [doi: 10.1080/01634372.2018.1546786] [Medline: 30422754]
54. Ryan AA, McCauley CO, Laird EA, Gibson A, Mulvenna MD, Bond R, et al. 'There is still so much inside': the impact of personalised reminiscence, facilitated by a tablet device, on people living with mild to moderate dementia and their family carers. Dementia (London) 2020 May;19(4):1131-1150. [doi: 10.1177/1471301218795242] [Medline: 30193535]
55. Span M, Smits C, Jukema J, Groen-van de Ven L, Janssen R, Vernooij-Dassen M, et al. An interactive web tool for facilitating shared decision-making in dementia-care networks: a field study. Front Aging Neurosci 2015;7:128 [FREE Full text] [doi: 10.3389/fnagi.2015.00128] [Medline: 26217221]
56. Banbury A, Parkinson L, Gordon S, Wood D. Implementing a peer-support programme by group videoconferencing for isolated carers of people with dementia. J Telemed Telecare 2019 Oct;25(9):572-577. [doi: 10.1177/1357633X19873793] [Medline: 31631761]
57. Breebaart H, Broese van Groenou M. A groupware tool to facilitate caregiving for home-dwelling frail older persons in the Netherlands: mixed-methods study. JMIR Aging 2018 Dec 07;1(2):e10697 [FREE Full text] [doi: 10.2196/10697] [Medline: 31518286]
58. Brown EL, Ruggiano N, Page TF, Roberts L, Hristidis V, Whiteman KL, et al. CareHeroes web and android™ apps for dementia caregivers: a feasibility study. Res Gerontol Nurs 2016;9(4):193-203 [FREE Full text] [doi: 10.3928/19404921-20160229-02] [Medline: 29977440]
59. Dam AE, Christie HL, Smeets CM, van Boxtel MP, Verhey FR, de Vugt ME. Process evaluation of a social support platform 'Inlife' for caregivers of people with dementia. Internet Interv 2019 Mar;15:18-27 [FREE Full text] [doi: 10.1016/j.invent.2018.09.002] [Medline: 30510911]
60. Sikder AT, Yang FC, Schafer R, Dowling GA, Traeger L, Jain FA. Mentalizing imagery therapy mobile app to enhance the mood of family dementia caregivers: feasibility and limited efficacy testing. JMIR Aging 2019 Mar 21;2(1):e12850 [FREE Full text] [doi: 10.2196/12850] [Medline: 31518275]
61. Costa Stutzel M, Filippo MP, Sztajnberg A, da Costa RM, Brites AD, da Motta LB, et al. Multi-part quality evaluation of a customized mobile application for monitoring elderly patients with functional loss and helping caregivers. BMC Med Inform Decis Mak 2019 Jul 22;19(1):140 [FREE Full text] [doi: 10.1186/s12911-019-0839-3] [Medline: 31331309]

XSL•FO
RenderX

62. Tyack C, Camic PM, Heron MJ, Hulbert S. Viewing art on a tablet computer: a well-being intervention for people with dementia and their caregivers. J Appl Gerontol 2017 Jul;36(7):864-894. [doi: 10.1177/0733464815617287] [Medline: 26675353]
63. Davis NJ, Clark PC, Johnson TM, Wyman JF. Feasibility of Tele-Prompt: a tablet-based prompted voiding intervention to support informal caregivers of older adults with urinary incontinence. Geriatr Nurs 2020;41(4):411-420. [doi: 10.1016/j.gerinurse.2020.01.002] [Medline: 31987698]
64. Park E, Park H, Kim EK. The effect of a comprehensive mobile application program (CMAP) for family caregivers of home-dwelling patients with dementia: a preliminary research. Jpn J Nurs Sci 2020 Oct;17(4):e12343. [doi: 10.1111/jjns.12343] [Medline: 32363664]
65. Bédard M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001 Oct;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
66. Kaye LW. Assessing the efficacy of a self-help support group program for older women. J Women Aging 1996 Feb 07;7(4):11-30. [doi: 10.1300/j074v07n04_03]
67. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Human Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
68. Anderson JG, Flatt JD, Jabson Tree JM, Gross AL, Rose KM. Characteristics of sexual and gender minority caregivers of people with dementia. J Aging Health 2021 Dec;33(10):838-851 [FREE Full text] [doi: 10.1177/08982643211014767] [Medline: 33998313]
69. Boehmer U, Clark MA, Heeren TC, Showalter EA, Fredman L. Differences in caregiving outcomes and experiences by sexual orientation and gender identity. LGBT Health 2018;5(2):112-120. [doi: 10.1089/lgbt.2017.0144] [Medline: 29364755]
70. Schulz R, Eden J. Family caregiving roles and impacts. In: Families Caring for an Aging America. Washington, DC: The National Academies Press; 2016.
71. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. J Am Med Inform Assoc 2018 Aug 01;25(8):1080-1088. [doi: 10.1093/jamia/ocy052] [Medline: 29788380]
72. Chesser AK, Keene Woods N, Smothers K, Rogers N. Health literacy and older adults: a systematic review. Gerontol Geriatr Med 2016;2:2333721416630492 [FREE Full text] [doi: 10.1177/2333721416630492] [Medline: 28138488]
73. Lindeman DA, Kim KK, Gladstone C, Apesoa-Varano EC. Technology and caregiving: emerging interventions and directions for research. Gerontologist 2020 Feb 14;60(Suppl 1):S41-S49 [FREE Full text] [doi: 10.1093/geront/gnz178] [Medline: 32057082]
74. Caregivers & technology: what they want and need. AARP. URL: https://www.aarp.org/content/dam/aarp/research/surveys_statistics/ltc/2018/caregivers-technology-needs.doi.10.26419-2Fres.00191.002.pdf [accessed 2021-08-16]
75. Cosco TD, Firth J, Vahia I, Sixsmith A, Torous J. Mobilizing mHealth data collection in older adults: challenges and opportunities. JMIR Aging 2019 Mar 19;2(1):e10019 [FREE Full text] [doi: 10.2196/10019] [Medline: 31518253]
76. Sharma N, Chakrabarti S, Grover S. Gender differences in caregiving among family - caregivers of people with mental illnesses. World J Psychiatry 2016 Mar 22;6(1):7-17 [FREE Full text] [doi: 10.5498/wjp.v6.i1.7] [Medline: 27014594]
77. Dilworth-Anderson P, Brummett BH, Goodwin P, Williams SW, Williams RB, Siegler IC. Effect of race on cultural justifications for caregiving. J Gerontol B Psychol Sci Soc Sci 2005 Sep;60(5):S257-S262. [doi: 10.1093/geronb/60.5.s257] [Medline: 16131626]
78. Williams A, Sethi B, Duggleby W, Ploeg J, Markle-Reid M, Peacock S, et al. A Canadian qualitative study exploring the diversity of the experience of family caregivers of older adults with multiple chronic conditions using a social location perspective. Int J Equity Health 2016 Mar 02;15:40 [FREE Full text] [doi: 10.1186/s12939-016-0328-6] [Medline: 26936057]
79. Sidani S, Guruge S, Miranda J, Ford-Gilboe M, Varcoe C. Cultural adaptation and translation of measures: an integrated method. Res Nurs Health 2010 Apr;33(2):133-143. [doi: 10.1002/nur.20364] [Medline: 20069612]
80. Cristiano A, Musteata S, De Silvestri S, Bellandi V, Ceravolo P, Cesari M, et al. Older adults' and clinicians' perspectives on a smart health platform for the aging population: design and evaluation study. JMIR Aging 2022 Feb 28;5(1):e29623 [FREE Full text] [doi: 10.2196/29623] [Medline: 35225818]
81. Portz JD, Ford KL, Elsbernd K, Knoepke CE, Flint K, Bekelman DB, et al. "I like the idea of it…but probably wouldn't use it" - health care provider perspectives on heart failure mHealth: qualitative study. JMIR Cardio 2020 Sep 04;4(1):e18101 [FREE Full text] [doi: 10.2196/18101] [Medline: 32885785]
82. Jaana M, Paré G. Comparison of mobile health technology use for self-tracking between older adults and the general adult population in Canada: cross-sectional survey. JMIR Mhealth Uhealth 2020 Nov 27;8(11):e24718 [FREE Full text] [doi: 10.2196/24718] [Medline: 33104517]
83. Brewer LC, Fortuna KL, Jones C, Walker R, Hayes SN, Patten CA, et al. Back to the future: achieving health equity through health informatics and digital health. JMIR Mhealth Uhealth 2020 Jan 14;8(1):e14512 [FREE Full text] [doi: 10.2196/14512] [Medline: 31934874]
84. Ibrahim S, Sidani S. Strategies to recruit minority persons: a systematic review. J Immigr Minor Health 2014 Oct;16(5):882-888. [doi: 10.1007/s10903-013-9783-y] [Medline: 23338906]

XSL•FO
RenderX

85. Welch V, Petticrew M, Ueffing E, Benkhalti Jandu M, Brand K, Dhaliwal B, et al. Does consideration and assessment of effects on health equity affect the conclusions of systematic reviews? A methodology study. PLoS One 2012;7(3):e31360 [FREE Full text] [doi: 10.1371/journal.pone.0031360] [Medline: 22427804]
86. Furler J, Magin P, Pirotta M, van Driel M. Participant demographics reported in "Table 1" of randomised controlled trials: a case of "inverse evidence"? Int J Equity Health 2012 Mar 19;11:14 [FREE Full text] [doi: 10.1186/1475-9276-11-14] [Medline: 22429574]
87. Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, PRISMA-Equity Bellagio group. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. PLoS Med 2012;9(10):e1001333 [FREE Full text] [doi: 10.1371/journal.pmed.1001333] [Medline: 23222917]

Abbreviations

mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROGRESS-Plus: place of residence, race, occupation, gender, religion, education, social capital, socioeconomic status–plus age, disability, and sexual orientation
PROSPERO: International Prospective Register of Systematic Reviews

Edited by J Wang; submitted 24.08.21; peer-reviewed by A Sikder, E Brown, J Wolff; comments to author 02.12.21; revised version received 11.04.22; accepted 23.05.22; published 08.07.22.

Please cite as:
Garnett A, Northwood M, Ting J, Sangrar R
mHealth Interventions to Support Caregivers of Older Adults: Equity-Focused Systematic Review
JMIR Aging 2022;5(3):e33085
URL: https://aging.jmir.org/2022/3/e33085
doi:10.2196/33085
PMID:35616514

©Anna Garnett, Melissa Northwood, Justine Ting, Ruheena Sangrar. Originally published in JMIR Aging (https://aging.jmir.org), 08.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Online Peer Support for People With Parkinson Disease: Narrative Synthesis Systematic Review

Esther Vera Gerritzen[1*], BSc, MSc; Abigail Rebecca Lee[1*], BSc, MSc; Orii McDermott[1*], ARCM, BA, MMT, PhD; Neil Coulson[2*], MA, MSc, PhD; Martin Orrell[1*], BMedSci, MBBS, PhD

[1]Institute of Mental Health, Mental Health and Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, United Kingdom
[2]Population and Lifespan Health, School of Medicine, University of Nottingham, Nottingham, United Kingdom
[*]all authors contributed equally

Corresponding Author:
Esther Vera Gerritzen, BSc, MSc
Institute of Mental Health
Mental Health and Clinical Neuroscience, School of Medicine
University of Nottingham
University of Nottingham Innovation Park
Triumph Road
Nottingham, NG7 2TU
United Kingdom
Phone: 44 115 748 4262
Email: Esther.Gerritzen@nottingham.ac.uk

Abstract

Background: Parkinson disease (PD) significantly impacts the lives of people with the diagnosis and their families. In addition to the physical symptoms, living with PD also has an emotional impact. This can result in withdrawal from social roles, increasing the risk for social isolation and loneliness. Peer support is a way to stay socially connected, share experiences, and learn new coping skills. Peer support can be provided both in person and on the internet. Some of the advantages of online peer support are that it overcomes geographical barriers and provides a form of anonymity; moreover, support can be readily available when needed. However, the psychosocial impact of PD is still underresearched, and there is no systematic synthesis of online peer support for people with PD.

Objective: This review aims to explore the benefits and challenges of online peer support and identify successful elements of online peer support for people with PD.

Methods: The method selected for this systematic review is narrative synthesis. A total of 6 databases were systematically searched in April 2020 for articles published between 1989 and 2020. The quality of the included studies was assessed using the Critical Appraisal Skills Programme qualitative research checklist and the Downs and Black quality checklist.

Results: A total of 10,987 unique articles were identified through a systematic database search. Of these 10,987 articles, 8 (0.07%) were included in this review. Of the 8 studies, 5 (63%) were of good or high quality, 2 (25%) were of medium or fair quality, and 1 (13%) study was of poor quality. Web-based platforms included discussion forums, a web-based virtual world, and Facebook groups. Most papers reported on text-based communication. The included studies reported on sharing social support and personal experiences. Successful elements included increasing similarity between members and offering the opportunity to directly ask questions to a physician. Challenges included members leaving without a warning and PD symptoms hindering the use of technology.

Conclusions: Peer support can improve social support and help people with PD in living meaningful and satisfying lives. Peer support is unique and cannot be replaced by family members, friends, or health care professionals. Online peer support can be a solution for those who do not have access to an in-person support group or whose PD symptoms restrict them from travelling. However, research on the personal experiences of those who engage in online peer support and potential barriers in accessing it remains limited. Future research could use qualitative methods to explore these fields further.

(JMIR Aging 2022;5(3):e35425) doi:10.2196/35425

XSL•FO
RenderX

KEYWORDS

Parkinson disease; web-based health community; online peer support; narrative synthesis; systematic review

Introduction

Background

Parkinson disease (PD) is a chronic and progressive, neurodegenerative condition which is characterized by motor symptoms such as tremor, bradykinesia, and rigidity. In addition to the motor symptoms, many patients experience nonmotor symptoms, such as sleep disturbances, depression, and constipation [1]. It was estimated that in 2016, 6.1 million people had PD globally [2]. People with PD typically start developing symptoms in their 60s; however, it can also occur at a younger age [3]. In 2018, in the United Kingdom, >145,000 people were living with PD, of whom 19,690 were younger than 65 years [4].

PD significantly impacts the lives of people with the diagnosis and their families [5,6]. Owing to its chronic and neurodegenerative nature, people with PD need ongoing care and support [6]. In addition to the physical symptoms associated with PD, receiving the diagnosis and living with the condition also has an emotional impact. This includes anxiety for the future, difficulties in managing the condition in daily life, and the impact on the family [7]. PD can affect people's social lives and how they are involved in different roles, such as their role within the family, social circles, or at work. Receiving a diagnosis of PD and living with the condition can result in withdrawal from such social roles, increasing the risk of social isolation and loneliness [8].

The psychosocial impact of PD can be discussed within the social health framework [9,10]. In this framework, health is viewed in the social domain and includes three dimensions: (1) being able to fulfill potential and obligations, (2) managing life with some level of independence despite living with a health condition, and (3) being able to participate in social activities and work. When focusing on coping strategies and finding a balance between limitations and one's abilities, people can successfully adapt to living with a chronic condition and still live meaningful and satisfying lives [10]. Dröes et al [9] discussed how the concept of social health relates to people living with dementia, suggesting that it is possible for people with dementia to still participate in the 3 dimensions of social health and perceive a good quality of life. Within the PD context, Vescovelli et al [11] touch upon the social health framework by emphasizing the importance of social support for the well-being of people with PD. Social support is a term used to describe receiving care and help from others. It is often linked to social connectedness and being part of a social network [12]. In their systematic reviews, Vescovelli et al [11] found that for people with PD, social support is associated with greater social inclusion and opportunities to remain involved with work, supporting people to keep living meaningful and satisfying lives despite their PD. Thus, social support could improve the social health of people with PD [11]. However, despite these findings, Hellqvist et al [8] and Vescovelli et al [11] conclude that the psychosocial impact of PD is still underresearched.

One way in which people can stay socially connected and thus improve their social health is through peer support [13]. Peer support can be defined as the exchange of support between those (also referred to as *peers*) who share a similar health condition or life experience [14,15]. Peers can provide one another with social support; more specifically, there is reciprocity of support, meaning that people can develop a relationship in which they can both receive and provide support. This can increase feelings of empowerment [16,17]. Furthermore, peers can share experiential knowledge, which includes information and perspectives that people have because of their personal experiences of living with a certain condition [17]. These elements are unique to peer support and cannot be provided by health care professionals or others who are not living with PD [14,15].

Peer support can be provided in different ways, including web-based settings. The internet is an important source of health-related information and provides a platform for the creation and spread of web-based patient communities [16]. Since the 1990s, the number of web-based patient communities for a variety of health conditions has been increasing [18,19]. Such communities can function as self-help groups in which members share experiences and emotions and provide mutual support and empathy [16,20,21]. Some of the advantages of online peer support compared with in-person support groups include that it overcomes geographical barriers [19,22]; provides a form of anonymity, which can be particularly suitable for people with stigmatized conditions [22,23]; and online peer support can be readily available at any time of the day when needed [19,22]. Research has been conducted on online peer support communities for a variety of health conditions, including chronic conditions such as multiple sclerosis (MS) [24,25], HIV or AIDS [26,27], and amyotrophic lateral sclerosis (ALS) [28,29]. The review by Kingod et al [13] shows that online peer support communities can offer people with chronic conditions emotional, social, and practical support in managing their condition in their daily lives. Chronic conditions covered by this review include type 1 diabetes, HIV or AIDS, and chronic pain [13].

Web-based health communities and peer support in web-based settings is a rapidly growing field [16,18,19]. Especially during the COVID-19 pandemic and national lockdowns, connecting with others on the web has become increasingly important. However, knowledge of the long-term effects of online peer support, how it impacts users' health and self-management, and what particular elements make it useful and meaningful need further research [19,30]. Research into online peer support for people affected by PD is also growing [31,32]; however, to the best of our knowledge, there is no systematic synthesis of this research yet.

Objectives

This narrative synthesis systematic review aimed to (1) explore the benefits and challenges of online peer support and (2) identify successful elements of online peer support for people

XSL•FO
RenderX

with PD. In this review, the challenges cover things that make it more difficult for a person with PD to use online peer support. This can include aspects related to technology as well as PD-related challenges. Understanding the successful elements can be helpful in improving existing and developing new online peer support opportunities for people with PD as well as other conditions. Elements of online peer support were deemed successful if studies identified positive outcomes for the people with PD engaging in online peer support.

Methods

Narrative Synthesis

The method that was selected for this systematic review was narrative synthesis, using the procedures outlined by Popay et al [33]. This entails including the following elements: (1) theory development, (2) development of a preliminary synthesis, (3) exploration of relationships in the data, and (4) assessment of robustness of the synthesis. With a narrative synthesis, the presentation of the findings is mainly words- and text-based, and it is a useful method to identify elements of best practice [33]. Furthermore, this review followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 guidelines [34]. More details on the narrative synthesis methods can be found in Multimedia Appendix 1 [9,10,14,16,33,35].

Search Strategy

A systematic database search was conducted in April 2020. The search strategy was developed with the help of 2 librarians and NC, who is an academic expert on online peer support. The initial search was part of a wider appraisal of the literature and included PD, MS, ALS, and Huntington disease. This paper will only present the results for patients with PD. A total of 6 databases were searched: CINAHL, Cochrane Library, EMBASE MEDLINE, PsycINFO, Scopus, and Web of Science. The keywords used for the searches are presented in Textbox 1. A search filter for the year of publication, 1989 to 2020, was applied. This was because the World Wide Web was introduced in 1989. No filters on the study design were applied. Finally, the reference lists of the included papers were searched manually. This did not result in any new papers being added.

Textbox 1. Search terms.

Search term 1

- parkinson* disease
- parkinson*

Search term 2

- online
- digital
- web-based
- app-based
- internet
- social media
- peer
- peer support
- support group
- social support
- online support group
- online support commun*
- discussion forum*
- bulletin board
- chat room*
- computer-mediated support
- internet support group*
- internet support commun*
- online self-help
- web-based support group*
- web-based support commun*

XSL•FO
RenderX

Textbox 2 lists the inclusion and criteria followed while selecting papers for this review.

Textbox 2. Inclusion and exclusion criteria.

Inclusion criteria

- The study population included people living with Parkinson disease or a blend of people living with Parkinson disease and caregivers.
- The intervention included online peer support. For this review, online peer support was regarded as communication via the internet between peers in a web- or app-based environment that is designed to facilitate social contact using either an asynchronous or synchronous text- or text and video-based platform (eg, social media platforms, forums, or chat rooms).
- Publication between 1989 and 2020.
- Publication in peer-reviewed journals.

Exclusion criteria

- The study focused solely on caregiver perspectives.
- The intervention included online peer support that was part of a program that also included in-person or telephone-based peer support.
- The study did not report on peer-to-peer interactions. This exclusion criterion was added after initial screening. See the *Study Selection* section for more details.
- Literature reviews, opinion pieces, protocols, editorials, or conference abstracts.
- Papers written in a language other than English if a translation was not available.

Study Selection

The search results were imported into EndNote (Clarivate), after which all duplicates were removed. The primary reviewer (EVG) reviewed each title and abstract against the eligibility criteria. The primary reviewer consulted a second reviewer (ARL) on the titles and abstracts that she was unsure about. The title and abstract screening was followed by a full-text analysis of the potentially relevant papers. The initial full-text analysis was conducted by the primary reviewer. The same procedures as used for the title and abstract screening were followed. At this stage, the main reason for labeling a paper as unsure was that although the paper met the inclusion and exclusion criteria, it mainly focused on other outcomes (eg, quality of life) rather than peer-to-peer interactions. Following a discussion with a third reviewer (OM), it was decided to refine the exclusion criteria and add the criterion that papers could be excluded if they did not report on peer-to-peer interactions. The papers that were included up to that point were reassessed against the newly added exclusion criterion.

Data Extraction

Following the study selection, the primary reviewer (EVG) extracted the data using standardized data extraction forms. Data were extracted on (1) study information, (2) study characteristics, (3) population characteristics, (4) characteristics of the web- or app-based platform, (5) outcomes, and (6) implications for future research. ARL provided a second independent review of the completed data extraction forms.

Quality Assessment

In all, 2 quality assessment tools were used to assess the risk of bias in individual studies. EVG completed the initial quality assessment and ARL provided a second independent review. For the assessment of the risk of bias in qualitative studies, the Critical Appraisal Skills Programme (CASP) qualitative research checklist was used [36]. This checklist consists of 10 questions related to “rigour, credibility and relevance” [37]. For studies that could not be assessed using the CASP checklist, the Downs and Black quality checklist was used. This tool consists of 27 items and is suitable for both randomized and nonrandomized studies [38]. Both the CASP checklist and the Downs and Black quality checklist were recommended by the Centre for Reviews and Dissemination guidance for undertaking reviews in health care [37] and have been successfully used in previous systematic reviews [39,40].

For the CASP checklist, studies will be graded *high*, if they met or partially met 8 to 10 items; *medium*, if they met or partially met 5 to 7 items; and *low*, if they met or partially met 0 to 4 items [41]. For the Downs and Black quality checklist, papers are labeled *excellent*, if they have 24 to 28 points; *good*, if they have 19 to 23 points; *fair*, with 14 to 18 points; and *poor*, when they have less than 14 points [42].

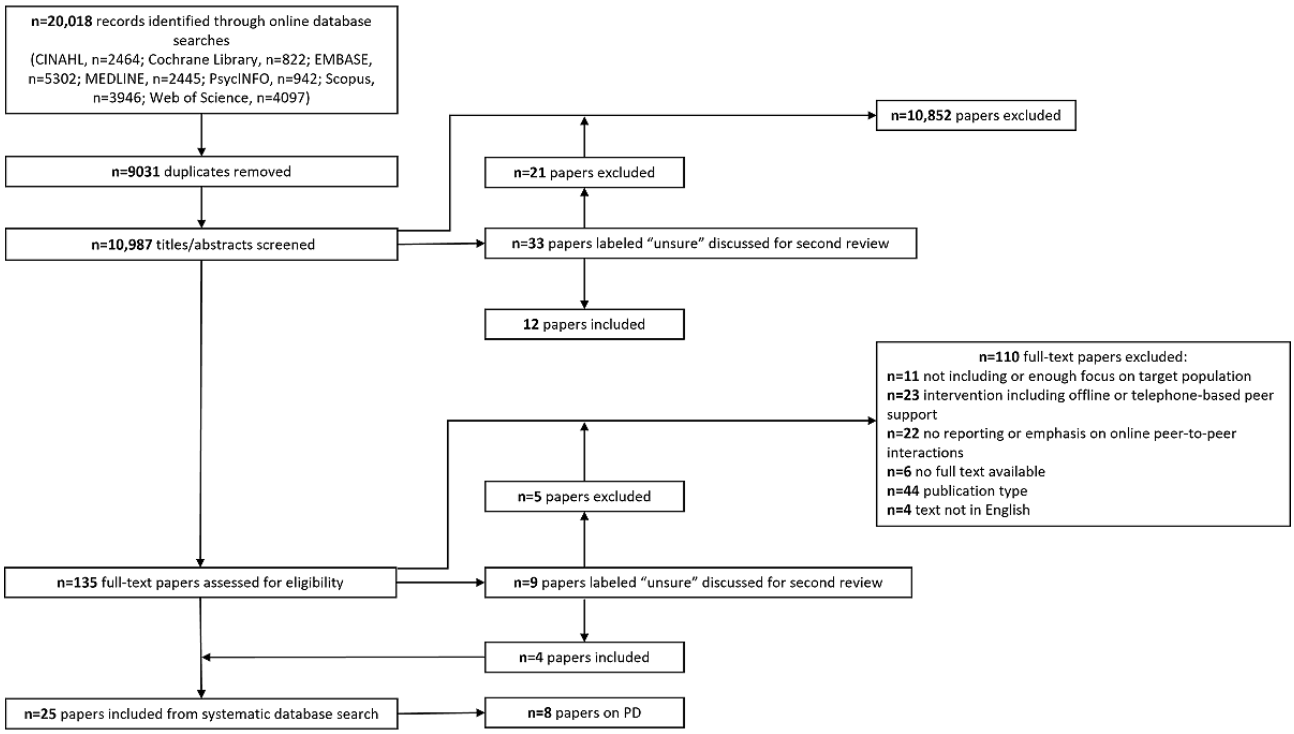
Results

Overview

The results section covers element 2 of a narrative synthesis: developing a preliminary synthesis. A web-based database search returned 10,987 unique titles and abstracts. After screening of the titles, abstracts, and full texts, of the 10,987 papers, 8 (0.07%) met the inclusion criteria for this review. An overview of the web-based database search and screening process can be found in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram of the search and review process. PD: Parkinson disease.



Study Characteristics

An overview of the study characteristics is presented in Table 1. This review includes a variety of methods. Of the 8 papers, 3 (38%) papers used a qualitative content analysis of posts on a discussion forum [31,32,43], 3 (38%) papers reported the findings of a pilot study [44-46], 1 (13%) paper conducted an ethnographic study in a virtual world [47], and 1 (13%) paper conducted a survey and interviews [48].

XSL•FO
RenderX

Table 1. Study characteristics.

Study (year)	Aim or aims	Design (methods)	Intervention	Setting (country)	Study population	Eligibility criteria	Sample	QA[a] score
Attard and Coulson [31] (2012)	Experiences of PD[b] forum users	Qualitative content analysis of posts on 4 discussion forums	On the internet, public, asynchronous discussion forum	Study conducted in the United Kingdom; data collected from the United States, Canada, and Australia	People living with PD	PD online support groups with a discussion forum	A total of 4 web-based communities; 1000 to 10,000 members per group; approximately 100 active members per group; age unknown (only what members decided to share); more women than men; 1013 messages (approximately 250 per group)	9 (high)
Bakke et al [32] (2018)	Interaction between professional and personal expertise in web-based PD community	Qualitative content analysis of posts on discussion forum	On the internet, public, asynchronous discussion forum	Unknown	People living with PD and carers	Physician-moderated forum for PD	In all, 1 web-based community: 107 threads, 409 individual comments; age and gender unknown (only what members decided to share)	8 (high)
Stewart Loane et al [43] (2014)	Social support and consumer value in web-based health communities	Qualitative content analysis of posts on discussion forum	On the internet, asynchronous discussion forum	Unknown	People with PD	Not reported	PD community: 35 members, 30 threads, 137 posts; age and gender not reported	8 (high)
Davis and Boellstorff [47] (2016)	Creativity of people with PD in a virtual world	Qualitative ethnographic web-based study in a virtual world	Second Life, a virtual world	Study conducted in the United States (based on ethics approval)	People living with PD	Members of a PD community in a virtual world (recruited through prior fieldwork in 2004)	A total of 2 people living with PD (1 male and 1 female); female patient with young onset PD, male patient with unknown onset	7 (medium)
Lieberman et al [46] (2005)	Impact of group composition and utility of computer-based text analysis in developing web-based groups	Pre-post measurement study comparing homogeneous and heterogeneous groups	A total of 6 web-based PD support groups delivered by professionals; weekly meetings for 20 weeks; 3 homogeneous groups (2 young onset, aged <60 years; 1 newly diagnosed in the last 2 years); 3 heterogeneous groups (mix of age and time since diagnosis)	Study conducted in the United States	People living with PD	People living with PD in California and attending web-based PD support groups, described in the study by Lieberman et al [46]	A total of 66 participants: 12 were unable to attend, 12 dropouts from homogenous groups, and 9 dropouts from heterogeneous groups; homogeneous groups: mean age 55.6 (SD 6.4) years, 77.8% female; heterogeneous groups: mean age 63.9 (SD 8.5) years, 46.2% female	16 (good)

XSL•FO
RenderX

Study (year)	Aim or aims	Design (methods)	Intervention	Setting (country)	Study population	Eligibility criteria	Sample	QA[a] score
Lieberman et al [45] (2006) (same population [46])	Willingness to participate in professionally led web-based groups; characteristics of participants; outcomes; group composition	Pilot study of effectiveness of professionally led web-based PD support groups	See above for Lieberman et al [46]	Study conducted in the United States	People living with PD	People living with PD in California	A total of 66 participants: 32 completed pre-post measurements; mean age 60.2 (SD 9.2) years, 68% male	16 (good)
Lieberman [44] (2007; same population [46])	Characteristics of people with PD in online support groups and impact of fear on dropout rates	Pilot study	See above for Lieberman et al [46]; weekly meetings, 90 min per meeting, 25 weeks; premature termination: attending <10 meetings	Study conducted in the United States	People living with PD	People living with PD	A total of 66 participants: 26 premature terminators and 40 continuers	15 (fair)
Martínez-Pérez et al [48] (2014)	Characteristics of Facebook groups and Twitter and their purposes and functions	Mixed methods survey and interviews with Facebook and Twitter users	Facebook and Twitter groups for PD	Unknown	People affected by PD	Facebook and Twitter focused on prevention, treatment, fund raising, cures, or general information	A total of 257 Facebook groups and 100 Twitter groups; no demographic information about group members was presented	4 (low)

[a]QA: quality assessment.

[b]PD: Parkinson disease.

Summary of Interventions

In all studies, the mode of communication between the participants was text-based. In 50% (4/8) of studies, communication was asynchronous [31,32,43,48], meaning that participants did not necessarily communicate with each other in real time. This is one of the characteristics of discussion forums, where people can post a message and others can respond at a time that is convenient for them. A total of 50% (4/8) of studies [44-47] used real-time communication (synchronous). Other than in a study [48], all online peer support communities analyzed in this review were moderated. This means that one or multiple people either guided the discussion or monitored posts. Although 75% (6/8) of studies only included people living with a PD diagnosis [31,43-47], 25% (2/8) of studies included both caregivers and people with a PD diagnosis [32,48].

Quality Assessment

Of the 8 papers, 5 (63%) were of good or high quality, 2 (25%) were labeled medium or fair quality, and 1 (13%) paper was labeled as poor quality. In total, 63% (5/8) of papers were assessed using the CASP checklist. Of these5 papers, 3 (60%) were labeled as high quality [31,32,43], 1 (20%) as medium [47], and 1 (20%) was assessed to be of low quality [48]. The 38% (3/8) of remaining papers were assessed using the Downs and Black quality checklist. Of these 3 papers, 2 (67%) were labeled as good [45,46] and 1 (33%) was labeled as fair [44]. An overview of the CASP checklist, Downs and Black quality checklist, and the scores for each study can be found in Multimedia Appendix 2 [31,32,43-48].

Key Findings

Overview

An overview of the web-based platform characteristics is presented in Table 2. An overview of the study outcomes is presented in Table 3.

XSL•FO
RenderX

Table 2. Web-based platform characteristics.

Study	Platform	Communication	Moderation
Attard and Coulson [31]	Discussion forums	Text-based (asynchronous)	Yes
Bakke et al [32]	WebMD (discussion forum)	Text-based (asynchronous)	Physician
Loane et al [43]	Discussion forum	Text-based (asynchronous)	Unknown
Davis and Boellstorff [47]	Virtual world	Verbal (synchronous)	Researchers
Lieberman et al [46]	Online support group in chat room	Text-based (synchronous)	Professional
Lieberman et al [45]	Online support group in chat room	Text-based (synchronous)	Professional
Lieberman [44]	Online support group in chat room	Text-based (synchronous)	Professional
Martínez-Pérez et al [48]	Facebook and Twitter	Text-based (asynchronous)	Unknown

XSL•FO
RenderX

Table 3. Study outcomes.

Study	Reported outcomes	Successful elements	Implications

Attard and Coulson [31]	• Positives:
	  • Social support, mutual understanding, and empathy
	  • Sharing experiences and advice
	  • Being part of a community, feeling less alone, and friendship
	  • Encouragement, positive thinking, and resilience
	• Negatives:
	  • Lack of replies
	  • Symptoms restricting ability to use computer
	  • Lack of personal information
	  • Absence of nonverbal communication
	  • Members leaving could be distressing for other members
		• Variety in experience, opinions, and advice
		• Tailored advice to individual members in simple, nonmedical language
		• Writing may help people to reflect on their situation and share things that are difficult to express face to face
		• Anonymous nature may help members to discuss taboo topics more openly
			• Explore the use of voice tools for people with PD[a] who have difficulties typing because of their symptoms
			• Ask users directly about experiences
			• Evaluate:
			  • accuracy of shared information
			  • impact of public nature of forum on members' experience and concerns about privacy
			  • impact of the presence of professional moderators

Bakke [32]	• Role of professional expertise:
	  • Trust in physician's opinion
	  • Acknowledging value of lived experience
	• Role of lay expertise:
	  • Value and trust peer's experiences. Mutual understanding and empathy
	  • Sharing personal experiences
	  • Reciprocity in answering questions and info sharing
	  • Referring to physician for advice
	  • Trust increased over time as members shared more
		• Having a physician moderator
		  • Opportunity to directly ask questions to physician
		  • Physician using understanding and supportive tone
		• Peer interaction, receiving advice from others going through something similar
		• Forum design: clearly labeling posts and profiles of physicians may play a role in building trust
			• For designing future forums:
			  • include badges and ratings to add validity to forum users' contributions
			  • clear norms and values pinned to home page
			• Moderation (professional or nonprofessional)

Stewart Loane et al [43]	• Information support most frequent, emotional support second.
	  • Initial posts often request information. Responses include answers and network and emotional support
	  • When sharing info, the posters receive positive feedback
	  • Spiritual support (expression of gratitude and feelings of connectedness)
	  • Ethics and morality: participants refusing to provide a diagnosis or medical advice
	  • Sharing poems and photos, humor, and banter. Sense of community
		• People with PD developed value through discussion without needing health care professionals to be present. This is helpful for health care professionals and managers.
		• Web-based discussion forums can remove barriers of information asymmetry and they create value and support for people with PD.
			• Using different methods to directly explore members' experiences
			• Further explore what features of a web-based community promote a sense of community among members
			• Explore a variety of web-based communities to identify whether specific features lead to greater value for members

Davis and Boellstorff [47]	• Users:
	  • discovered new ways of creativity
	  • continued creative parts of previous jobs which gave sense of purpose
	  • created art works in the platform to express what it feels like to have PD
	  • felt part of a community beyond PD
	  • learned new web-based skills
		• The Second Life platform was used for offline work purposes
		• Art works created in Second Life to express how it feels to have PD can be used for educational purposes
		• It can be difficult to find age-appropriate in-person support groups for younger people with PD. Web-based platforms are accessible to people from different areas
			• Explore the influence of factors such as gender, age, and young onset or late-onset PD on creativity
			• Explore to what extent creativity is experienced as a community or an individual phenomenon

XSL•FO
RenderX

Study	Reported outcomes	Successful elements	Implications

Lieberman et al [46]	• Quality of life of all groups improved
	• Homogeneous groups:
	  • were more committed to their group
	  • had higher levels of commitment and attraction, and positive feelings in initial 5 meetings
	  • had significantly greater positive changes compared with heterogeneous groups
		• Homogeneous groups based on age or time since diagnosis
		• The internet makes it easier to create homogeneous groups, with access to a larger group of patients
		• Lurking (reading posts but not creating own posts) can help with learning more about the group and finding similarities with other members
			• Explore:
			  • the impact of writing in online peer support groups
			  • the impact of the absence of visual and auditory cues
			• Internet support groups could target a more specific audience to enhance similarity between members
			• Option for subgroups

Lieberman et al [45]	• Members of web-based groups:
	  • had lower average age
	  • were living with diagnosis for fewer years
	  • had better scores for depression and QoL[b] before and after the intervention
	  • felt freer to talk about certain topics compared with in-person groups
	• Only homogeneous groups continued to stay in touch after intervention
	• Most participants heard about the online support groups through the internet, only a small percentage through their physician.
		• Homogeneous groups based on age or time since diagnosis	• Explore why people drop out of online support groups
			• Explore opportunities of using voice recognition software

Lieberman [44]	• Participants who dropped out:
	  • had higher levels of anxiety
	  • did not score differently on depression, quality of life, and intensity of PD symptoms measurements
		• Homogeneous groups showed significantly greater improvement compared with heterogeneous groups	• Explore what effective strategies are to prevent people from dropping out (eg, group structure, group composition, and preparation)

Martínez-Pérez et al [48]	• On Facebook, the majority was self-help groups • On Twitter, the goals of people were to share information and create awareness • There is a need for dedicated networking sites for peer support	• N/A[c]	• Directly explore the experiences of users

[a]PD: Parkinson disease.

[b]QoL: quality of life.

[c]N/A: not applicable.

Social Support

One of the main characteristics of online and in-person peer support is social support [14,16]. This finding also came forward in this review, and studies reported on different elements of social support. Through content analysis of discussion forums, studies [31,32,43] observed mutual understanding and empathy among the members of the forum and an exchange of different types of support. This was observed through members sharing personal experiences and both providing and receiving support. The most frequently observed types of support were emotional and informational support. Examples of emotional support and expressions of understanding and empathy from the work of Bakke [32] are as follows:

> *Hi, I feel your fear and confusion.*
>
> *[...] I am responding to you mainly because I wanted to tell you that you are NOT alone with your medication problems.*

An example of informational support was provided in the work of Stewart Loane et al [43]. A person asked the following:

> *Does anyone ever experience freezing that lasts for hours on end? Please reply urgently.*

Another member responded quickly, and the person who asked the question replied as follows:

> *[...] I tried several of the methods that you suggested and I have found one that works for me. I'm telling you it WORKS. I'm so excited! I have been so worried*

XSL•FO
RenderX

> *about what would happen if I were alone and I froze, and now I have a new freedom. Thank you.*

Stewart Loane et al [43] observed that new posts on the forum often started with a request for information and that in their responses, other members shared information, personal experiences, and emotional support. Overall, the authors of all 3 papers observed a real sense of community, belonging, and friendship on each of the platforms, which can be described as network support [43]. An example that illustrates this type of support was seen in the work of Attard and Coulson [31]: "I am glad I found this forum, makes me feel like I am not alone."

In the study by Lieberman et al [46], the authors researched the impact of group composition. Participants were divided into homogeneous (based on age or time since diagnosis) and heterogeneous groups. Although all groups improved on quality of life scores, participants in the homogeneous groups showed significant improvement in depression and PD symptoms compared with heterogeneous groups. These findings suggest that similarities between group members can improve the outcomes of peer support [46].

Benefits of Online Peer Support

Davis and Boellstorff [47] observed how 2 people with PD used the Second Life web-based platform. Through their ethnographic study they found that both participants were able to express themselves creatively on the platform. Through their web-based artworks and creative expressions, both people with PD were able to continue with creative parts of their previous jobs, and they also used art to express what it feels like to have PD. A sense of community was also observed here. Furthermore, one of the participants was living in a rural area, where it was difficult to find in-person support groups. In this case, the web-based platform provided a way to connect with other people with PD [47]. The work of Lieberman et al [45] showed that people with PD who participated in web-based groups felt freer to talk about certain topics compared with in-person groups. A participant shared the following [45]:

> *In an internet group, you are much freer to talk about things that you probably wouldn't in a F2F [face to face]. We got into discussion of sex [meds affecting sexual desire]. I know I wouldn't have discussed in a F2F.*

Challenges of Online Peer Support

Of the 4 studies, only 1 (25%) reported on the challenges related to online peer support communities for people with PD—a qualitative content analysis of a PD discussion forum [31]. Challenges were related to online peer support and the use of technology in general. Some were related to the behavior of group members, such as a lack of replies to posts and group members leaving without warning. This could be distressing for other members. An example that illustrates this is, "If you are out there please respond. I have searched the net for you dear friend and I would like to talk to you again" [31]. Other challenges were more related to the nature of discussion forums and web-based support in general, such as the absence of nonverbal communication, which at times could lead to misunderstandings, and the lack of personal information. Finally, some posts showed that, at times, it was difficult for people with PD to use a computer or other types of technology because of their symptoms: "Sometimes my PD prevents my fingers from being able to type. At other times they work fine, but my brain is a blob!" [31].

Furthermore, a study investigated the reasons why people would drop out of online PD support groups. Findings show that people who dropped out of the online peer support sessions had similar scores on depression, quality of life, and PD symptoms scales but had higher levels of anxiety before starting their participation [44].

Successful Elements of Online Peer Support

Several successful elements of online peer support for people with PD have been identified in this review. First, writing may help people reflect on their own situation and share things that may be difficult to express face to face [31]. Second, having homogeneous groups based on age or time since diagnosis leads to increased benefits for members [44-46]. The findings of Lieberman et al [46] show that people who participated in the homogeneous groups felt more committed to their group and had more positive feelings about the group during the first 5 meetings. Furthermore, only members from the homogeneous groups continued to stay in touch after the intervention ended [45]. Finally, although most studies included in this review analyzed moderated platforms, the study by Bakke [32] specifically looked at a physician-moderated platform. The author observed that members appreciated the opportunity to ask questions directly to a professional. A helpful feature in the forum design was clearly labeling the physician's comments [32].

Discussion

Principal Findings

This section presents the summary and interpretation of the findings, covering narrative synthesis element 3: exploring relationships within and between studies. To the best of our knowledge, this is the first review to systematically synthesize the literature on online peer support for people with PD. This review shows that online peer support can be a way for people with PD to stay socially connected, share experiences, and exchange support for managing daily life with PD. Furthermore, this review identified the successful elements of online peer support.

Benefits and Successful Elements of Online Peer Support

Overview

The main positive elements related to peer support are reciprocity and social support [14,16]. This finding has also been identified in this review, indicating that the benefits of peer support are not limited to in-person settings. Despite not knowing each other in person and not being physically close, this review shows that people with PD can find emotional support, mutual understanding, and empathy through web-based communities. Moreover, people with PD can build new friendships and expand their social networks. People can share

XSL•FO
RenderX

their personal experiences and provide and receive informational support and advice from others in similar situations. For example, people can share experiences with medication or how they manage PD symptoms in daily life. This is based on experiential knowledge, which is a combination of unique knowledge and expertise that people have because of their personal experiences of living with PD [17]. Sharing knowledge and learning from others' experience can contribute to developing coping skills for living with PD. This, in turn, can support people in living meaningful and satisfying lives despite having PD [10]. Similar findings have been published on online peer support groups for other conditions, including people with chronic illnesses [13] and Huntington disease [49,50]. This review supports previous research in that the benefits of peer support are not limited to a physical, in-person setting but can also be transferred via the internet. Elements that can make online peer support successful include having homogeneous groups [44-46] and having the option for participants to directly ask questions to a physician [32]. However, different people have different needs and preferences. Some who engage in online support may still miss in-person human interactions such as having a cup of tea together or being able to give someone a hug when they are upset [51].

There are also additional benefits to peer support in a web- or app-based setting. First, online peer support groups are available to a wide range of people, including those living in remote areas. For these people, it might be difficult to find in-person peer support groups in their local areas. PD symptoms may also impose additional challenges on people to travel to in-person peer support groups. Finally, the internet provides a form of anonymity. The anonymous nature of online peer support groups can make it easier for people to discuss taboo topics that are difficult to talk about in an in-person setting [23,45].

Challenges of Online Peer Support

Only a few studies in this review provided information on users' age or gender [44,46,47], whereas for the other studies, it was unknown. Information on group composition and personal information, such as age, gender, or time since diagnosis, is often unknown. A lack of such information can make it difficult to determine the extent to which members have things in common. This also highlights the challenge for people with PD in finding more specific peer support groups, such as young onset PD groups or groups for people who are newly diagnosed. The importance of similarity between group members was presented in the work of Lieberman et al [46]. These findings highlight a key element of peer support and something that defines whether someone is a peer: sharing similarities [14]. A lack of personal information was mostly the case for papers analyzing discussion forums, which could be because of the anonymous nature of such forums. The studies in this review that analyzed a discussion forum all used a publicly accessible platform. Reasons for using publicly accessible forums include ethical issues regarding informed consent and respect for members' privacy [31]. It could be that because of the public nature, either members did not have the option to share more personal information or members chose not to share that information [19].

Impact of Research Methods

Qualitative content analysis was conducted in 38% (3/8) of papers included in this review. Although this method provides insights into what is happening and being shared on the platform, it does not provide information about members' personal experiences. A number of aspects of this methodology remain unknown. First, the findings are highly dependent on researchers' interpretations. Although researchers can interpret the intention or underlying meaning of a post, it is often not possible to directly contact the author of the post and ask if this was indeed how they intended their message. Similarly, it is often not possible to directly contact the intended receiver of the post to confirm if they perceived the message in the way that the researcher interpreted it. These challenges can be addressed using qualitative research methods to directly explore users' experiences, as was done by Davis and Boellstorff [47] and Martínez-Pérez et al [48], or by setting up an online peer support intervention and performing pre-post measurements, as was done by Lieberman et al [45]. Second, on discussion forums and social media pages, all group members can often read all posts (besides private messages). This means that not only the intended receiver but also other members can read the posts. Many people can read it, but not everyone will respond to or participate in the discussion. When using a content analysis method, it remains unknown how people who only read the posts but not interact, also called *lurkers*, interpret the message and experience it [23]. Steadman and Pretorius [52] explored the impact of a Facebook group for people with MS on nonactive members. During individual interviews, people expressed that they still experienced social support despite not being actively involved in the discussions [52].

Third, the research into online peer support presented in this review might show an overly positive image of the online peer support group, as people who are active on the platform and post messages are often the ones that enjoy being part of the community. In many web-based communities, people can come and go when they want, and those who have negative experiences can leave the group without giving a reason. This means that negative experiences and potentially harmful aspects of online peer support groups remain underresearched. A potential negative experience identified in this review is the lack of response to messages [31], which has also been identified in another systematic review [23]. The authors stated that new members of an online peer support group are especially at risk of withdrawing after not receiving a response to their messages. The reason for this could be that new members may be more psychologically vulnerable and have certain expectations when joining the online peer support group [23]. When selecting a specific platform for research on online peer support, there is a risk of presenting an overly positive view of the platform and the experiences of its users. An alternative could be to explore the experiences with and opinions on online peer support in the wider PD community, for example, through a survey.

Limitations

This section covers narrative synthesis element 4: assessing the robustness of the synthesis. This systematic review only included studies on written communication between people with PD on

XSL•FO
RenderX

publicly available platforms. The database search did not identify any papers that included other platforms that can potentially be used for online peer support, such as videoconferencing platforms or social media platforms such as WhatsApp or Instagram. Therefore, the findings of this review are limited to the platforms covered in this review (discussion forums and Facebook groups) and cannot be generalized beyond these. Moreover, of the 8 studies, only 1 (13%) study has included findings on the potential challenges of online peer support [31]. As a result, this review may overrepresent the positive and beneficial aspects of online peer support and may not provide an accurate picture of the real-world experiences of people with PD who are part of such communities. In addition, within the studies, it was sometimes difficult to identify the contributions of technological, social, and individual elements to how people experienced online peer support. Finally, people have different preferences and needs, and online peer support may not be suitable for everyone living with PD. In addition, the physical symptoms of PD may be a barrier for people to use technology and to access online peer support communities. The views and experiences of people who are unable or do not want to engage in online peer support groups have not been presented in this review.

Recommendations for Future Research

For this review, no papers were identified that covered videoconferencing platforms that can be used for peer support; for example, Zoom, Skype, or Microsoft Teams. As these platforms have become more widely used since the COVID-19 pandemic, future research could explore how widely they are used among people with PD, and if and how they are used for peer support. Furthermore, research could focus on how people experience this form of online peer support and how it impacts their lives, as it is different in nature than what has been discussed in this review. More specifically, videoconferencing platforms include synchronous and verbal communication, often where one sees the other members. This reduces anonymity and adds a face-to-face element, in which nonverbal communication can be more prevalent.

Future research could also focus on using different methodologies for analyzing online peer support for people with PD. Direct assessment of users' personal experiences was also recommended by some of the studies included in this review [31,43,48]. Examples of these methods include individual interviews, focus groups, or surveys. It is necessary to learn how people with PD truly experience being part of an online peer support community and what the impact is on their daily lives. Furthermore, future research is needed to explore potential negative experiences people may have with online peer support, as these are currently underresearched. Qualitative methods, such as individual interviews and open-question surveys, can be used for this purpose. In addition, there is a group of people who are unable to access online peer support or use technology, for example, because of their PD symptoms. It is important to explore in more detail the barriers that people face and how they could overcome them. Some of the studies included in this review recommended investigating the use of voice assistive tools for people with PD [31,45]. Research into the use of such assistive tools for online peer support has already been conducted for people with ALS; for example, in the work of Caron and Light [53].

Conclusions

Peer support can be an extremely valuable source of social support for people with PD. More specifically, peer support can improve social health and support people with PD in living meaningful and satisfying lives, despite their condition. Sharing experiences with peers can improve feelings of empowerment and social connectedness and help people with PD develop new coping skills. Peer support is unique and cannot be replaced by family members, friends, or health care professionals who do not live with PD. The benefits of peer support are not limited to physical, in-person support groups but can be transferred via the internet. Online peer support is accessible to a wide range of people and is not limited by geographical barriers. This could make online peer support particularly suitable for those who do not have an in-person peer support group in their local area, or whose PD symptoms hinder them from traveling. However, research on the personal experiences of those who engage in online peer support and potential barriers to accessing online peer support remains limited. Future research could use qualitative methods, such as individual interviews, focus groups, and open-question surveys to explore these fields further.

Acknowledgments

The authors would like to thank Emma Young (deputy head of knowledge services) and Naomi Thorpe (senior information specialist) from the Library and Knowledge Services at Nottinghamshire Healthcare National Health Service Foundation Trust for their support in defining the search strategy and running the pilot searches for this review. The research presented in this paper was conducted as part of the Marie Curie Innovative Training Network action, H2020-MSCA-ITN-2018, under grant agreement 813196.

Authors' Contributions

EVG conducted the review and narrative synthesis and was the lead in writing of the manuscript. ARL provided a second review during all stages. OM was consulted as a third reviewer during the screening process. NC provided support for developing the search strategy. OM and MO provided feedback regarding the narrative synthesis process. All authors contributed to the writing of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Narrative synthesis methods.
[DOCX File , 25 KB - aging_v5i3e35425_app1.docx]

Multimedia Appendix 2
Quality assessment tables.
[DOCX File , 20 KB - aging_v5i3e35425_app2.docx]

References

1. Halli-Tierney AD, Luker J, Carroll DG. Parkinson disease. Am Fam Physician 2020 Dec 01;102(11):679-691 [FREE Full text] [Medline: 33252908]
2. GBD 2016 Parkinson's Disease Collaborators. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 2018 Nov;17(11):939-953 [FREE Full text] [doi: 10.1016/S1474-4422(18)30295-3] [Medline: 30287051]
3. Chand P, Litvan I. Parkinson's disease. In: Birren JE, editor. Encyclopedia of Gerontology. Amsterdam, The Netherlands: Elsevier; 2007:322-333.
4. The incidence and prevalence of Parkinson's in the UK. Results from the Clinical Practice Research Datalink Summary report. Parkinson's UK. 2018. URL: https://www.parkinsons.org.uk/sites/default/files/2018-01/CS2960%20Incidence%20and%20prevalence%20report%20branding%20summary%20report.pdf [accessed 2022-06-23]
5. Ambrosio L, Portillo MC, Rodriguez-Blazquez C, Rojo JM, Martinez-Martin P, EC-PC Validation Group. Influencing factors when living with Parkinson's disease: a cross-sectional study. J Clin Nurs 2019 Sep;28(17-18):3168-3176. [doi: 10.1111/jocn.14868] [Medline: 30938889]
6. Beaudet L, Ducharme F, L'Écuyer N, Chouinard S, Jodoin N, Panisset M. Development and evaluation of a dyadic intervention for elderly couples living with moderate-stage Parkinson disease. Appl Nurs Res 2015 Nov;28(4):e21-e27. [doi: 10.1016/j.apnr.2015.02.004] [Medline: 26324119]
7. Hellqvist C, Berterö C, Hagell P, Dizdar N, Sund-Levander M. Effects of self-management education for persons with Parkinson's disease and their care partners: a qualitative observational study in clinical care. Nurs Health Sci 2020 Sep;22(3):741-748. [doi: 10.1111/nhs.12721] [Medline: 32270898]
8. Perepezko K, Hinkle JT, Shepard MD, Fischer N, Broen MP, Leentjens AF, et al. Social role functioning in Parkinson's disease: a mixed-methods systematic review. Int J Geriatr Psychiatry 2019 Aug;34(8):1128-1138 [FREE Full text] [doi: 10.1002/gps.5137] [Medline: 31069845]
9. Dröes RM, Chattat R, Diaz A, Gove D, Graff M, Murphy K, INTERDEM sOcial Health Taskforce. Social health and dementia: a European consensus on the operationalization of the concept and directions for research and practice. Aging Ment Health 2017 Jan;21(1):4-17. [doi: 10.1080/13607863.2016.1254596] [Medline: 27869503]
10. Huber M, Knottnerus JA, Green L, van der Horst H, Jadad AR, Kromhout D, et al. How should we define health? BMJ 2011 Jul 26;343:d4163. [doi: 10.1136/bmj.d4163] [Medline: 21791490]
11. Vescovelli F, Sarti D, Ruini C. Subjective and psychological well-being in Parkinson's Disease: a systematic review. Acta Neurol Scand 2018 Jul;138(1):12-23. [doi: 10.1111/ane.12946] [Medline: 29696627]
12. Taylor SE. Social support: a review. In: The Oxford Handbook of Health Psychology. Oxford, UK: Oxford University Press; 2011:189-214.
13. Kingod N, Cleal B, Wahlberg A, Husted GR. Online peer-to-peer communities in the daily lives of people with chronic illness: a qualitative systematic review. Qual Health Res 2017 Jan;27(1):89-99. [doi: 10.1177/1049732316680203] [Medline: 27956659]
14. Keyes SE, Clarke CL, Wilkinson H, Alexjuk EJ, Wilcockson J, Robinson L, et al. "We're all thrown in the same boat … ": a qualitative analysis of peer support in dementia care. Dementia (London) 2016 Jul;15(4):560-577. [doi: 10.1177/1471301214529575] [Medline: 24742876]
15. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehabil J 2004;27(4):392-401. [doi: 10.2975/27.2004.392.401] [Medline: 15222150]
16. Barak A, Boniel-Nissim M, Suler J. Fostering empowerment in online support groups. Comput Human Behav 2008 Sep;24(5):1867-1883. [doi: 10.1016/j.chb.2008.02.004]
17. Dennis CL. Peer support within a health care context: a concept analysis. Int J Nurs Stud 2003 Mar;40(3):321-332. [doi: 10.1016/s0020-7489(02)00092-5] [Medline: 12605954]
18. Meier A. Technology-mediated groups. In: Garvin CD, Gutiérrez LM, Galinsky MJ, editors. Handbook of Social Work with Groups. New York, NY, USA: The Guilford Press; 2004:479-503.

XSL•FO
RenderX

19. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
20. Demiris G. The diffusion of virtual communities in health care: concepts and challenges. Patient Educ Couns 2006 Aug;62(2):178-188. [doi: 10.1016/j.pec.2005.10.003] [Medline: 16406472]
21. Wang X, Zhao K, Street N. Analyzing and predicting user participations in online health communities: a social support perspective. J Med Internet Res 2017 Apr 24;19(4):e130 [FREE Full text] [doi: 10.2196/jmir.6834] [Medline: 28438725]
22. Walther JB, Boyd S. Attraction to computer-mediated social support. In: Lin CA, Atkin D, editors. Communication Technology and Society: Audience Adoption and Uses. Cresskill, NJ, USA: Hampton Press; 2002:153-188.
23. Gatos D, Günay A, Kırlangıç G, Kuscu K, Yantac AE. How HCI bridges health and design in online health communities: a systematic review. In: Proceedings of the 2021 Designing Interactive Systems Conference. 2021 Presented at: DIS '21; June 28-July 2, 2021; Virtual p. 970-983. [doi: 10.1145/3461778.3462100]
24. Della Rosa S, Sen F. Health topics on Facebook groups: content analysis of posts in multiple sclerosis communities. Interact J Med Res 2019 Feb 11;8(1):e10146 [FREE Full text] [doi: 10.2196/10146] [Medline: 30741640]
25. Shavazi MA, Morowatisharifabad MA, Shavazi MT, Mirzaei M, Ardekani AM. Online social support for patients with multiple sclerosis: a thematic analysis of messages posted to a virtual support community. Int J Community Based Nurs Midwifery 2016 Jul;4(3):188-198 [FREE Full text] [Medline: 27382585]
26. Flickinger TE, DeBolt C, Waldman AL, Reynolds G, Cohn WF, Beach MC, et al. Social support in a virtual community: analysis of a clinic-affiliated online support group for persons living with HIV/AIDS. AIDS Behav 2017 Nov;21(11):3087-3099. [doi: 10.1007/s10461-016-1587-3] [Medline: 27766448]
27. Mo PK, Coulson NS. Are online support groups always beneficial? A qualitative exploration of the empowering and disempowering processes of participation within HIV/AIDS-related online support groups. Int J Nurs Stud 2014 Jul;51(7):983-993. [doi: 10.1016/j.ijnurstu.2013.11.006] [Medline: 24342449]
28. Hargreaves S, Bath PA, Duffin S, Ellis J. Sharing and empathy in digital spaces: qualitative study of online health forums for breast cancer and motor neuron disease (amyotrophic lateral sclerosis). J Med Internet Res 2018 Jun 14;20(6):e222 [FREE Full text] [doi: 10.2196/jmir.9709] [Medline: 29903695]
29. Kazmer MM, Lustria M, Cortese J, Burnett G, Kim JH, Ma J, et al. Distributed knowledge in an online patient support community: authority and discovery. J Assn Inf Sci Tec 2014 Feb 26;65(7):1319-1334. [doi: 10.1002/asi.23064]
30. Lu Y, Wu Y, Liu J, Li J, Zhang P. Understanding health care social media use from different stakeholder perspectives: a content analysis of an online health community. J Med Internet Res 2017 Apr 07;19(4):e109 [FREE Full text] [doi: 10.2196/jmir.7087] [Medline: 28389418]
31. Attard A, Coulson NS. A thematic analysis of patient communication in Parkinson's disease online support group discussion forums. Comput Human Behav 2012 Mar;28(2):500-506. [doi: 10.1016/j.chb.2011.10.022]
32. Bakke A. Trust-building in a patient forum: the interplay of professional and personal expertise. J Tech Writ Commun 2018 May 21;49(2):156-182. [doi: 10.1177/0047281618776222]
33. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC methods programme. Lancaster University. 2006 Apr. URL: https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersion1-April2006.pdf [accessed 2022-06-23]
34. 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. Int J Surg 2021 Apr;88:105906. [doi: 10.1016/j.ijsu.2021.105906] [Medline: 33789826]
35. Lakey B, Cohen S. Social support theory and measurement. In: Cohen S, Underwood LG, Gottlieb BH, editors. Social Support Measurement and Intervention: a Guide for Health and Social Scientists. Oxford, UK: Oxford University Press; 2000:29-52.
36. CASP Qualitative Checklist. Critical Appraisal Skills Programme. 2018. URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2022-06-23]
37. Centre for Reviews and Dissemination. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care. York, UK: Centre for Reviews and Dissemination; 2009:226.
38. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 1998 Jun;52(6):377-384 [FREE Full text] [doi: 10.1136/jech.52.6.377] [Medline: 9764259]
39. Crellin NE, Orrell M, McDermott O, Charlesworth G. Self-efficacy and health-related quality of life in family carers of people with dementia: a systematic review. Aging Ment Health 2014;18(8):954-969 [FREE Full text] [doi: 10.1080/13607863.2014.915921] [Medline: 24943873]
40. McDermott O, Crellin N, Ridder HM, Orrell M. Music therapy in dementia: a narrative synthesis systematic review. Int J Geriatr Psychiatry 2013 Aug;28(8):781-794. [doi: 10.1002/gps.3895] [Medline: 23080214]

XSL•FO
RenderX

41. Bayliss A, Currie L, McIntosh T, Gallagher R, Project Board, Infusion Therapy Standards, Watson M, et al. Infusion therapy standards: rapid evidence review. Royal College of Nursing. 2016. URL: https://www.eunetips.eu/fileadmin/pdf/Infusion_therapy_standards_Rapid_evidence_review.pdf [accessed 2022-06-23]
42. O'Connor SR, Tully MA, Ryan B, Bradley JM, Baxter GD, McDonough SM. Failure of a numerical quality assessment scale to identify potential risk of bias in a systematic review: a comparison study. BMC Res Notes 2015 Jun 06;8:224 [FREE Full text] [doi: 10.1186/s13104-015-1181-1] [Medline: 26048813]
43. Stewart Loane S, Webster CM, D'Alessandro S. Identifying consumer value co-created through social support within online health communities. J Macromarketing 2014 May 29;35(3):353-367. [doi: 10.1177/0276146714538055]
44. Lieberman MA. Psychological characteristics of people with Parkinson's disease who prematurely drop out of professionally led Internet chat support groups. Cyberpsychol Behav 2007 Dec;10(6):741-748. [doi: 10.1089/cpb.2007.9956] [Medline: 18085960]
45. Lieberman MA, Winzelberg A, Golant M, Wakahiro M, DiMinno M, Aminoff M, et al. Online support groups for Parkinson's patients: a pilot study of effectiveness. Soc Work Health Care 2005;42(2):23-38. [doi: 10.1300/J010v42n02_02] [Medline: 16390834]
46. Lieberman MA, Wizlenberg A, Golant M, Di Minno M. The impact of group composition on Internet support groups: homogeneous versus heterogeneous Parkinson's groups. Group Dyn Theory Res Pract 2005;9(4):239-250. [doi: 10.1037/1089-2699.9.4.239]
47. Davis DZ, Boellstorff T. Compulsive creativity: virtual worlds, disability, and digital capital. Int J Commun 2016;10:23 [FREE Full text]
48. Martínez-Pérez B, de la Torre-Díez I, Bargiela-Flórez B, López-Coronado M, Rodrigues JJ. Content analysis of neurodegenerative and mental diseases social groups. Health Informatics J 2015 Dec;21(4):267-283 [FREE Full text] [doi: 10.1177/1460458214525615] [Medline: 24698768]
49. Coulson NS, Buchanan H, Aubeeluck A. Social support in cyberspace: a content analysis of communication within a Huntington's disease online support group. Patient Educ Couns 2007 Oct;68(2):173-178. [doi: 10.1016/j.pec.2007.06.002] [Medline: 17629440]
50. Smedley RM, Coulson NS. Genetic testing for Huntington's disease: a thematic analysis of online support community messages. J Health Psychol 2021 Mar;26(4):580-594. [doi: 10.1177/1359105319826340] [Medline: 30696276]
51. Giebel C, Cannon J, Hanna K, Butchard S, Eley R, Gaughan A, et al. Impact of COVID-19 related social support service closures on people with dementia and unpaid carers: a qualitative study. Aging Ment Health 2021 Jul;25(7):1281-1288. [doi: 10.1080/13607863.2020.1822292] [Medline: 32954794]
52. Steadman J, Pretorius C. The impact of an online Facebook support group for people with multiple sclerosis on non-active users. Afr J Disabil 2014 Nov 21;3(1):132 [FREE Full text] [doi: 10.4102/ajod.v3i1.132] [Medline: 28730005]
53. Caron J, Light J. "My world has expanded even though I'm stuck at home": experiences of individuals with amyotrophic lateral sclerosis who use augmentative and alternative communication and social media. Am J Speech Lang Pathol 2015 Nov;24(4):680-695. [doi: 10.1044/2015_AJSLP-15-0010] [Medline: 26254447]

Abbreviations

ALS: amyotrophic lateral sclerosis
CASP: Critical Appraisal Skills Programme
MS: multiple sclerosis
PD: Parkinson disease
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by J Wang; submitted 04.12.21; peer-reviewed by J Taylor, S Choemprayong; comments to author 14.03.22; revised version received 27.03.22; accepted 31.03.22; published 27.07.22.

Please cite as:
Gerritzen EV, Lee AR, McDermott O, Coulson N, Orrell M
Online Peer Support for People With Parkinson Disease: Narrative Synthesis Systematic Review
JMIR Aging 2022;5(3):e35425
URL: https://aging.jmir.org/2022/3/e35425
doi:10.2196/35425
PMID:35896025

©Esther Vera Gerritzen, Abigail Rebecca Lee, Orii McDermott, Neil Coulson, Martin Orrell. Originally published in JMIR Aging (https://aging.jmir.org), 27.07.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution

XSL•FO
RenderX

License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Capturing Cognitive Aging in Vivo: Application of a Neuropsychological Framework for Emerging Digital Tools

Katherine Hackett[1], MA; Tania Giovannetti[1], PhD

Department of Psychology and Neuroscience, Temple University, Philadelphia, PA, United States

Corresponding Author:
Tania Giovannetti, PhD
Department of Psychology and Neuroscience
Temple University
1701 N 13th Street
Weiss Hall 6th Floor
Philadelphia, PA, 19122
United States
Phone: 1 215 204 4296
Email: tania.giovannetti@temple.edu

Abstract

As the global burden of dementia continues to plague our healthcare systems, efficient, objective, and sensitive tools to detect neurodegenerative disease and capture meaningful changes in everyday cognition are increasingly needed. Emerging digital tools present a promising option to address many drawbacks of current approaches, with contexts of use that include early detection, risk stratification, prognosis, and outcome measurement. However, conceptual models to guide hypotheses and interpretation of results from digital tools are lacking and are needed to sort and organize the large amount of continuous data from a variety of sensors. In this viewpoint, we propose a neuropsychological framework for use alongside a key emerging approach—digital phenotyping. The Variability in Everyday Behavior (VIBE) model is rooted in established trends from the neuropsychology, neurology, rehabilitation psychology, cognitive neuroscience, and computer science literature and links patterns of intraindividual variability, cognitive abilities, and everyday functioning across clinical stages from healthy to dementia. Based on the VIBE model, we present testable hypotheses to guide the design and interpretation of digital phenotyping studies that capture everyday cognition in vivo. We conclude with methodological considerations and future directions regarding the application of the digital phenotyping approach to improve the efficiency, accessibility, accuracy, and ecological validity of cognitive assessment in older adults.

(JMIR Aging 2022;5(3):e38130) doi:10.2196/38130

KEYWORDS

digital phenotyping; neuropsychology; aging; dementia; smartphone; neurological; psychological; older adults; aging

Introduction

The global burden of dementia, a clinical syndrome associated with cognitive deficits that impair everyday functioning, poses a tremendous and growing challenge to our healthcare system. As the worldwide population of older adults continues to increase and becomes more medically complex and diverse, the number of people that will develop Alzheimer's disease and related dementias (ADRD) without current pharmacologic treatments to improve cognition and function [1] is expected to triple from 55 million in 2021 to over 139 million by 2050 [2]. Estimates of disability-adjusted life years (ie, sum of years lost due to premature mortality and years lived with disability) indicate that ADRD is extremely burdensome to individuals diagnosed, their families, and their caregivers, ranking among the top 10 most burdensome diseases in the United States [3]. Early diagnosis and intervention before neuronal degeneration and functional disability begin presents one promising route to meaningfully delay disease burden and promote aging in place [4-6]. From a health economics perspective, it is estimated that early detection of the prodromal, mild cognitive impairment (MCI) stage [7] could save $7.9 trillion in the United States alone [8]. Novel digital methods have great potential for efficient, accessible, reliable, and accurate assessment of early cognitive changes reflecting ADRD. However, to be most effective, digital tools should be informed by conceptual models that explain and predict early cognitive changes.

In this viewpoint, we focus on the application of digital phenotyping to assess age-related changes in functional cognition, with contexts of use that include early detection, risk

XSL•FO
RenderX

stratification, prognosis, and outcome measurement. We begin by outlining current approaches to detecting pathological cognitive change along with their notable drawbacks. The digital phenotyping approach is introduced as a promising complementary method. We then present a neuropsychological framework of everyday cognitive and functional changes, termed the Variability in Everyday Behavior (VIBE) model, which can be used to inform studies and generate testable hypotheses in the context of digital phenotyping. Supporting literature that was integrated to develop the VIBE model is also summarized. We conclude with methodological considerations and future directions regarding the digital phenotyping approach.

Current Approaches to Early Detection of Pathological Change

Neurodegenerative pathology may be directly measured in the brain tissue and detected in cerebrospinal fluid (CSF) and blood [9]; biological measures are classified using biomarker-based diagnostic frameworks for ADRD [10,11]. Importantly, existing methods of biomarker testing are expensive, not widely available, and may be invasive depending on the methodology (eg, lumbar puncture). More concerning, however, is that biological indicators of neurodegenerative disease yield limited information on clinical outcomes such as progression, cognitive abilities, and everyday functioning [12,13]. For example, approximately 30% of individuals with substantial amyloid burden—a core Alzheimer's disease (AD) biomarker—fail to show clinical symptoms of dementia, whereas up to 25% of individuals who meet clinical criteria for AD have no or limited amyloid burden [14]. The prioritization of biological outcomes is also concerning given that clinical outcomes such as cognitive and functional abilities are most predictive of quality of life, cost of care, and independence, which are precisely the outcomes that individuals diagnosed, their caregivers, healthcare professionals, and policy makers most value [15].

Neuropsychological assessment is less expensive and invasive compared to biomarker testing and is currently used for clinical staging, differential diagnosis, tracking change in cognitive functioning over time, and informing personalized recommendations. The neuropsychological measures that are used for clinical assessments have undergone extensive psychometric validation and are informed by cognitive neuroscience theories. At present, neuropsychological test results are a key component of clinical diagnostic criteria for dementia and MCI [10,16,17] and serve as a primary end point in most clinical trials [18]. In recent years, several neuropsychological tests and composite measures have shown promise in identifying very early, subtle changes that occur in presymptomatic disease stages [19-21].

Nevertheless, current assessment methods present methodological drawbacks, including lengthy and resource-intensive in-person testing sessions that are often inaccessible to individuals from underserved or rural communities, highly controlled testing environments that foster limited ecological validity and test anxiety, burdensome and error-prone scoring procedures, and limited data sharing infrastructures. Traditional assessments take place on a single occasion representing a one-time snapshot that may not reflect an individual's typical range of performance or intervening contextual factors [22,23]. Even when repeat testing is performed, practice effects between sessions may obscure subtle but meaningful cognitive decline [24].

New mobile and computerized platforms with enhanced efficiency and sensitivity, such as repeat ambulatory cognitive assessments, address some of these methodological drawbacks [25] and have been examined in various studies among populations comprising healthy and community-based older adults [22,26], those with preclinical AD [27], and those with MCI or early dementia [28]. However, many of these methods continue to be (A) modeled after traditional tests that measure isolated cognitive domains with limited ecological validity; (B) susceptible to practice effects [29]; (C) influenced by socioeconomic status and cultural factors [30-32]; and (D) prone to challenges with adherence even among highly motivated and engaged individuals, particularly with longer study durations [27,33,34]. Thus, while tremendous advances have been made in the realm of digital cognitive assessment, existing methods continue to show limited generalizability to diverse populations and real-world settings, even when used at home outside of the clinic. The strengths and weaknesses of the current approaches are summarized in Table 1.

XSL•FO
RenderX

Table 1. Strengths and weaknesses of current approaches to detect pathological change.

Approaches	Strengths	Weaknesses
Biomarker testing	• Objective measurement of disease presence in the body • Good sensitivity/early detection of target pathology • Ability to localize pathology • Ability to identify specific pathology	• High cost • Limited accessibility • Potentially invasive (CSF[a] and blood biomarkers) • Limited correspondence with functional outcomes • Limited prognostic value • Interpretation can be subjective
Traditional neuropsychological assessment	• Extensively validated and informed by cognitive neuroscience theories • Noninvasive • Measure discrete cognitive abilities • Inform personalized recommendations • Moderate correspondence with functional outcomes	• Limited accessibility • Lengthy and error-prone administration and scoring procedures • Highly controlled environment and tasks/limited ecological validity • Limited sensitivity to early decline • Single time point without context • Practice effects at repeat administration • Influenced by socioeconomic and cultural factors
Mobile cognitive assessment	• Brief administration • Improved accessibility • Potential for increased sensitivity • Noninvasive • Ability to assess cognition in everyday context and across multiple time points • Possible reduction in test anxiety	• Challenges in adherence • Practice effects at repeat administration • Impact of hardware and software differences when personal devices are used • Continued impact of socioeconomic/cultural factors • Uncontrolled testing environment may lead to increased measurement error/noise

[a]CSF: cerebrospinal fluid.

Digital Phenotyping

Emerging digital tools lend a unique opportunity to address many of the drawbacks of traditional, computerized, and mobile cognitive testing. One such method is digital phenotyping, an innovative approach that utilizes the "moment-by-moment quantification of the individual-level human phenotype in situ" based on interactions with technology, including smartphones and smart home devices, to capture social and behavioral data passively, continuously, and with minimal interference [35-37]. Because most everyday tasks require the coordinated effort of multiple cognitive processes and are highly context dependent, digital phenotyping data collected in this passive manner may provide a more naturalistic, comprehensive, and nuanced understanding of behavior and cognition as compared to traditional active assessment methods that take place in the clinic, the lab, or during a discrete period of time. Contrary to standardized neuropsychological tasks that are highly related to educational quality [38] and other sociocultural factors [39], digital proxies of everyday behavior captured in someone's natural environment may yield a less biased measure of cognition and function, particularly when methods rely on longitudinal monitoring of individual change. Furthermore, high frequency continuous data have the potential to improve sensitivity and reliability and reduce the sample size requirements needed to detect subtle differences between groups or among individuals over time [40].

Smartphones, which are ubiquitous, are equipped with a host of embedded sensors that are common across different devices and may be leveraged to passively assess everyday activities and behaviors. Preliminary studies have investigated smartphone-based digital biomarkers (via sensor and application use data) to measure specific behaviors and offer support and validation for call and text message logs [41] as well as call reciprocity [42] as measures of social patterns; WiFi/Bluetooth signals as a proxy for social engagement (time spent proximal to other people) [37]; GPS movement trajectories and keystroke data as measures of mood [43-45]; and accelerometer data to infer sleep patterns [46]. The validity of smartphone digital phenotyping has been demonstrated in mental health and neurological populations, with results supporting the predictive utility of a range of smartphone data for daily stress levels [47], changes in depression and loneliness [43,46,47], psychosis onset and relapse [48-50], suicide risk [47], speech changes [51], and biological rhythms [52].

Other studies have attempted to identify digital markers that reflect underlying cognitive abilities. A 2018 study of 27 healthy young adults [53] followed by a 2019 study of 84 healthy older adults [54] demonstrated significant associations between smartphone metrics (eg, number of apps used, usage by hour of day, swipes, and keystroke events) and performance on standard cognitive tests. Of note, these studies were exploratory in nature and lacked a priori hypotheses to guide analyses. A separate pilot study of adults with and without bipolar disorder examined performance on a digital trail making test and found associations between smartphone typing speed and typing speed variability and test performance, suggesting a possible link between executive functioning and keystroke measures [29]. In the context of MCI and dementia, a feasibility study employing multiple sensor streams and machine learning models identified 5 digital features that discriminated symptomatic (MCI, mild AD) from asymptomatic groups; these features included typing speed, regularity in behavior (via first and last phone use), number of received text messages, reliance on helper apps, and

XSL•FO
RenderX

survey compliance [55]. As noted by the authors of the aforementioned pilot and feasibility studies, a major limitation was the small sample sizes, which limited interpretability.

Indeed, although preliminary studies have laid the groundwork for exploring relationships between passive digital variables and standard measures of cognition, the lack of integrative theoretical models to inform interpretation of large continuous datasets represents a major gap [23]. As digital tools and machine learning approaches become increasingly sophisticated, it is critical that theoretically sound models are developed to avoid scattershot analyses and spurious findings and to facilitate interpretability [56]. Furthermore, as technologies inevitably continue to evolve, the development of testable models that are agnostic to hardware and software differences is key to the continued validation of passive approaches [56-58]. Therefore, we propose a neuropsychological framework to guide studies using emerging digital tools to assess age-related cognitive and functional decline. The VIBE model integrates established findings regarding intraindividual variability, cognitive abilities, and everyday functioning in the context of aging and ADRD. Importantly, the VIBE model generates specific, testable hypotheses grounded in theory that may inform the design and interpretation of future digital phenotyping studies and represents a preliminary step toward establishing conceptual guidelines for the field.

Approach to Framework Development

The VIBE model resulted from an in-depth review of the neuropsychology, neurology, neuroscience, rehabilitation psychology, and computer science literature. Consistent findings in both performance level and intraindividual variability were identified across the spectrum of cognitive impairment and interpreted in the context of known patterns of cognitive change and their underlying mechanisms. The literature review was used to conceptualize changes in everyday behavior across the spectrum from healthy aging to ADRD and how these changes would be captured by digital phenotyping approaches. For example, the increased variability in standardized cognitive testing and everyday task performance in people with MCI is expected to result in meaningful variability in passive smartphone sensor data in digital phenotyping studies. Without a framework to guide analyses, aggregate data might be prioritized over meaningful variability, which could be misinterpreted as a nuisance (ie, “noise”). Therefore, the VIBE model integrates and extends existing findings to provide structure, guidance, and optimize digital phenotyping study designs.

Our Proposed Framework

Early stages of pathological aging (ie, MCI) are associated with mild isolated decrements on standardized cognitive tests, subtle difficulties with complex activities of daily living, and increased variability in both cognitive and functional measures. Later stages (ie, dementia) are characterized by greater cognitive and functional impairment, reduced activity and task accomplishment, and less variability in cognitive and functional performance. Table 2 provides a summary of these trends. Multimedia Appendix 1 contains a comprehensive review of the supporting literature [59-107].

Table 2. Summary of background literature supporting the Variability in Everyday Behavior (VIBE) framework.

	Healthy aging	Early decline (MCI[a])	Later decline (dementia)
Cognitive ability	• Subtle declines within normative limits	• Impaired performance on 1+ domain according to normative scores	• Impaired performance on multiple domains according to normative scores
Cognitive variability	• Increased variability versus younger adults	• Increased variability versus healthy older adults • Increased variability predicts further decline and poorer cognition	• Less variability than MCI for complex tasks at floor • Increased variability than MCI for simple tasks
Everyday functioning	• Subtle changes/inefficient behaviors (microerrors) • Fully independent	• Difficulty with complex tasks • Independent with some compensatory strategy use • Inefficient (commission errors) and more variable than healthy older adults	• Impaired for basic and complex tasks • Dependent • Outright failure to complete tasks (omission errors)

[a]MCI: mild cognitive impairment.

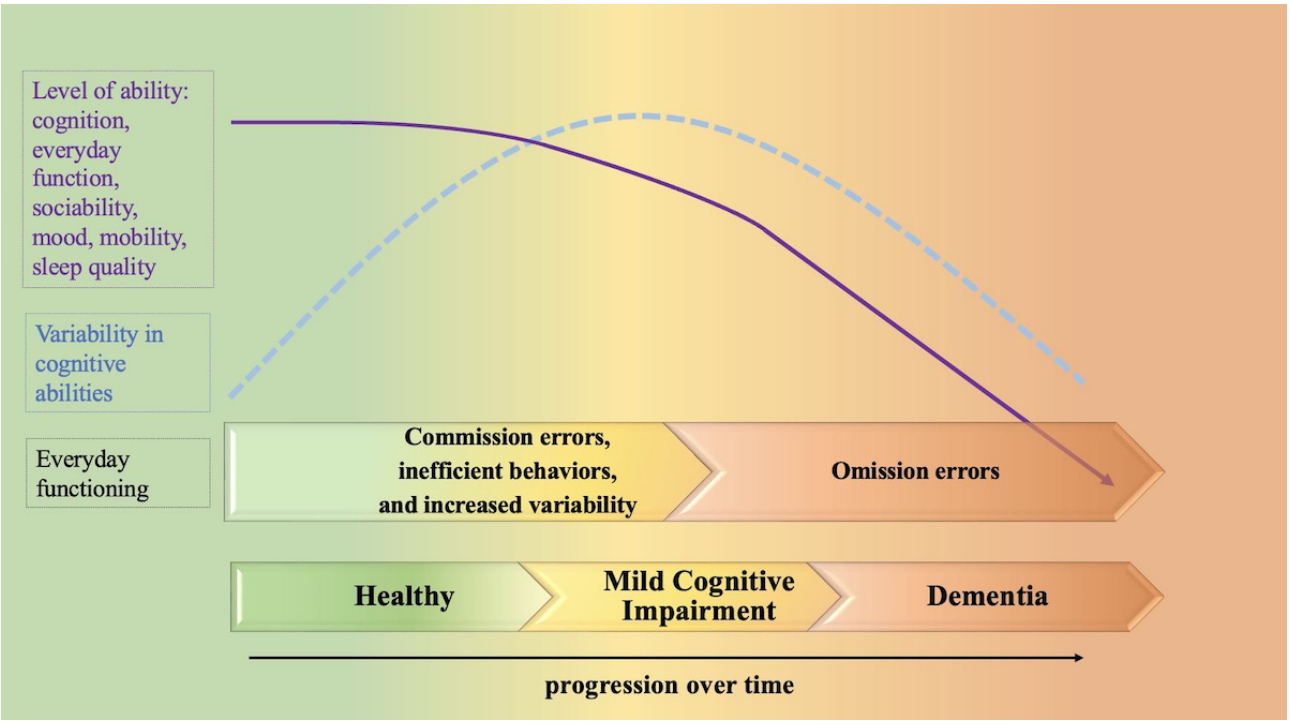
Theoretical models from computational science offer a useful framework for understanding changes in ability level and variability in the progression of pathological aging. The term “graceful degradation” is used to characterize the way in which complex systems maintain functionality in the face of mild damage or problematic changes in the environment [108]. From a neuropsychological perspective, increased inefficiency and variability in the early stage of decline may stem from faulty executive control mechanisms governed by the prefrontal cortex and associated white matter projections, which, according to a framework proposed by Giovannetti and colleagues [109], are essential to modulate goal activations, enable smooth transitions between goals, and inhibit inappropriate activations from internal or external distractors during everyday tasks. Reductions in extrastriatal dopaminergic neuromodulation required for consistent cognitive control in early stages of dementia support this framework [110-112]. Indeed, long-standing explanations for the link between inconsistency and neurologic disease

XSL•FO
RenderX

include impaired neural networks, functional connectivity, and executive functioning [113-115]. An alternative framework from which to interpret early patterns of inefficiency and variability, particularly in the absence of executive function deficits, is the resource theory [116], which originates from the cognitive rehabilitation literature. This theory posits that early damage to any nonspecific brain region depletes overall cognitive resources and leads to errors in task performance and that the level—not the type—of cognitive impairment is critical in determining functioning [109]. As a result of mild resource depletion, compensatory strategies are engaged to allow the system to function, but at a cost (ie, inefficiently, slowly, and inconsistently). In moderate-to-severe stages, greater decrements are observed across multiple cognitive domains, basic activities of daily living are impaired, and patterns of variability are less clear because people are generally less active.

Considering this, we propose the Variability in Everyday Behavior (VIBE) model as a dual-pronged neuropsychological framework that integrates trends in variability (see Figure 1, blue dotted line showing a U-shaped pattern peaking at MCI) and declining ability level (see Figure 1, solid purple line showing a negative linear trend) that are observed across the cognitive aging spectrum. The VIBE model proposes a theoretical foundation from which to evaluate metrics of everyday behavior and cognition captured by the digital phenotyping approach, both in studies examining cross-sectional differences in individuals with different levels of cognitive impairment, and over time in individuals with progressive neurodegenerative disease in longitudinal designs. For example, decreasing cognitive abilities may be indexed by decreases in social activity [117,118], technology usage [119,120], positive mood (ie, increased depressive symptoms [121]), and range of movement/physical activity [122], which can all be inferred from passive sensor metrics. These activity metrics tend to remain stable in earlier stages and begin to decline more notably in the transition from MCI to dementia. A simultaneous examination of intraindividual variability within these metrics across a longitudinal study period may reflect increased day-to-day variability as early as the healthy to MCI transition stage, as individuals engage reserve mechanisms and compensatory strategies to combat mild difficulties and inefficiencies (eg, commission errors). On metrics/activities where dementia-level performance is at floor (eg, movement trajectories outside the home, text messaging, other complex activities where compensatory mechanisms have failed and task goals are no longer achieved; ie, omission errors), we expect variability to decrease below that which we observe in MCI (Figure 1, blue dotted line).

Figure 1. The Variability in Everyday Behavior (VIBE) model of intraindividual variability, cognitive abilities, and everyday functioning for pathological cognitive decline in older adults.



The existing literature is less clear on patterns of variability in the transition from MCI to dementia [123], and we acknowledge the possibility that for relatively simple activities that individuals with mild dementia still perform (eg, movement trajectories within the home, incoming phone calls, sleep/wake cycle), variability may continue to increase in the mild dementia stage followed by eventual decline as abilities further decline. Thus, model predictions should be tested and interpreted with attention to task demands, as well as other contextual features, including the time of day [29], mood, and technology use habits. In other words, the progression from increased variability to decreased variability and complete failure to act depicted in Figure 1 is expected with increasing severity of impairment, though impairment level is determined by more than just clinical status. There may be some period—likely at the transition between MCI and dementia—where contextual factors (eg, task complexity, time of day, external distractors) interact with clinical status to influence level of impairment on metrics of

XSL•FO
RenderX

everyday behavior. For example, a person with mild dementia may show marked impairment and decreased variability in financial tasks but may show only mild impairment and increased variability in meal preparation until later in the course of their illness when both tasks are impaired, and variability is diminished. Thus, task effects should be carefully considered, particularly at the boundary of MCI and dementia.

Application of the VIBE Framework Through Digital Phenotyping Studies

Digital phenotyping using personal smartphone devices represents a promising method to examine age-related changes in functional cognition according to our proposed framework. Study designs may take a variety of forms, but initial studies should include collection of clinically relevant validation measures and longitudinal monitoring. One potential protocol would involve comprehensive baseline assessment to gather gold-standard validation data on function, cognition, mood, and other participant features such as demographics, attitudes toward and experience with technology, and technology use habits that are likely to influence resulting digital data. A period of passive longitudinal monitoring using available, open-source digital phenotyping applications (eg, Beiwe [37], mindLAMP [124]) would involve collection of a host of sensor and application data, including the examples listed in Textbox 1.

The VIBE framework enables systematic selection and analysis of the mobility, sociability, and device activity features from Textbox 1 to obtain activity and variability metrics that could be tested according to a priori hypotheses. A list of nonexhaustive, sample hypotheses derived from the VIBE model that are appropriate for cross-sectional studies of older adults across the cognitive aging spectrum is included in Table 3.

XSL•FO
RenderX

Textbox 1. Example digital phenotyping metrics across 3 feature domains.

Mobility

- Time spent at home
- Distance traveled
- Radius of gyration
- Maximum diameter
- Maximum distance from home
- Number of significant locations
- Average flight length
- Standard deviation of flight length
- Average flight duration
- Standard deviation of flight duration
- Fraction of the day spent stationary
- Significant location entropy
- Minutes of GPS data missing
- Physical circadian rhythm
- Average sleep duration
- Standard deviation of sleep duration

Sociability

- Number of outgoing texts
- Total outgoing text length
- Number of incoming texts
- Total incoming text length
- Texting reciprocity
- Texting responsiveness
- Number of outgoing calls
- Total outgoing call duration
- Number of incoming calls
- Total incoming call durations
- Call reciprocity
- Call responsiveness

Device activity

- Average battery level
- Total phone off/on events
- Total charge initiations
- Total screen on/off events
- Total application launches
- Application switches
- Central processing unit (CPU) utilization
- Time to view daily survey
- Time to complete daily survey
- Time of first/last screen on event

XSL•FO
RenderX

Table 3. Sample hypotheses informed by the Variability in Everyday Behavior (VIBE) model.

Digital phenotyping feature domain	Total activity level metrics	Across-day variability metrics
Mobility	Average distance traveled from home will decline from healthy to MCI[a] to dementia.	Variability in distance traveled from home will be highest in MCI versus healthy/dementia.
Sociability	Average number of outgoing calls will decline from healthy to MCI to dementia.	Variability in daily average outgoing text length will be highest in MCI versus healthy/dementia.
Device activity	Average number of application launches will decline from healthy to MCI to dementia.	Variability in daily number of screen on/off events will be greater in MCI versus healthy/dementia.
Time of day effects	Average time of first phone use will decline from healthy (earlier) to MCI to dementia (later).	Time of first phone use will be most variable in MCI versus healthy/dementia.

[a]MCI: mild cognitive impairment.

Methodological Considerations of the Digital Phenotyping Approach

There are a host of important methodological factors that must be thoughtfully considered when conducting such studies, many of which remain unresolved. Cross-device compatibility is a concern that becomes relevant when participants use their own personal devices for data collection. Individual devices may differ in operating system, screen size, sensor sampling rates, and more. These device differences impact user interactions and the quality of data that is collected; they are also related to socioeconomic status and other important participant features and thus cannot be simply covaried in analyses. A single study-issued device may be provided to all participants to standardize data collection and ensure that individuals from underserved groups have an equal opportunity to participate in such studies. However, introducing new technology creates a deviation from participants' routines, diminishing ecological validity and posing more demands on everyday functioning. Therefore, the personal versus study-provided device decision must be weighed according to the study population and specific aims [27,33]. Although there is a critical concern that studies employing personal digital devices will serve to widen existing health disparities, rates of smartphone ownership—particularly among diverse individuals—have skyrocketed in recent years to include a total of 85% of Americans as of 2021, up from just 35% in 2011 [125]. This rate is consistent across individuals who identify as White (85%), Black (83%), and Hispanic (85%) and is only slightly lower (76%) for individuals with a household annual income less than US $30,000. Therefore, although careful attention must be paid to ensure smartphone studies are equitable, accessible, and generalizable to all, the increased affordability of smartphones may alleviate this concern. Relatedly, recruitment efforts should ensure diverse representation within digital phenotyping studies to investigate the generalizability of these methods. Updates to hardware, software, and allowable permissions (ie, which sensors an app can collect) are occurring at increasingly frequent rates as technology evolves, presenting an additional challenge to the continued validation and generalizability of such approaches. Thus, a device- and operating system–agnostic theoretical model, such as the VIBE model, from which to continually evaluate new data is critically important.

The naturalistic and passive collection of data in a completely unstandardized fashion presents an additional challenge in making between-group comparisons [56], and it remains undetermined whether between-group differences in metrics of interest will emerge despite individual differences in everyday phone use. The most powerful insights from the digital phenotyping approach may be realized by monitoring intraindividual change over longer periods of time, which would require theoretically informed statistical models to make generalizable claims in n-of-1 trials [56]. Another open question relates to the various sampling rates that can be selected to collect raw data from phone sensors and applications, which should be considered in the context of the scientific question and device battery limitations. Although most software platforms include default settings for sensor sampling (eg, GPS sampled at 1 Hz when the phone is in motion, WiFi signals recorded every 5 minutes), they also allow for customization of sampling rates [37]. A variety of GPS sampling rates have been applied across several studies of primarily young adult participants [48,49], and statistical approaches for imputing missing mobility data have been developed [126]. However, limited studies have examined the incremental utility of increased sampling rates across sensors other than mobility for making predictions of interest. Older adult phone users may require less frequent sampling due to less activity, though this may result in a restricted range of variability and impact findings. Determining the minimum necessary sampling frequency for smartphone data is directly tied to feasibility and is critical to inform the design of future studies, as greater frequencies come with greater costs (ie, increasingly expensive sensors, decreased battery life, increased storage needs). This also applies to the optimal length of the data collection period and the study sample size, which may differ depending on the population of interest and the study design [120], and are not appropriately determined using traditional power calculation methods. Barnett and colleagues [127] recommend the use of generalized linear mixed models and change point detection methods to inform the sample size and study duration necessary to achieve adequate power in such studies.

Digital phenotyping studies may employ a combination of passive and active data streams. In active data collection, users are prompted to complete a standardized test or survey on their smartphones, which can be used to yield key contextual information to inform the interpretation of passive sensor data [23,37,128-130]. However, this type of engagement detracts

XSL•FO
RenderX

from the unobtrusive, naturalistic nature of pure passive monitoring, and it is unclear which types of active data are most useful when attempting to infer cognition from passive digital data. These methodological questions around sampling frequency and active data collection have not yet been explored in a population of older adult phone users, whose usage patterns may differ and may require increased sampling frequencies or increased active data than younger adults to accurately infer clinically relevant information.

It is also important to establish the context of use of the digital phenotyping approach and determine whether it is best applied as a risk, diagnostic, monitoring, prognostic, or outcome measurement tool. Regulatory agencies like the US Food and Drug Administration and pharmaceutical companies have increasingly recognized the potential of digital devices as a source of "real-world data" and "real-world evidence," with the capability to monitor health status and clinical response over time and yield new insights about long-term health outcomes in the real world, outside of traditional randomized controlled trials [131]. However, as thoughtfully outlined by O'Bryant and colleagues [9], there are many challenges associated with translating new biomarker discoveries from research domains to routine clinical settings. For this to occur, standardization of the underlying platforms and data frameworks is critical to help make these data more uniform, interoperable, reproducible, and actionable [124]. Questions of scalability, manufacturability, intellectual property law, and regulatory considerations, including inconsistent governance of entities conducting digital health research [132], should not be disregarded [9]. In particular, the point at which mobile digital phenotyping applications are considered "software as a medical device" is ambiguous in the face of rapidly evolving regulatory guidance [133]. Finally, and most importantly, privacy and security concerns must be addressed, and protections of confidentiality must be clearly and continuously communicated to users and participants. Deidentification using study identification numbers, industry-standard encryption methods, storage of encrypted data on secure severs, and ongoing transparency and control over personal data are examples of privacy considerations that should be carefully addressed at the study design phase. Given the extent of personal and sensitive health information involved, prospective risk assessment using tools like the Digital Health Checklist for Researchers should be completed beforehand to evaluate risks and benefits and ensure safe and responsible use of digital tools [132,134]. Importantly, the development and enforcement of privacy standards that are applied consistently across studies will be key to the success of this burgeoning field [35].

Benefits of the Digital Phenotyping Approach

Despite the numerous unresolved challenges and considerations outlined above, the potential for the digital phenotyping approach to yield ecologically valid and sensitive information on changes in everyday cognition is increasingly apparent. The benefits of emerging digital approaches are outlined in detail in the recent American Psychological Association Handbook of Neuropsychology [57]. To reiterate a few, sample size requirements are reduced when using continuous largescale data, and subtle fluctuations can be captured when data are sampled at such high rates, lending a highly sensitive scale that is captured in vivo. The use of personal smartphone devices represents a complex activity of daily living, thus creating an ideal platform to capture changes that occur early in the disease phase. Early detection of decline provides an opportunity for early intervention, which can lead to notable cost savings and reduced disability-adjusted life years, as noted earlier. Increased smartphone ownership lends increased accessibility compared to traditional methods. Passive data are objective and do not rely on current or retrospective self-report. However, it is possible that the most optimal application of this approach involves a blend of passive phenotyping, ecological momentary assessment for context, and burst cognitive testing to uncover the mechanisms of how changes in cognition within and across days relate to changes in behavior. Additionally, within-person n-of-1 designs may be increasingly sensitive and may address the interpretive challenges of between-groups designs. Finally, emerging digital methods should be considered complementary to traditional neuropsychological evaluations that remain the gold standard tool for diagnosis and intervention. If shown to be valid, emerging digital tools may represent a sensitive and accessible first line measure for early detection, risk stratification, and change in response to interventions.

Conclusions

Traditional approaches to measuring age-related changes in cognition and function provide valuable and distinct insights. Notable strengths of biomarker, traditional, and mobile cognitive assessments include extensive validation, measurement of discrete cognitive abilities, and localization of pathology (Table 1). At the same time, these approaches present many drawbacks that have become increasingly apparent in the face of technological advances that offer innovative solutions. The digital phenotyping approach is just 1 example of a novel tool that can serve as an increasingly accessible, efficient, sensitive, and personalized complement. Importantly, digital phenotyping remains in its infancy, and many methodological considerations warrant careful attention. Multiple sources of within-person differences (eg, hardware, software, technology habits, daily routines), as well as interpretive challenges of large-scale continuous datasets, make comparisons across individuals and across studies near impossible without a sound theoretical model from which to design and interpret such studies. The VIBE model, supported by decades of cross-discipline literature in neuropsychology, neurology, neuroscience, rehabilitation psychology, and computer science, proposes testable hypotheses (see Figure 1 and Table 3) that can be used in future digital phenotyping studies to provide novel, valuable, and clinically interpretable insights into meaningful changes in everyday behavior and cognition.

XSL•FO
RenderX

Acknowledgments

We acknowledge funding from the National Institutes of Health (NIH) and National Institute on Aging (NIA), including the following grants: F31AG069444 (author KH) and R21AG066771, R21AG060422, and R01AG062503 (author TG). We also acknowledge Ian Barnett for his cosponsorship of KH's F31 grant, including consultation on digital phenotyping methodologies and related literature.

Authors' Contributions

KH and TG conceptualized the proposed framework and collaboratively reviewed the pertinent literature. KH drafted the manuscript. TG reviewed and provided conceptual and editorial advice for the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supporting background literature.
[DOCX File , 19 KB - aging_v5i3e38130_app1.docx]

References

1. Branca C, Oddo S. Paving the way for new clinical trials for Alzheimer's disease. Biol Psychiatry 2017 Jan 15;81(2):88-89. [doi: 10.1016/j.biopsych.2016.10.016] [Medline: 27938878]
2. Dementia key facts 2021. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2022-04-01]
3. 2022 Alzheimer's disease facts and figures. Alzheimers Dement 2022 Apr 14;18(4):700-789. [doi: 10.1002/alz.12638] [Medline: 35289055]
4. Schelke MW, Hackett K, Chen JL, Shih C, Shum J, Montgomery ME, et al. Nutritional interventions for Alzheimer's prevention: a clinical precision medicine approach. Ann N Y Acad Sci 2016 Mar;1367(1):50-56 [FREE Full text] [doi: 10.1111/nyas.13070] [Medline: 27116241]
5. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of 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):280-292 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.003] [Medline: 21514248]
6. Sperling RA, Rentz DM, Johnson KA, Karlawish J, Donohue M, Salmon DP, et al. The A4 study: stopping AD before symptoms begin? Sci Transl Med 2014 Mar 19;6(228):228fs13 [FREE Full text] [doi: 10.1126/scitranslmed.3007941] [Medline: 24648338]
7. 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]
8. Alzheimer's Association. 2019 Alzheimer's disease facts and figures. Alzheimers Dement 2019 Mar;15(3):321-387. [doi: 10.1016/j.jalz.2019.01.010]
9. O'Bryant SE, Mielke MM, Rissman RA, Lista S, Vanderstichele H, Zetterberg H, et al. Blood-based biomarkers in Alzheimer disease: Current state of the science and a novel collaborative paradigm for advancing from discovery to clinic. Alzheimers Dement 2017 Jan 18;13(1):45-58 [FREE Full text] [doi: 10.1016/j.jalz.2016.09.014] [Medline: 27870940]
10. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Contributors. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. Alzheimers Dement 2018 Apr;14(4):535-562 [FREE Full text] [doi: 10.1016/j.jalz.2018.02.018] [Medline: 29653606]
11. McKeith IG, Boeve BF, Dickson DW, Halliday G, Taylor J, Weintraub D, et al. Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB Consortium. Neurology 2017 Jul 04;89(1):88-100 [FREE Full text] [doi: 10.1212/WNL.0000000000004058] [Medline: 28592453]
12. Negash S, Wilson RS, Leurgans SE, Wolk DA, Schneider JA, Buchman AS, et al. Resilient brain aging: characterization of discordance between Alzheimer's disease pathology and cognition. Curr Alzheimer Res 2013 Oct;10(8):844-851 [FREE Full text] [doi: 10.2174/15672050113109990157] [Medline: 23919768]
13. Brookmeyer R, Abdalla N. Estimation of lifetime risks of Alzheimer's disease dementia using biomarkers for preclinical disease. Alzheimers Dement 2018 May 22;14(8):981-988. [doi: 10.1016/j.jalz.2018.03.005]
14. Glymour MM, Brickman AM, Kivimaki M, Mayeda ER, Chêne G, Dufouil C, et al. Will biomarker-based diagnosis of Alzheimer's disease maximize scientific progress? Evaluating proposed diagnostic criteria. Eur J Epidemiol 2018 Jul 9;33(7):607-612 [FREE Full text] [doi: 10.1007/s10654-018-0418-4] [Medline: 29948371]

XSL•FO
RenderX

15. Tochel C, Smith M, Baldwin H, Gustavsson A, Ly A, Bexelius C, ROADMAP consortium. What outcomes are important to patients with mild cognitive impairment or Alzheimer's disease, their caregivers, and health-care professionals? A systematic review. Alzheimers Dement 2019 Dec 09;11(1):231-247 [FREE Full text] [doi: 10.1016/j.dadm.2018.12.003] [Medline: 30906845]
16. 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;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
17. 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;7(3):263-269 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
18. Kozauer N, Katz R. Regulatory innovation and drug development for early-stage Alzheimer's disease. N Engl J Med 2013 Mar 28;368(13):1169-1171. [doi: 10.1056/NEJMp1302513] [Medline: 23484795]
19. Langbaum JB, Hendrix SB, Ayutyanont N, Chen K, Fleisher AS, Shah RC, et al. An empirically derived composite cognitive test score with improved power to track and evaluate treatments for preclinical Alzheimer's disease. Alzheimers Dement 2014 Nov 21;10(6):666-674 [FREE Full text] [doi: 10.1016/j.jalz.2014.02.002] [Medline: 24751827]
20. Papp KV, Rentz DM, Orlovsky I, Sperling RA, Mormino EC. Optimizing the preclinical Alzheimer's cognitive composite with semantic processing: The PACC5. Alzheimers Dement 2017 Nov 09;3(4):668-677 [FREE Full text] [doi: 10.1016/j.trci.2017.10.004] [Medline: 29264389]
21. Papp KV, Rentz DM, Mormino EC, Schultz AP, Amariglio RE, Quiroz Y, et al. Cued memory decline in biomarker-defined preclinical Alzheimer disease. Neurology 2017 Apr 11;88(15):1431-1438 [FREE Full text] [doi: 10.1212/WNL.0000000000003812] [Medline: 28283594]
22. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and Validity of Ambulatory Cognitive Assessments. Assessment 2018 Jan;25(1):14-30 [FREE Full text] [doi: 10.1177/1073191116643164] [Medline: 27084835]
23. Weizenbaum E, Torous J, Fulford D. Cognition in context: understanding the everyday predictors of cognitive performance in a new era of measurement. JMIR Mhealth Uhealth 2020 Jul 23;8(7):e14328. [doi: 10.2196/14328]
24. Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials. Alzheimers Dement 2015 Mar;1(1):103-111 [FREE Full text] [doi: 10.1016/j.dadm.2014.11.003] [Medline: 27239497]
25. Miller J, Barr W. The technology crisis in neuropsychology. Arch Clin Neuropsychol 2017 Aug 01;32(5):541-554. [doi: 10.1093/arclin/acx050] [Medline: 28541383]
26. Buckley RF, Sparks KP, Papp KV, Dekhtyar M, Martin C, Burnham S, et al. Computerized Cognitive Testing for Use in Clinical Trials: a comparison of the NIH Toolbox and Cogstate C3 batteries. J Prev Alzheimers Dis 2017;4(1):3-11 [FREE Full text] [doi: 10.14283/jpad.2017.1] [Medline: 29188853]
27. Öhman F, Hassenstab J, Berron D, Schöll M, Papp KV. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. Alzheimers Dement 2021 Jul 20;13(1):e12217 [FREE Full text] [doi: 10.1002/dad2.12217] [Medline: 34295959]
28. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov;4(6):428-437 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
29. Ross MK, Demos AP, Zulueta J, Piscitello A, Langenecker SA, McInnis M, et al. Naturalistic smartphone keyboard typing reflects processing speed and executive function. Brain Behav 2021 Nov 06;11(11):e2363 [FREE Full text] [doi: 10.1002/brb3.2363] [Medline: 34612605]
30. Heaton R, Miller S, Taylor M, Grant I. Revised Comprehensive Norms for an Expanded Halstead-Reitan Battery: Demographically Adjusted Neuropsychological Norms for African American and Caucasian Adults. Lutz, FL: Psychological Assessment Resources; 2004.
31. 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]
32. 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]
33. Hassenstab J, Aschenbrenner AJ, Balota DA, McDade E, Lim YY, Fagan AM, et al. Remote cognitive assessment approaches in the Dominantly Inherited Alzheimer Network (DIAN). Alzheimers Dement 2020 Dec 07;16(S6). [doi: 10.1002/alz.038144]
34. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020 Feb 17;3(1):21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
35. Insel TR. Digital phenotyping: technology for a new science of behavior. JAMA 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]

XSL•FO
RenderX

36. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
37. Torous J, Kiang MV, Lorme J, Onnela J. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health 2016 May 05;3(2):e16 [FREE Full text] [doi: 10.2196/mental.5165] [Medline: 27150677]
38. Fyffe D, Mukherjee S, Barnes L, Manly J, Bennett D, Crane P. Explaining differences in episodic memory performance among older African Americans and whites: the roles of factors related to cognitive reserve and test bias. J Int Neuropsychol Soc 2011 May 06;17(4):625-638. [doi: 10.1017/s1355617711000476]
39. Byrd DA, Rivera-Mindt MG. Neuropsychology's race problem does not begin or end with demographically adjusted norms. Nat Rev Neurol 2022 Mar;18(3):125-126. [doi: 10.1038/s41582-021-00607-4] [Medline: 35017715]
40. Dodge HH, Zhu J, Mattek NC, Austin D, Kornfeld J, Kaye JA. Use of high-frequency in-home monitoring data may reduce sample sizes needed in clinical trials. PLoS One 2015 Sep 17;10(9):e0138095 [FREE Full text] [doi: 10.1371/journal.pone.0138095] [Medline: 26379170]
41. Onnela J, Arbesman S, González MC, Barabási AL, Christakis NA. Geographic constraints on social network groups. PLoS One 2011 Apr 05;6(4):e16939 [FREE Full text] [doi: 10.1371/journal.pone.0016939] [Medline: 21483665]
42. Onnela J, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology 2016 Dec;41(7):1691-1696 [FREE Full text] [doi: 10.1038/npp.2016.7] [Medline: 26818126]
43. Saeb S, Zhang M, Kwasny MM, Karr CJ, Kording K, Mohr DC. The relationship between clinical, momentary, and sensor-based assessment of depression. Int Conf Pervasive Comput Technol Healthc 2015 Aug;2015 [FREE Full text] [doi: 10.4108/icst.pervasivehealth.2015.259034] [Medline: 26640739]
44. Zulueta J, Piscitello A, Rasic M, Easter R, Babu P, Langenecker SA, et al. Predicting mood disturbance severity with mobile phone keystroke metadata: a BiAffect digital phenotyping study. J Med Internet Res 2018 Dec 20;20(7):e241 [FREE Full text] [doi: 10.2196/jmir.9775] [Medline: 30030209]
45. Vesel C, Rashidisabet H, Zulueta J, Stange JP, Duffecy J, Hussain F, et al. Effects of mood and aging on keystroke dynamics metadata and their diurnal patterns in a large open-science sample: A BiAffect iOS study. J Am Med Inform Assoc 2020 Jul 01;27(7):1007-1018. [doi: 10.1093/jamia/ocaa057] [Medline: 32467973]
46. Faherty LJ, Hantsoo L, Appleby D, Sammel MD, Bennett IM, Wiebe DJ. Movement patterns in women at risk for perinatal depression: use of a mood-monitoring mobile application in pregnancy. J Am Med Inform Assoc 2017 Jul 01;24(4):746-753. [doi: 10.1093/jamia/ocx005] [Medline: 28339686]
47. Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. Psychiatr Rehabil J 2015 Sep;38(3):218-226 [FREE Full text] [doi: 10.1037/prj0000130] [Medline: 25844912]
48. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela J. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology 2018 Jul;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z] [Medline: 29511333]
49. Torous J, Staples P, Barnett I, Sandoval LR, Keshavan M, Onnela J. Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia. NPJ Digit Med 2018 Apr 6;1(1):1-9 [FREE Full text] [doi: 10.1038/s41746-018-0022-8] [Medline: 31304300]
50. Wang R, Aung M, Abdullah S, Brian R, Campbell A, Choudhury T, et al. CrossCheck: toward passive sensing and detection of mental health changes in people with schizophrenia. 2016 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12-16; Heidelberg, Germany p. 886-897. [doi: 10.1145/2971648.2971740]
51. Berry JD, Paganoni S, Carlson K, Burke K, Weber H, Staples P, et al. Design and results of a smartphone-based digital phenotyping study to quantify ALS progression. Ann Clin Transl Neurol 2019 May;6(5):873-881 [FREE Full text] [doi: 10.1002/acn3.770] [Medline: 31139685]
52. Murnane E, Abdullah S, Matthews M, Kay M, Kientz J, Choudhury T, et al. Mobile Manifestations of Alertness: Connecting Biological Rhythms with Patterns of Smartphone App Use. In: Proceedings of the 18th international conference on human-computer interaction with mobile devices and services. USA: ACM; 2016 Sep Presented at: MobileHCI '16; September 2016; Florence, Italy p. 465-477 URL: https://europepmc.org/abstract/MED/30931436 [doi: 10.1145/2935334.2935383]
53. Dagum P. Digital biomarkers of cognitive function. NPJ Digit Med 2018 Mar 28;1(1):10 [FREE Full text] [doi: 10.1038/s41746-018-0018-4] [Medline: 31304295]
54. Gordon M, Gatys L, Guestrin C, Bigham J, Trister A, Patel K. App usage predicts cognitive ability in older adults. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems Internet. USA: ACM; 2019 Presented at: CHI '19; May 2019; Glasgow, UK p. 1-12. [doi: 10.1145/3290605.3300398]
55. Chen R, Jankovic F, Marinsek N, Foschini L, Kourtis L, Signorini A, et al. Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. 2019 Jul Presented at: KDD '19; July 25; Anchorage p. 2145-2155. [doi: 10.1145/3292500.3330690]

XSL•FO
RenderX

56. Onnela J. Opportunities and challenges in the collection and analysis of digital phenotyping data. Neuropsychopharmacology 2021 Jan 17;46(1):45-54 [FREE Full text] [doi: 10.1038/s41386-020-0771-3] [Medline: 32679583]
57. Giovannetti T, Hackett K, Tassoni M, Mis R, Simone S. Assessing and predicting everyday function. In: Gregory B, Bruce C, Kathleen H, Tricia K, editors. APA Handbook of Neuropsychology. Washington, DC: American Psychological Association; Nov 2022.
58. Germine L, Reinecke K, Chaytor NS. Digital neuropsychology: Challenges and opportunities at the intersection of science and software. Clin Neuropsychol 2019 Feb;33(2):271-286. [doi: 10.1080/13854046.2018.1535662] [Medline: 30614374]
59. Tucker-Drob EM. Neurocognitive functions and everyday functions change together in old age. Neuropsychology 2011 May;25(3):368-377 [FREE Full text] [doi: 10.1037/a0022348] [Medline: 21417532]
60. Hultsch D, Strauss E, Hunter M, MacDonald S. Intraindividual variability, cognition, and aging. In: The Handbook of Aging and Cognition. Hove, England: Psychology Press; 2011:491-556.
61. Hultsch DF, MacDonald SWS, Dixon RA. Variability in reaction time performance of younger and older adults. J Gerontol B Psychol Sci Soc Sci 2002 Mar 01;57(2):P101-P115. [doi: 10.1093/geronb/57.2.p101] [Medline: 11867658]
62. Buckner RL. Memory and executive function in aging and AD: multiple factors that cause decline and reserve factors that compensate. Neuron 2004 Sep 30;44(1):195-208 [FREE Full text] [doi: 10.1016/j.neuron.2004.09.006] [Medline: 15450170]
63. Raz N, Lindenberger U, Rodrigue KM, Kennedy KM, Head D, Williamson A, et al. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. Cereb Cortex 2005 Nov;15(11):1676-1689 [FREE Full text] [doi: 10.1093/cercor/bhi044] [Medline: 15703252]
64. 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]
65. 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. Am J Geriatr Psychiatry 2009 May;17(5):368-375 [FREE Full text] [doi: 10.1097/JGP.0b013e31819431d5] [Medline: 19390294]
66. Edmonds EC, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW. Subtle cognitive decline and biomarker staging in preclinical Alzheimer's disease. J Alzheimers Dis 2015;47(1):231-242 [FREE Full text] [doi: 10.3233/JAD-150128] [Medline: 26402771]
67. Lövdén M, Li S, Shing YL, Lindenberger U. Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: Longitudinal data from the Berlin Aging Study. Neuropsychologia 2007;45(12):2827-2838. [doi: 10.1016/j.neuropsychologia.2007.05.005]
68. MacDonald SW, Nyberg L, Bäckman L. Intra-individual variability in behavior: links to brain structure, neurotransmission and neuronal activity. Trends Neurosci 2006 Aug;29(8):474-480. [doi: 10.1016/j.tins.2006.06.011] [Medline: 16820224]
69. Troyer AK, Vandermorris S, Murphy KJ. Intraindividual variability in performance on associative memory tasks is elevated in amnestic mild cognitive impairment. Neuropsychologia 2016 Sep;90:110-116. [doi: 10.1016/j.neuropsychologia.2016.06.011] [Medline: 27297728]
70. MacDonald SWS, Hultsch DF, Dixon RA. Performance variability is related to change in cognition: evidence from the Victoria Longitudinal Study. Psychol Aging 2003 Sep;18(3):510-523. [doi: 10.1037/0882-7974.18.3.510] [Medline: 14518812]
71. Christensen H, Mackinnon AJ, Korten A, Jorm A, Henderson A, Jacomb P. Dispersion in cognitive ability as a function of age: a longitudinal study of an elderly community sample. Neuropsychol Dev Cogn B 2010 Aug 09;6(3):214-228. [doi: 10.1076/anec.6.3.214.779]
72. Holtzer R, Verghese J, Wang C, Hall CB, Lipton RB. Within-person across-neuropsychological test variability and incident dementia. JAMA 2008 Aug 20;300(7):823-830 [FREE Full text] [doi: 10.1001/jama.300.7.823] [Medline: 18714062]
73. Roalf DR, Quarmley M, Mechanic-Hamilton D, Wolk DA, Arnold SE, Moberg PJ. Within-individual variability: an index for subtle change in neurocognition in mild cognitive impairment. J Alzheimers Dis 2016 Aug 10;54(1):325-335 [FREE Full text] [doi: 10.3233/JAD-160259] [Medline: 27567827]
74. Bangen KJ, Weigand AJ, Thomas KR, Delano-Wood L, Clark LR, Eppig J, et al. Cognitive dispersion is a sensitive marker for early neurodegenerative changes and functional decline in nondemented older adults. Neuropsychology 2019 Jul;33(5):599-608 [FREE Full text] [doi: 10.1037/neu0000532] [Medline: 30896235]
75. Farias ST, Mungas D, Reed BR, 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]
76. Jekel K, Damian M, Wattmo C, Hausner L, Bullock R, Connelly PJ, et al. Mild cognitive impairment and deficits in instrumental activities of daily living: a systematic review. Alzheimers Res Ther 2015 Mar 18;7(1):17 [FREE Full text] [doi: 10.1186/s13195-015-0099-0] [Medline: 25815063]
77. 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]

XSL•FO
RenderX

78. Schmitter-Edgecombe M, Parsey CM. Cognitive correlates of functional abilities in individuals with mild cognitive impairment: comparison of questionnaire, direct observation, and performance-based measures. Clin Neuropsychol 2014;28(5):726-746 [FREE Full text] [doi: 10.1080/13854046.2014.911964] [Medline: 24766574]
79. Triebel KL, Martin R, Griffith HR, Marceaux J, Okonkwo OC, Harrell L, et al. Declining financial capacity in mild cognitive impairment: A 1-year longitudinal study. Neurology 2009 Sep 22;73(12):928-934 [FREE Full text] [doi: 10.1212/WNL.0b013e3181b87971] [Medline: 19770468]
80. Fauth EB, Schwartz S, Tschanz JT, Østbye T, Corcoran C, Norton MC. Baseline disability in activities of daily living predicts dementia risk even after controlling for baseline global cognitive ability and depressive symptoms. Int J Geriatr Psychiatry 2013 Jun 12;28(6):597-606 [FREE Full text] [doi: 10.1002/gps.3865] [Medline: 22968965]
81. Luck T, Luppa M, Angermeyer MC, Villringer A, König H, Riedel-Heller SG. Impact of impairment in instrumental activities of daily living and mild cognitive impairment on time to incident dementia: results of the Leipzig Longitudinal Study of the Aged. Psychol Med 2011 May;41(5):1087-1097. [doi: 10.1017/S003329171000142X] [Medline: 20667169]
82. 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;25(4):359-365. [doi: 10.1159/000121005] [Medline: 18340108]
83. Dodge HH, Mattek NC, Austin D, Hayes TL, Kaye JA. 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]
84. Akl A, Taati B, Mihailidis A. Autonomous unobtrusive detection of mild cognitive impairment in older adults. IEEE Trans Biomed Eng 2015 May;62(5):1383-1394. [doi: 10.1109/tbme.2015.2389149]
85. Gillain S, Dramé M, Lekeu F, Wojtasik V, Ricour C, Croisier J, et al. Gait speed or gait variability, which one to use as a marker of risk to develop Alzheimer disease? A pilot study. Aging Clin Exp Res 2016 Apr;28(2):249-255. [doi: 10.1007/s40520-015-0392-6] [Medline: 26076908]
86. Hayes TL, Abendroth F, Adami A, Pavel M, Zitzelberger TA, Kaye JA. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. Alzheimers Dement 2008 Nov;4(6):395-405 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.004] [Medline: 19012864]
87. Montero-Odasso M, Casas A, Hansen KT, Bilski P, Gutmanis I, Wells JL, et al. Quantitative gait analysis under dual-task in older people with mild cognitive impairment: a reliability study. J Neuroeng Rehabil 2009 Sep 21;6(1):35 [FREE Full text] [doi: 10.1186/1743-0003-6-35] [Medline: 19772593]
88. Verghese J, Robbins M, Holtzer R, Zimmerman M, Wang C, Xue X, et al. Gait dysfunction in mild cognitive impairment syndromes. J Am Geriatr Soc 2008 Jul;56(7):1244-1251 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.01758.x] [Medline: 18482293]
89. Verghese J, Wang C, Lipton RB, Holtzer R, Xue X. Quantitative gait dysfunction and risk of cognitive decline and dementia. J Neurol Neurosurg Psychiatry 2007 Sep;78(9):929-935 [FREE Full text] [doi: 10.1136/jnnp.2006.106914] [Medline: 17237140]
90. Benson RR, Guttmann CRG, Wei X, Warfield SK, Hall C, Schmidt JA, et al. Older people with impaired mobility have specific loci of periventricular abnormality on MRI. Neurology 2002 Jan 08;58(1):48-55. [doi: 10.1212/wnl.58.1.48] [Medline: 11781405]
91. Onen F, Henry-Feugeas M, Roy C, Baron G, Ravaud P. Mobility decline of unknown origin in mild cognitive impairment: an MRI-based clinical study of the pathogenesis. Brain Res 2008 Jul 30;1222:79-86. [doi: 10.1016/j.brainres.2008.05.027] [Medline: 18579121]
92. Silbert LC, Nelson C, Howieson DB, Moore MM, Kaye JA. Impact of white matter hyperintensity volume progression on rate of cognitive and motor decline. Neurology 2008 Jul 08;71(2):108-113 [FREE Full text] [doi: 10.1212/01.wnl.0000316799.86917.37] [Medline: 18606964]
93. Austin J, Klein K, Mattek N, Kaye J. Variability in medication taking is associated with cognitive performance in nondemented older adults. Alzheimers Dement (Amst) 2017;6:210-213 [FREE Full text] [doi: 10.1016/j.dadm.2017.02.003] [Medline: 28349120]
94. Seelye A, Hagler S, Mattek N, Howieson DB, Wild K, Dodge HH, et al. Computer mouse movement patterns: A potential marker of mild cognitive impairment. Alzheimers Dement (Amst) 2015 Dec 01;1(4):472-480 [FREE Full text] [doi: 10.1016/j.dadm.2015.09.006] [Medline: 26878035]
95. Reckess G, Varvaris M, Gordon B, Schretlen D. Within-person distributions of neuropsychological test scores as a function of dementia severity. Neuropsychology 2014 Mar;28(2):254-260. [doi: 10.1037/neu0000017] [Medline: 24219602]
96. Hultsch DF, MacDonald SWS, Hunter MA, Levy-Bencheton J, Strauss E. Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. Neuropsychology 2000 Oct;14(4):588-598. [doi: 10.1037/0894-4105.14.4.588]
97. Giebel CM, Sutcliffe C, Challis D. Activities of daily living and quality of life across different stages of dementia: a UK study. Aging Ment Health 2015 Jan;19(1):63-71. [doi: 10.1080/13607863.2014.915920] [Medline: 24831511]
98. Fitz A, Teri L. Depression, cognition, and functional ability in patients with Alzheimer's disease. J Am Geriatr Soc 1994 Feb;42(2):186-191. [doi: 10.1111/j.1532-5415.1994.tb04950.x] [Medline: 8126334]

XSL•FO
RenderX

99. Moon H, Rote S, Haley WE. Factors that contribute to remaining in the community among older adults. Aging Ment Health 2018 Nov;22(11):1502-1509. [doi: 10.1080/13607863.2017.1363869] [Medline: 28799417]
100. Sajjad A, Freak-Poli RL, Hofman A, Roza SJ, Ikram MA, Tiemeier H. Subjective measures of health and all-cause mortality – the Rotterdam Study. Psychol Med 2017 Mar 13;47(11):1971-1980. [doi: 10.1017/s0033291717000381]
101. Lin C, Shih P, Ku LE. Activities of daily living function and neuropsychiatric symptoms of people with dementia and caregiver burden: The mediating role of caregiving hours. Arch Gerontol Geriatr 2019 Mar;81:25-30. [doi: 10.1016/j.archger.2018.11.009] [Medline: 30496871]
102. 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]
103. Kessler R, Giovannetti T, MacMullen L. 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]
104. Rusted J, Sheppard L. Action-based memory in Alzheimer's disease: a longitudinal look at tea making. Neurocase 2002;8(1):111-126. [doi: 10.1076/neur.8.1.111.16240]
105. Schmitter-Edgecombe M, Parsey CM. Assessment of functional change and cognitive correlates in the progression from healthy cognitive aging to dementia. Neuropsychology 2014 Nov;28(6):881-893 [FREE Full text] [doi: 10.1037/neu0000109] [Medline: 24933485]
106. Suzuki T, Murase S, Tanaka T, Okazawa T. New approach for the early detection of dementia by recording in-house activities. Telemed J E Health 2007 Feb;13(1):41-44. [doi: 10.1089/tmj.2006.0033] [Medline: 17309353]
107. Mattek N, Thomas NW, Sharma N, Beattie Z, Marcoe J, Riley T, et al. TD‐P‐17: home‐based digital activity biomarkers remotely monitor relevant activities of MCI and Alzheimer's disease patients and their care partners. Alzheimer's Dement 2019 Jul;15(7S_Part_3):160. [doi: 10.1016/j.jalz.2019.06.4328]
108. Herlihy M, Wing J. Specifying graceful degradation. IEEE Trans. Parallel Distrib. Syst 1991 Jan;2(1):93-104. [doi: 10.1109/71.80192]
109. Giovannetti T, Mis R, Hackett K, Simone S, Ungrady M. 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]
110. Li S, Lindenberger U, Sikström S. Aging cognition: from neuromodulation to representation. Trends Cogn Sci 2001 Nov;5(11):479-486. [doi: 10.1016/s1364-6613(00)01769-1]
111. Li S, Brehmer Y, Shing YL, Werkle-Bergner M, Lindenberger U. Neuromodulation of associative and organizational plasticity across the life span: empirical evidence and neurocomputational modeling. Neurosci Biobehav Rev 2006 Jan;30(6):775-790. [doi: 10.1016/j.neubiorev.2006.06.004] [Medline: 16930705]
112. MacDonald SW, Cervenka S, Farde L, Nyberg L, Bäckman L. Extrastriatal dopamine D2 receptor binding modulates intraindividual variability in episodic recognition and executive functioning. Neuropsychologia 2009 Sep;47(11):2299-2304. [doi: 10.1016/j.neuropsychologia.2009.01.016] [Medline: 19524093]
113. Gleason CE, Norton D, Anderson ED, Wahoske M, Washington DT, Umucu E, et al. Cognitive variability predicts incident Alzheimer's disease and mild cognitive impairment comparable to a cerebrospinal fluid biomarker. J Alzheimers Dis 2018;61(1):79-89 [FREE Full text] [doi: 10.3233/JAD-170498] [Medline: 29125485]
114. Kelly AMC, Uddin LQ, Biswal BB, Castellanos FX, Milham MP. Competition between functional brain networks mediates behavioral variability. Neuroimage 2008 Jan 01;39(1):527-537. [doi: 10.1016/j.neuroimage.2007.08.008] [Medline: 17919929]
115. West R, Murphy KJ, Armilio ML, Craik FIM, Stuss DT. Lapses of intention and performance variability reveal age-related increases in fluctuations of executive control. Brain Cogn 2002 Aug;49(3):402-419. [doi: 10.1006/brcg.2001.1507] [Medline: 12139961]
116. 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]
117. Bassuk SS, Glass TA, Berkman LF. Social disengagement and incident cognitive decline in community-dwelling elderly persons. Ann Intern Med 1999 Aug 03;131(3):165-173. [doi: 10.7326/0003-4819-131-3-199908030-00002] [Medline: 10428732]
118. Dickinson WJ, Potter GG, Hybels CF, McQuoid DR, Steffens DC. Change in stress and social support as predictors of cognitive decline in older adults with and without depression. Int J Geriatr Psychiatry 2011 Dec 02;26(12):1267-1274 [FREE Full text] [doi: 10.1002/gps.2676] [Medline: 21370277]
119. Kaye J, Mattek N, Dodge HH, Campbell I, Hayes T, Austin D, et al. Unobtrusive measurement of daily computer use to detect mild cognitive impairment. Alzheimers Dement 2014 Jan;10(1):10-17 [FREE Full text] [doi: 10.1016/j.jalz.2013.01.011] [Medline: 23688576]
120. Bernstein JPK, Dorociak K, Mattek N, Leese M, Trapp C, Beattie Z, et al. Unobtrusive, in-home assessment of older adults' everyday activities and health events: associations with cognitive performance over a brief observation period. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2021 Apr 18:1-18. [doi: 10.1080/13825585.2021.1917503] [Medline: 33866939]

XSL•FO
RenderX

121. Chodosh J, Kado DM, Seeman TE, Karlamangla AS. Depressive symptoms as a predictor of cognitive decline: MacArthur studies of successful aging. Am J Geriatr Psychiatry 2007 May;15(5):406-415. [doi: 10.1097/01.jgp.0b013e31802c0c63]
122. Gallaway P, Miyake H, Buchowski M, Shimada M, Yoshitake Y, Kim A, et al. Physical activity: a viable way to reduce the risks of mild cognitive impairment, Alzheimer's disease, and vascular dementia in older adults. Brain Sci 2017 Feb 20;7(2):22 [FREE Full text] [doi: 10.3390/brainsci7020022] [Medline: 28230730]
123. Hilborn JV, Strauss E, Hultsch DF, Hunter MA. Intraindividual variability across cognitive domains: investigation of dispersion levels and performance profiles in older adults. J Clin Exp Neuropsychol 2009 May;31(4):412-424. [doi: 10.1080/13803390802232659] [Medline: 18720183]
124. Vaidyam A, Halamka J, Torous J. Actionable digital phenotyping: a framework for the delivery of just-in-time and longitudinal interventions in clinical healthcare. Mhealth 2019;5:25 [FREE Full text] [doi: 10.21037/mhealth.2019.07.04] [Medline: 31559270]
125. Mobile fact sheet internet. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2022-04-01]
126. Barnett I, Torous J, Staples P, Keshavan M, Onnela J. Beyond smartphones and sensors: choosing appropriate statistical methods for the analysis of longitudinal data. J Am Med Inform Assoc 2018 Dec 01;25(12):1669-1674. [doi: 10.1093/jamia/ocy121] [Medline: 30272176]
127. Barnett I, Torous J, Reeder H, Baker J, Onnela JP. Determining sample size and length of follow-up for smartphone-based digital phenotyping studies. J Am Med Inform Assoc 2020 Dec 09;27(12):1844-1849 [FREE Full text] [doi: 10.1093/jamia/ocaa201] [Medline: 33043370]
128. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/ wearable devices opportunity. NPJ Digit Med 2019;2 [FREE Full text] [doi: 10.1038/s41746-019-0084-2] [Medline: 31119198]
129. Taylor KI, Staunton H, Lipsmeier F, Nobbs D, Lindemann M. Outcome measures based on digital health technology sensor data: data- and patient-centric approaches. NPJ Digit Med 2020 Jul 23;3(1):97 [FREE Full text] [doi: 10.1038/s41746-020-0305-8] [Medline: 32715091]
130. Teipel S, König A, Hoey J, Kaye J, Krüger F, Robillard JM, et al. Use of nonintrusive sensor-based information and communication technology for real-world evidence for clinical trials in dementia. Alzheimers Dement 2018 Sep;14(9):1216-1231 [FREE Full text] [doi: 10.1016/j.jalz.2018.05.003] [Medline: 29936147]
131. Real-world data and real-world evidence supporting clinical decision making. Professional Society for Health Economics and Outcomes Research. URL: https://www.ispor.org/images/default-source/strategic-initiative/pfizer-bms-ispor-infographic_finalf1e4521f586d4354b69ff9700bd2fb93.jpg?sfvrsn=508911d8_2 [accessed 2022-04-01]
132. Bartlett Ellis R, Wright J, Miller LS, Jake-Schoffman D, Hekler EB, Goldstein CM, et al. Lessons learned: beta-testing the Digital Health Checklist for Researchers prompts a call to action by behavioral scientists. J Med Internet Res 2021 Dec 22;23(12):e25414 [FREE Full text] [doi: 10.2196/25414] [Medline: 34941548]
133. Policy for device software functions and mobile medical applications. Guidance for industry and Food and Drug Administration staff. US Food and Drug Administration. 2019 Sep. URL: https://www.fda.gov/media/80958/download [accessed 2022-04-01]
134. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med 2020 Oct 08;10(4):1004-1015 [FREE Full text] [doi: 10.1093/tbm/ibz074] [Medline: 31120511]

Abbreviations

AD: Alzheimer’s disease
ADRD: Alzheimer’s disease and related dementias
CSF: cerebrospinal fluid
MCI: mild cognitive impairment
NIA: National Institute on Aging
NIH: National Institutes of Health
VIBE: Variability in Everyday Behavior

XSL•FO
RenderX

Edited by T Leung; submitted 21.03.22; peer-reviewed by E Weizenbaum, G Jones; comments to author 29.06.22; revised version received 19.07.22; accepted 31.07.22; published 07.09.22.

Please cite as:
Hackett K, Giovannetti T
Capturing Cognitive Aging in Vivo: Application of a Neuropsychological Framework for Emerging Digital Tools
JMIR Aging 2022;5(3):e38130
URL: https://aging.jmir.org/2022/3/e38130
doi:10.2196/38130
PMID:36069747

©Katherine Hackett, Tania Giovannetti. Originally published in JMIR Aging (https://aging.jmir.org), 07.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Web-Based Delivery of the Caregiving Essentials Course for Informal Caregivers of Older Adults in Ontario: Mixed Methods Evaluation Study

Shelley Rottenberg[1], BSc, MA; Allison Williams[1], BA, MA, PhD

School of Earth, Environment & Society, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Shelley Rottenberg, BSc, MA
School of Earth, Environment & Society
McMaster University
1280 Main Street West
Hamilton, ON, L8S 4L8
Canada
Phone: 1 905 525 9140
Email: rottensm@mcmaster.ca

Related Article:

Correction of: https://aging.jmir.org/2021/2/e25671

(JMIR Aging 2022;5(3):e42215) doi:10.2196/42215

In "Web-Based Delivery of the Caregiving Essentials Course for Informal Caregivers of Older Adults in Ontario: Mixed Methods Evaluation Study" (JMIR Aging 2021;4(2):e25671), the authors made a change in the corresponding author's contact information.

In the corrected version, the corresponding author's phone number has been updated to "1 905 525 9140."

The correction will appear in the online version of the paper on the JMIR Publications website on August 31, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 26.08.22; this is a non–peer-reviewed article;accepted 29.08.22; published 31.08.22.

Please cite as:
Rottenberg S, Williams A
Correction: Web-Based Delivery of the Caregiving Essentials Course for Informal Caregivers of Older Adults in Ontario: Mixed Methods Evaluation Study
JMIR Aging 2022;5(3):e42215
URL: https://aging.jmir.org/2022/3/e42215
doi:10.2196/42215
PMID:36044752

©Shelley Rottenberg, Allison Williams. Originally published in JMIR Aging (https://aging.jmir.org), 31.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: The Effect of Cognitive Function Health Care Using Artificial Intelligence Robots for Older Adults: Systematic Review and Meta-analysis

Hocheol Lee[1,2], PhD; Min Ah Chung[1,2], MPH; Hyeji Kim[1,2], BA; Eun Woo Nam[1,2,3,4], MPH, PhD

[1]Healthy City Research Center, Yonsei University, Wonju, Republic of Korea

[2]Yonsei Global Health Center, Yonsei University, Wonju, Republic of Korea

[3]Center of Evidence Based Medicine, Institute of Convergence Science, Yonsei University, Wonju, Republic of Korea

[4]Department of Health Administration, College of Software and Digital Healthcare Convergence, Yonsei University, Wonju, Republic of Korea

Corresponding Author:
Eun Woo Nam, MPH, PhD
Department of Health Administration
College of Software and Digital Healthcare Convergence
Yonsei University
Unit 412, Chang-jo gwan
1 Yonseidae-gil
Wonju, 26493
Republic of Korea
Phone: 82 33 760 2413
Fax: 82 33 762 9562
Email: ewnam@yonsei.ac.kr

Related Article:

Correction of: https://aging.jmir.org/2022/2/e38896

(JMIR Aging 2022;5(3):e42312) doi:10.2196/42312

In "The Effect of Cognitive Function Health Care Using Artificial Intelligence Robots for Older Adults: Systematic Review and Meta-analysis" (JMIR Aging 2022;5(2):e38896), the authors made one addition.

The following information was added to the *Acknowledgments* section:

> *This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1C1C2005464).*

The correction will appear in the online version of the paper on the JMIR Publications website on September 14, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 06.09.22; this is a non–peer-reviewed article;accepted 06.09.22; published 14.09.22.

Please cite as:
Lee H, Chung MA, Kim H, Nam EW
Correction: The Effect of Cognitive Function Health Care Using Artificial Intelligence Robots for Older Adults: Systematic Review and Meta-analysis
JMIR Aging 2022;5(3):e42312
URL: https://aging.jmir.org/2022/3/e42312
doi:10.2196/42312
PMID:36103706

XSL•FO
RenderX

©Hocheol Lee, Min Ah Chung, Hyeji Kim, Eun Woo Nam. Originally published in JMIR Aging (https://aging.jmir.org), 14.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: A Web-Based Platform (CareVirtue) to Support Caregivers of People Living With Alzheimer Disease and Related Dementias: Mixed Methods Feasibility Study

Justin J Boutilier[1], PhD; Priya Loganathar[1], MS; Anna Linden[1], MSc; Eleanore Scheer[2], BSc; Sofia Noejovich[1], BSc; Christian Elliott[3], BSc; Matthew Zuraw[3], MBA; Nicole E Werner[4], PhD

[1]Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, United States

[2]Department of Engineering Systems and Environment, University of Virginia, Charlottesville, VA, United States

[3]Whiplash Technology, Inc, Palm Springs, CA, United States

[4]Department of Health and Wellness Design, Indiana University School of Public Health-Bloomington, Bloomington, IN, United States

Corresponding Author:
Nicole E Werner, PhD
Department of Health and Wellness Design
Indiana University School of Public Health-Bloomington
1025 E 7th St
Bloomington, IN, 47405
United States
Phone: 1 703 340 7177
Email: newerner@iu.edu

Related Article:

Correction of: https://aging.jmir.org/2022/3/e36975

(JMIR Aging 2022;5(3):e41912) doi:10.2196/41912

In "Web-Based Platform (CareVirtue) to Support Caregivers of People Living With Alzheimer Disease and Related Dementias: Mixed Methods Feasibility Study" (JMIR Aging 2022;5(3):e36975), the authors noted an error in the attribution of affiliations. Authors Justin J Boutilier, Priya Loganathar, Anna Linden, and Sofia Noejovich had the following incorrect affiliation attributed to their names:

> *Department of Health and Wellness Design, Indiana University School of Public Health-Bloomington, Bloomington, IN, United States*

The correct affiliation for these authors, which did not appear in the original article, is as follows:

> *Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, US*

This affiliation appears in the corrected article as affiliation 1 and the remaining affiliations have been reordered accordingly.

The full list of authors and affiliations in the originally published article was as follows:

> *Justin J Boutilier[1], PhD; Priya Loganathar[1], MS; Anna Linden[1], MSc; Eleanore Scheer[2], BSc; Sofia Noejovich[1], BSc; Christian Elliott[3], BSc; Matthew Zuraw [3], MBA; Nicole E Werner[1], PhD*
>
> *[1]Department of Health and Wellness Design, Indiana University School of Public Health-Bloomington, Bloomington, IN, United States*
>
> *[2]Department of Engineering Systems and Environment, University of Virginia, Charlottesville, VA, United States*
>
> *[3]Whiplash Technology, Inc, Palm Springs, CA, US*

The full list of authors and affiliations in the corrected version of the article appears as follows:

> *Justin J Boutilier[1], PhD; Priya Loganathar[1], MS; Anna Linden[1], MSc; Eleanore Scheer[2], BSc; Sofia Noejovich[1], BSc; Christian Elliott[3], BSc; Matthew Zuraw[3], MBA; Nicole E Werner[4], PhD*
>
> *[1]Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, WI, US*
>
> *[2]Department of Engineering Systems and Environment, University of Virginia, Charlottesville, VA, US*
>
> *[3]Whiplash Technology, Inc., Palm Springs, CA, US*
>
> *[4]Department of Health and Wellness Design, Indiana University School of Public Health-Bloomington, Bloomington, IN, US*

XSL•FO
RenderX

The correction will appear in the online version of the paper on the JMIR Publications website on September 14, 2022, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 14.08.22; this is a non–peer-reviewed article;accepted 18.08.22; published 14.09.22.

Please cite as:
Boutilier JJ, Loganathar P, Linden A, Scheer E, Noejovich S, Elliott C, Zuraw M, Werner NE
Correction: A Web-Based Platform (CareVirtue) to Support Caregivers of People Living With Alzheimer Disease and Related Dementias: Mixed Methods Feasibility Study
JMIR Aging 2022;5(3):e41912
URL: https://aging.jmir.org/2022/3/e41912
doi:10.2196/41912
PMID:36103711

©Justin J Boutilier, Priya Loganathar, Anna Linden, Eleanore Scheer, Sofia Noejovich, Christian Elliott, Matthew Zuraw, Nicole E Werner. Originally published in JMIR Aging (https://aging.jmir.org), 14.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Shared Decision-Making Training for Home Care Teams to Engage Frail Older Adults and Caregivers in Housing Decisions: Stepped-Wedge Cluster Randomized Trial

Évèhouénou Lionel Adisso[1,2,3], MSc; Monica Taljaard[4,5], PhD; Dawn Stacey[4,6], PhD; Nathalie Brière[7], PhD; Hervé Tchala Vignon Zomahoun[1,2,3,8,9], PhD; Pierre Jacob Durand[2,3,7], PhD; Louis-Paul Rivest[10,11], PhD; France Légaré[1,2,8,12], MD, PhD

[1]Tier 1 Canada Research Chair in Shared Decision Making and Knowledge Translation, Québec, QC, Canada

[2]VITAM – Centre de recherche en santé durable, Quebec, QC, Canada

[3]Department of Social and Preventive Medicine, Faculty of Medicine, Université Laval, Quebec, QC, Canada

[4]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

[5]School of Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

[6]School of Nursing, University of Ottawa, Ottawa, ON, Canada

[7]Direction des services multidisciplinaires, Centre intégré universitaire de santé et de services sociaux de la Capitale-Nationale, Quebec, QC, Canada

[8]Knowledge Translation and Implementation Component of the Quebec Strategy for Patient-Oriented Research - Support for People and Patient-Oriented Research and Trials Unit, Health and Social Services Systems, Quebec, QC, Canada

[9]School of Physical and Occupational Therapy, Faculty of Medicine, McGill University, Montreal, QC, Canada

[10]Department of Mathematics and Statistics, Université Laval, Quebec, QC, Canada

[11]Canada Research Chair in Statistical Sampling and Data Analysis, Laval University, Quebec, QC, Canada

[12]Department of Family Medicine and Emergency Medicine, Faculty of Medicine, Université Laval, Québec, QC, Canada

Corresponding Author:
France Légaré, MD, PhD
Tier 1 Canada Research Chair in Shared Decision Making and Knowledge Translation
2525 Chemin de la Canardière bureau A-3421
Québec, QC, G1J0A4
Canada
Phone: 1 4188219481
Email: france.legare@fmed.ulaval.ca

Abstract

Background: Frail older adults and caregivers need support from their home care teams in making difficult housing decisions, such as whether to remain at home, with or without assistance, or move into residential care. However, home care teams are often understaffed and busy, and shared decision-making training is costly. Nevertheless, overall awareness of shared decision-making is increasing. We hypothesized that distributing a decision aid could be sufficient for providing decision support without the addition of shared decision-making training for home care teams.

Objective: We evaluated the effectiveness of adding web-based training and workshops for care teams in interprofessional shared decision-making to passive dissemination of a decision guide on the proportion of frail older adults or caregivers of cognitively-impaired frail older adults reporting active roles in housing decision-making.

Methods: We conducted a stepped-wedge cluster randomized trial with home care teams in 9 health centers in Quebec, Canada. Participants were frail older adults or caregivers of cognitively impaired frail older adults facing housing decisions and receiving care from the home care team at one of the participating health centers. The intervention consisted of a 1.5-hour web-based tutorial for the home care teams plus a 3.5-hour interactive workshop in interprofessional shared decision-making using a decision guide that was designed to support frail older adults and caregivers in making housing decisions. The control was passive dissemination of the decision guide. The primary outcome was an active role in decision-making among frail older adults and caregivers, measured using the Control Preferences Scale. Secondary outcomes included decisional conflict and perceptions of how much care teams involved frail older adults and caregivers in decision-making. We performed an intention-to-treat analysis.

XSL•FO
RenderX

Results: A total of 311 frail older adults were included in the analysis, including 208 (66.9%) women, with a mean age of 81.2 (SD 7.5) years. Among 339 caregivers of cognitively-impaired frail older adults, 239 (70.5%) were female and their mean age was 66.4 (SD 11.7) years. The intervention increased the proportion of frail older adults reporting an active role in decision-making by 3.3% (95% CI –5.8% to 12.4%, *P*=.47) and the proportion of caregivers of cognitively-impaired frail older adults by 6.1% (95% CI -11.2% to 23.4%, *P*=.49). There was no significant impact on the secondary outcomes. However, the mean score for the frail older adults' perception of how much health professionals involved them in decision-making increased by 5.4 (95% CI −0.6 to 11.4, *P*=.07) and the proportion of caregivers who reported decisional conflict decreased by 7.5% (95% CI −16.5% to 1.6%, *P*=.10).

Conclusions: Although it slightly reduced decisional conflict for caregivers, shared decision-making training did not equip home care teams significantly better than provision of a decision aid for involving frail older adults and their caregivers in decision-making.

Trial Registration: ClinicalTrials.gov NCT02592525; https://clinicaltrials.gov/show/NCT02592525

(JMIR Aging 2022;5(3):e39386) doi:10.2196/39386

KEYWORDS

shared decision-making; home care; nursing homes; patient engagement

Introduction

Aging is associated with a higher risk of developing disabilities that can lead to loss of autonomy [1,2]. When frail older adults start to lose autonomy, one of the most difficult decisions they face is whether to remain at home, with or without assistance, or move into residential care [3]. When these older adults have cognitive impairment, caregivers may have to make the decision instead, often with little support [4]. Making this difficult decision [5] can lead to stress, decisional conflict, and regret [6].

Shared decision-making (SDM) is a process whereby health professional, patients, and their caregivers work together to make health care choices based on the best evidence and what matters most to patients [7]. SDM tools, such as decision guides, are associated with better decision quality and decision-making processes without damaging patient or health system outcomes [8]. Decision guides can increase the involvement of frail older adults and caregivers in decisions about their care while improving agreement between them and their home care teams [9].

In previous work, an interprofessional SDM (IP-SDM) training program for home care teams with a decision guide increased by 12% (compared to usual care) the proportion of caregivers who reported being active in making housing decisions for frail older adults with cognitive impairment [10]. However, other studies have shown that educational interventions may make little difference to the actual practice of SDM with older adults with cognitive impairment and their surrogate decision makers [11]. In addition, given that home care teams are already very busy, and overall awareness of SDM is increasing [12], passive dissemination of decision guides alone could be enough to increase patient engagement in decision-making [13]. However, the effectiveness of decision guides alone, compared to their use as part of a multifaceted intervention, is unknown.

The aim of this study was to evaluate the effect of adding a blended web-based and in-person training program in IP-SDM for home care teams to passive dissemination of a decision guide on the proportion of frail older adults or caregivers reporting an active role in making housing decisions, compared with passive dissemination of the decision guide alone. We hypothesized that the addition of a training program in IP-SDM to the passive dissemination of a decision guide would increase the proportion of frail older adults or caregivers reporting an active role in the decision-making process.

Methods

Ethics Approval

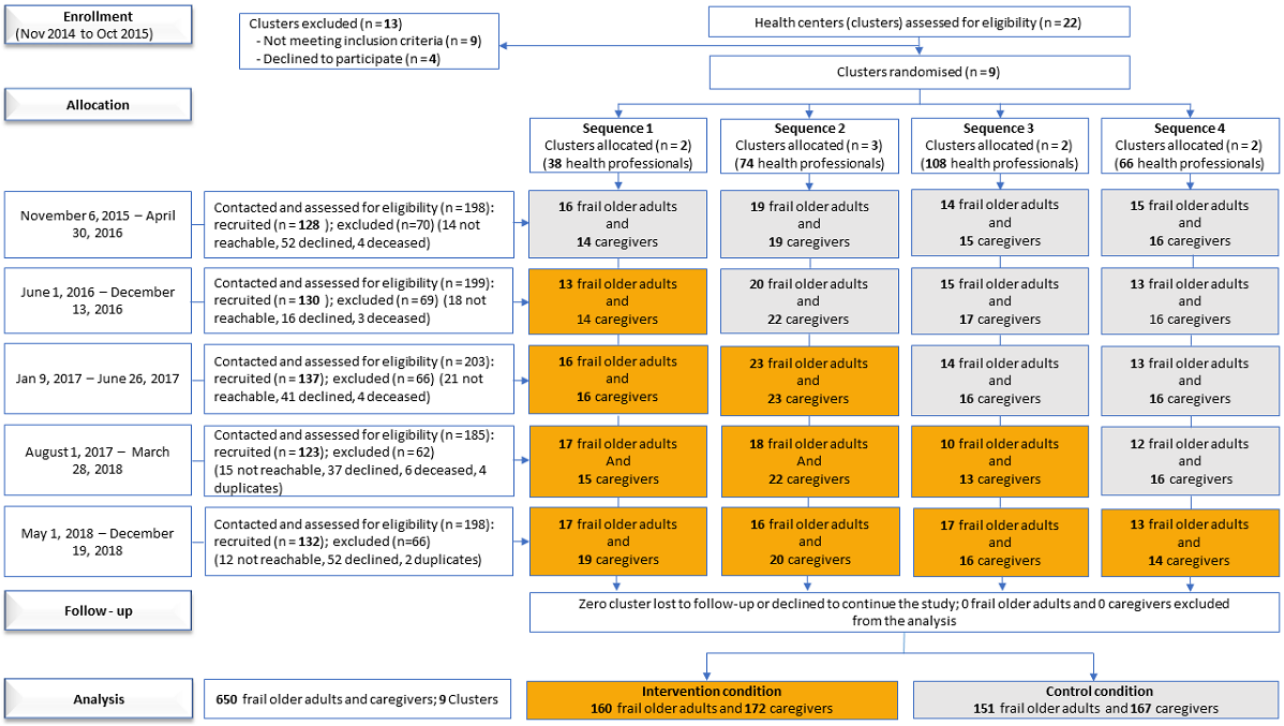
We reported this trial following the extension of Consolidated Standards of Reporting Trials (CONSORT) for stepped-wedge cluster randomized trials [14]. The trial was registered at ClinicalTrials.gov (NCT02592525) and the protocol was published [15]. Ethics committee review approval has been obtained from the Multicenter Ethics Committee of Centre intégré de santé et de services sociaux de Laval (2015-2016/01-01-E).

Study Design

We conducted a cross-sectional, stepped-wedge cluster randomized trial (the Inter Professional Shared Decision-Making-Stepped Wedge Study) from November 2014 to December 2018 with home care teams at health centers in Quebec, Canada. We chose cluster randomization because the intervention was delivered at the health-center level, precluding individual randomization. A stepped-wedge design was chosen to facilitate recruitment, as all health centers would ultimately receive the intervention [16]. This design also offers more statistical power than a traditional parallel cluster study when there are large cluster-level effects (or intracluster correlations) [17]. Health centers were randomly allocated to 1 of 4 intervention start times (sequences), with 5 data collection periods (Figure 1).

XSL•FO
RenderX

Figure 1. Flow chart of trial by allocated sequence and period_updated.



Participants and Eligibility

Study participants were frail older adults with loss of autonomy and caregivers of frail older adults with cognitive impairment who were recruited through the home care teams at the health centers. Home care teams were eligible if they (1) were involved in caring for frail older adults, (2) practiced in one of the health centers participating in the trial, and (3) were interprofessional (ie, involved more than 2 health professionals from different professions). Frail older adults were eligible if they (1) were aged ≥65 years; (2) were receiving care from one of the home care teams; (3) had made a decision about staying home or moving during the recruitment period; (4) were able to read, understand, and write French or English; and (5) were able to give informed consent. When frail older adults were cognitively impaired, their informal caregiver became the eligible participant. Caregivers were defined in this study as close relatives or friends and were eligible if they (1) were caring for a cognitively impaired older adult who was otherwise eligible; (2) were able to read, understand, and write French or English; and (3) provided informed consent to participate in the study. Frail older adults with cognitive impairment had been clinically evaluated by a health professional as no longer able to make decisions on their own.

Randomization

Health centers (clusters) were randomized to 1 of 4 sequences. Once participating home care teams had been identified, an independent biostatistician at the Ottawa Hospital Research Institute's Methods Centre performed randomization using computer-generated numbers. Given the nature of the intervention, the investigators, project coordinator, and research assistants (RAs) collecting the data were not blinded. However, the allocation list was concealed from the research team for as long as possible and the RAs were asked not to discuss this information with any frail older adult or caregiver and not to refer to the intervention. Frail older adults and caregivers were blinded to the intervention.

Control

Before baseline data collection, we asked managers at all the enrolled health centers to distribute (ie, perform passive dissemination of) a decision guide for home care teams supporting frail older adults or caregivers in making housing decisions [4]. Dissemination of the decision guide was passive in the sense that although it was distributed in the health centers, we did not train the teams in how to use it. The decision guide, adapted from an online family decision support tool designed for the context of the home, had French and English versions [4,18]. It has the potential to help health professionals discuss with frail older adults or caregivers of cognitively impaired frail older adults the decision about the location of care [4,9,13].

Intervention

The intervention consisted of (1) a 1.5-hour web-based tutorial, based on the Ottawa Decision Support Tutorial, [19] that was completed individually by the health professionals in the participating home care teams at the cluster level, followed by (2) a 3.5-hour live interactive workshop. The web-based tutorial ensured that all participants arrived at the workshop with a similar knowledge of SDM concepts. The workshop included a lecture reviewing SDM concepts (especially the IP-SDM approach); a video demonstrating the approach in a home care team with a frail older adult making a housing decision [20]; training in using the decision guide [4]; and role play using the decision guide with feedback from facilitators [15,20]. The workshop, based on adult education principles [21], included decision-making about housing decisions with frail older adults, communication techniques, and, for frail older adults with cognitive impairment, strategies for fostering their participation

XSL•FO
RenderX

or that of their caregivers in decision-making. All workshops were held at the health centers, had the same content, same materials, and same trainers, and were held as a single session [15]. All home care teams received the intervention at various time points. The decision guide distributed before the intervention was still available in sufficient quantities afterwards [15]. The digital format of the initial tutorial and the video were convenient and easily scalable to our 9 intervention sites and ensured that base elements of the training were standardized and identical. This is helpful in stepped-wedge trials, where control and intervention conditions are experienced at different times, there is implementation lag, and individuals are exposed to the intervention in different ways and locations. It also reduced time expenditure and costs, in contrast to in-person training, which had to be repeated at each crossover point [22]. However, our intervention overcame the disadvantages of web-based learning (mainly isolation) [23-25] with the in-person part of the training, which provided role play, feedback, and discussion opportunities for applying knowledge to skills and behavior [26].

Outcomes and Measurement

The primary outcome was the frail older adults' or caregivers' perception of the role they assumed in decision-making, as measured using a modified version of the Control Preferences Scale [27], a single question with five response categories: (1) "I made the decision," (2) "I made the decision after seriously considering the health care professionals' opinions," (3) "the health care professionals and I shared the responsibility for the decision making," (4) "the health care professionals made the decision after seriously considering my opinion," and (5) "the health care professionals made the decision." For sample size calculation and analysis, we dichotomized the primary outcome by collapsing categories 1, 2, and 3 into an "active" role and 4 and 5 into a "passive" role in decision-making.

Secondary outcomes assessed in frail older adults and caregivers were (1) their preferred option about whether the cognitively impaired older adult should stay at home or move to another location, and the actual decision made; (2) decisional conflict, assessed with the 16-item Decisional Conflict Scale [28,29]; (3) decision regret, assessed with the 5-item Decision Regret Scale [30]; and (4) perception of the extent to which health professionals involved them in decision-making, assessed with the Dyadic-OPTION scale, a 12-item instrument evaluating SDM behaviors during decision-making [31,32]. Other secondary outcomes included health-related quality of life, assessed only in the frail older adults with the 36 items of the Nottingham Health Profile [33-35], and burden of care, assessed only in the caregivers with the Zarit Burden Inventory scale [36-38].

Data Collection

Home care teams made lists of potentially eligible frail older patients. Trained RAs assigned to each health center contacted these patients or caregivers of frail older adults with cognitive impairment and asked if they wished to participate. The RAs then met all interested participants at their home or a place of their choice to complete informed consent and proceed with data collection. Data collection took place from November 2015 to December 2018. Due to practical constraints, some health centers started the intervention earlier or later than planned. The collected data included outcomes; the relationship between caregivers and frail older adults (when appropriate); and sociodemographic characteristics, including age, sex, and education, which were variables identified as predictors of our primary outcome, that is, that younger, female, well-educated (secondary school level or higher) people are more likely to take an active role in decisions about their health [27,39-41].

Sample Size

The sample size calculation was informed by preliminary data from another study [42]. We used the method developed by Hussey and Hughes [43] for stepped-wedge designs. We assumed an average of 8 frail older adults and 8 caregivers per health center in each data collection period and a time-independent intraclass correlation (ICC) of 0.05 [44]. To detect an absolute increase of 20% [45] in the primary outcome (ie, from 70% to 90%) with 80% power using a stepped-wedge design with 4 sequences and a 2-sided test at the 5% significance level, a total of 8 clusters (with a total of 320 caregivers) were required, [46] meaning 320 frail older adults and 320 caregivers of frail older adults with cognitive impairment. To account for potential loss to follow-up of clusters we recruited one more health center than planned.

Statistical Methods

We describe organizational settings and characteristics of the health professionals randomized to the trial and report the sociodemographics of the frail older adults and caregivers using frequencies and percentages, means and SD, or medians and IQR, as appropriate. We performed analyses with the intention-to-treat principle with the frail older adult or caregiver as the unit of analysis. The primary outcome was analyzed using a generalized linear mixed model (GLMM) with logit link. The prespecified primary analysis assumed a uniform within- and between-period correlation, adjusting for time effects (categorical) and specifying a random effect for cluster [43].

We performed secondary analyses by additionally adjusting for primary outcome predictors and for imbalanced baseline characteristics [47,48]. To explore the implications of bias due to misspecification of the correlation structure [49], we conducted analyses using 2 other correlation structures identified in the literature: nested exchangeable (specifying a random cluster effect and a random time by cluster interaction) [50,51] and exponential decay (an autoregressive between-period correlation) [52]. There are no guidelines for choosing the best-fitting covariance structure, so we used the pseudo–Akaike information criterion to select the best-fitting model and presented the results as sensitivity analyses. To estimate the absolute difference, as required by the CONSORT extension for stepped-wedge cluster randomized trials, [14] we applied GLMM using an identity link with the adaptive Gaussian–Hermite approximation to the likelihood maximum [53].

For binary secondary outcomes, we conducted similar analyses. For continuous secondary outcomes, we used a linear mixed model, and summarized the intervention effects as mean

XSL•FO
RenderX

differences. We obtained within-period intraclass correlation coefficients (WpICC), between-period intraclass correlation coefficients (BpICC), and cluster autocorrelation coefficients (CAC) for each outcome analyzed. We used α=.05 as the level of significance. All analyses were conducted using SAS (version 9.4, SAS Institute).

Results

Participants

Recruitment took place from November 2014 to December 2018. Interprofessional home care teams from 9 health centers with 281 health professionals participated in the study. Of 481 frail older adults contacted, 311 (64.6%) were recruited. Of 502 eligible caregivers contacted, 339 (67.5%) were recruited. There was no loss to follow-up of health centers, and no frail older adults, caregivers, or health centers were excluded (Figure 1). Sociodemographics of the frail older adults and caregivers were well balanced between allocated sequences (Multimedia Appendices 1 and 2).

Baseline Characteristics of Participants

Participating frail older adults were on average 81.2 (SD 7.5) years old; 66.9% (208/311) were female and 58.8% (183/311) had secondary education or higher. Baseline characteristics were well balanced between the intervention and control groups, except for education level (Table 1). Caregivers of frail older adults with cognitive impairment were on average 66.4 (SD 11.7) years old; 70.5% (239/339) were female and 87.3% (296/339) had secondary education or higher. Most caregivers (242/339, 71.4%) were retired or at home and 90.3% (306/339) were the child, spouse, or husband of the frail older adult. Among caregivers, baseline characteristics were well balanced between the intervention and control groups, except for age (Table 2).

Table 1. Baseline characteristics of frail older adults without cognitive impairment (N=311).

Characteristics	Control (n=151)	Intervention (n=160)
Age (years), mean (SD)	81.6 (7.6)	80.9 (7.4)[a]
Sex (female), n (%)	101 (66.9)	107 (66.9)
Education level, n (%)		
Primary school	44 (29.2)	84 (52.5)
Secondary school	73 (48.3)	51 (31.9)
Postsecondary	34 (22.5)	25 (15.6)
Marital status, n (%)		
Married/common-law partner	45 (29.8)	58 (36.3)
Widowed	72 (47.7)	60 (37.5)
Separated/divorced	20 (13.3)	25 (15.6)
Single	14 (9.2)	17 (10.6)
Household income (CAD $)[b], n (%)		
Less than 30,000	83 (55.0)	86 (53.8)
30,000-59,999	34 (22.5)	30 (18.8)
60,000 and more	4 (2.7)	7 (4.4)
I prefer not to answer/I do not know	30 (19.9)	37 (23.1)

[a]n=159

[b]A currency exchange rate of CAD $1=US $0.76 is applicable.

XSL•FO
RenderX

Table 2. Baseline characteristics of caregivers of cognitively-impaired frail older adults (N=339).

Characteristics	Control (n=167)	Intervention (n=172)
Age (years), mean (SD)	64.2 (11.9)	68.6 (11.2)
Sex (female), n (%)	122 (73.1)	117 (68.0)
Education level, n (%)		
Primary school	19 (11.4)	24 (14.0)
Secondary school	63 (37.7)	69 (40.1)
Postsecondary	85 (50.9)	79 (45.9)
Marital status, n (%)		
Married/common-law partner	129 (77.2)	132 (76.7)
Widowed	7 (4.2)	9 (5.2)
Separated/divorced	16 (9.6)	18 (10.5)
Single	15 (9.0)	13 (7.6)
Household income (CAD $)[a], n (%)		
Less than 30,000	37 (22.2)	43 (25.0)
30,000-59,999	54 (32.3)	50 (29.1)
60,000 or more	51 (30.5)	46 (26.7)
I prefer not to answer/I do not know	25 (15.0)	33 (19.2)
Caregivers' employment status, n (%)		
Retired	94 (56.3)	114 (66.3)
Employed	56 (33.5)	39 (22.7)
At home (eg, unemployed/job seeker)	17 (10.2)	17 (9.9)
Missing	0 (0.0)	2 (1.1)
Caregivers' relationship to frail older adult, n (%)		
Child	94 (56.3)	75 (43.6)
Wife/husband or common-law partner	59 (35.3)	78 (45.3)
Other (eg, family member or friend)	14 (8.4)	19 (11.1)

[a]A currency exchange rate of CAD $1=US $0.76 is applicable.

Primary Outcomes

At baseline (period 1), when no health center had yet received the intervention, but they had been exposed to passive dissemination of the decision guide (ie, the control condition), 92% (59/64) of frail older adults and 83% (53/64) of caregivers of frail older adults with cognitive impairment already reported an active role in decision-making (Multimedia Appendices 3 and 4). In all, 92.1% (139/151) of frail older adults recruited under the control condition reported an active role in decision-making versus 94.3% (149/160) of frail older adults recruited under the intervention condition, for an absolute increase of 3.3% (95% CI –5.8% to 12.4%, *P*=.47) after accounting for the secular trend (Table 3). Similarly, 77.8% (130/167) of caregivers recruited under the control condition reported an active role in decision-making versus 80.8% (139/172) under the intervention condition, for an absolute increase of 6.1% (95% CI –11.8% to 23.4%, *P*=.49) (Table 4) after accounting for the secular trend. The ICC (WpICC) and the CAC were, respectively, 0.051 and 0.627 in the frail older adults and 0.045 and 0.493 in the caregivers of frail older adults with cognitive impairment (Multimedia Appendices 5 and 6).

XSL•FO
RenderX

Table 3. Effect of the intervention on primary and secondary outcomes for frail older adults without cognitive impairment.

Outcomes	Values		Absolute scale effect size		Relative scale effect size	
	Control (n=151)	Intervention (n=160)	Proportion difference[a]/mean difference (95% CI)	*P* value	Odds ratio (95% CI)[b]	*P* value
Primary outcome, n (%)						
Role assumed (active)[c]	139 (92.1)	149 (94.3)	3.3 (–5.8 to 12.4)	.47	1.70 (0.28 to 10.4)	.56
Secondary outcomes						
Preferred housing option (stay at home),[d] n (%)	100 (66.7)	97 (60.6)	–9.4 (–27.0 to 8.2)	.29	0.65 (0.24 to 1.75)	.39
Housing decision made (stay at home),[d] n (%)	41 (27.3)	61 (38.1)	3.3 (–14.1 to 20.7)	.71	1.16 (0.28 to 4.85)	.84
Decisional conflict (yes; scale ≥37.5), n (%)	28 (18.5)	20 (12.5)	–2.2 (–15.3 to 10.8)	.73	0.87 (0.20 to 3.74)	.85
Decisional regret (yes; scale >0), n (%)	107 (70.9)	108 (67.5)	–13.9 (–31.3 to 3.6)	.12	0.50 (0.12 to 2.11)	.34
Involvement in decision-making (Dyadic-OPTION),[e] mean (SD)	65.8 (19.4)	67.9 (17.2)	5.8 (–0.5 to 12.1)[f]	.07	N/A[g]	N/A
Quality of life (0-100),[h] mean (SD)	72.9 (23.8)	75.1 (22.3)	–2.1 (–10.0 to 5.9)[g]	.61	N/A	N/A

[a]Generalized linear mixed model with logit link function including intervention as a binary variable, a fixed effect (categorical) for time, and specifying a random effect for cluster.

[b]Linear mixed model with dichotomous dependent variables to handle convergence issues and reported risk differences, which can be interpreted as a difference of proportions (dependent dichotomous variables coded 1/0) [54-56].

[c]n=149 and n=158 for the control and intervention groups, respectively.

[d]n=150 and n=159 for the control and intervention groups, respectively.

[e]Assessed on a continuous scale ranging from 0 to 100.

[f]Linear mixed model including intervention as binary variable, a fixed effect (categorical) for time, and specifying a random effect for cluster.

[g]N/A: not applicable.

[h]Assessed on a continuous scale ranging from 0 to 100.

XSL•FO
RenderX

Table 4. Effect of the intervention on primary and secondary outcomes for caregivers of cognitively-impaired frail older adults (primary analysis).

Outcomes	Outcome frequency		Absolute scale effect size		Relative scale effect size	
	Control (n=167)	Intervention (n=172)	Proportional difference[a]/mean difference (95% CI)	*P* value	Odds ratio (95% CI)[b]	*P* value
Primary outcome, n (%)						
Role assumed (active)	130 (77.8)	139 (80.8)	6.1 (–11.2 to 23.4)	.49	1.30 (0.66 to 2.55)	.45
Secondary outcomes						
Preferred housing option (stay at home), n (%)	82 (49.1)	83 (48.3)	–2.7 (–19.4 to 14.1)	.76	0.89 (0.41 to 1.95)	.77
Housing decision made (stay at home), n (%)	27 (16.2)	36 (20.9)	2.6 (–10.0 to 15.2)	.69	1.10 (0.46 to 2.62)	.83
Decisional conflict (yes: scale ≥37.5), n (%)	23 (13.8)	19 (11.1)	–7.5 (–16.5 to 1.6)	.10	0.46 (0.19 to 1.11)	.08
Decisional regret (yes: scale >0), n (%)	117 (70.1)	124 (72.1)	1.7 (–15.0 to 18.3)	.84	1.03 (0.32 to 3.31)	.96
Involvement in decision making (Dyadic-OPTION),[c] mean (SD)	69.3 (17.6)	69.4 (19.8)	1.2 (–5.2 to 7.6)[d]	.72	N/A[e]	N/A
Burden of care[f] (0-88), mean (SD)	34.6 (17.2)	31.3 (16.5)	–1.1 (–6.2 to 4.0)[d]	.66	N/A	N/A

[a]Generalized linear mixed model with adaptive Gaussian–Hermite approximation to the likelihood maximum using an identity link, including intervention as binary variable, a fixed effect (categorical) for time, and specifying a random effect for cluster.

[b]Generalized linear mixed model with logit link function, including intervention as binary variable, a fixed effect (categorical) for time, and specifying a random effect for cluster.

[c]Assessed on a continuous scale ranging from 0 to 100.

[d]Linear mixed model including intervention as binary variable, a fixed effect (categorical) for time, and specifying a random effect for cluster.

[e]N/A: not applicable.

[f]Assessed on continuous scale ranging from 0 to 88.

Secondary Outcomes

The intervention had no statistically significant effect on any secondary outcomes among the frail older adults or caregivers. Frail older adults' perception of the extent to which health professionals involved them in decision-making scored an average of 67 of 100 with a (nonsignificant) mean increase of 5.4 (95% CI –0.6 to 11.4; *P*=.07). For caregivers, there was a nonsignificant effect on decisional conflict: 13.8% (23/167) in the control group versus 11% (19/172) in the intervention group, for an absolute decrease of 7.5% (95% CI –16.5% to 1.6%, *P*=.10) (Tables 3 and 4).

Discussion

Principal Findings

This study evaluated the effectiveness of adding training in IP-SDM for home care teams to the passive dissemination of a decision guide on the proportion of frail older adults, or caregivers of frail older adults with cognitive impairment, who reported taking an active role in making a housing decision. In this pragmatic trial, we observed a nonsignificant increase in the proportion of participants reporting an active role in decision-making. We observed no significant effect on any secondary outcomes. However, for frail older adults, there was an absolute (nonsignificant) increase in the extent to which health professionals involved them in decision-making and an absolute (nonsignificant) decrease in decisional conflict among caregivers. These results lead us to make the following observations.

Interpretation and Comparison With Prior Work

First, the nonsignificant increase observed in the primary outcome in both categories of participants (frail older adults and caregivers) may be explained by the fact that at baseline, the control group scored higher than expected. In our control condition, all clusters had been exposed to passive dissemination of the decision guide. In the trial that informed our sample size calculation (caregivers only), where the control group received usual care (ie, without the decision guide), fewer participants reported playing an active role at baseline and there was more room for improvement [10]. In the earlier trial, caregivers were also younger, and other studies confirm that younger people want a more active role in decision-making [57]. Both trials were pragmatic, and the loss of efficacy in a real clinical practice setting was to be expected. Interestingly, in both trials with caregivers of frail older adults with cognitive impairment, regardless of the decision-making role they assumed at baseline, an active decision-making role postintervention seemed to reach a similar threshold and go no further: in the first study, 79.6% took an active role postintervention [10], compared to 80.8% (139/172) in the current study. This suggests that among caregivers there is a natural ceiling to the expectation or desire to be active in decision-making on behalf of frail older adults. This threshold could be linked to discomfort with the role of being an active proxy decision-maker for more difficult and preference-sensitive decisions. At these times, it may be less

XSL•FO
RenderX

stressful to surrender responsibility for decision-making to the clinician [58].

Second, we observed high staff turnover during the study. In a postintervention follow-up, we found that of the 281 health professionals who received the intervention, less than half remained, possibly due to a major restructuring of the Quebec health care system occurring at the time [46]. High staff turnover was identified as one of the main barriers to engaging in IP-SDM [59]. Thus, many participants were being cared for by staff who had not been exposed to the intervention, likely contributing to its ineffectiveness. Repeating the intervention with replacement staff could have remedied this [60]. Periodic reminders [61] and postintervention coaching could have increased the long-term effects of the intervention and fidelity to it [62]. Changing clinical, organizational, and policy-making environments can have major impacts on pragmatic trials such as ours.

Finally, the health professionals were under severe time constraints. Caregivers may have felt they should not take up too much time talking about their preferences and values, although this was suggested by the decision guide [63]. In addition, the home care teams may have felt that SDM as presented in the training would be too time-consuming, even though they may, in fact, have already been collaborating with patients and their caregivers in decision-making [7]. The perception of SDM as an issue related to the quantity of time needs to shift to a perception that SDM is rather an issue related to the quality of time [63]. Our results could be interpreted as showing that in this context, with overworked staff and high turnover, the decision aid was the most appropriate and practical intervention for increasing client involvement in decision-making.

A strength of this trial was that it was pragmatic, according to the pragmatic–explanatory continuum indicator summary (PRECIS-2) [64]. Pragmatic trials are more applicable to real clinical practice [65] and increase external validity [66]. Second, no health center was lost to follow-up, reducing selection bias and indicating that the study was relevant to its participants. Decision support for housing decisions was clearly already of great interest even before the COVID-19 pandemic and its catastrophic consequences for long-term care residents made housing decisions a policy priority [67]. Third, all analyses gave similar results, demonstrating their consistency (Multimedia Appendices 7, 8, 9, and 10).

Limitations

This study had a number of limitations. First, we assumed our sample size would give us enough power to detect a 20% increase in our primary outcome, but the increase was 6.1% (not statistically significant). This lack of power may also explain why our study failed to detect a significant difference between the study groups, given the large CIs around their point estimates [10]. Second, identifying and recruiting participants after randomization may have increased the risk of selection bias, which would have caused under- or overestimation of the effect. However, the fact that characteristics were overall well balanced between groups indicates that this bias was minimal; we also adjusted for imbalanced variables to mitigate their influence on the estimate. Third, health professionals may have selected compliant participants, thereby inducing selection bias [68]. However, this limitation would have affected both the intervention and control groups. Fourth, the decision guide was distributed to all health professionals in the workshop. A question in our survey as to whether older adults and caregivers had been shown the decision guide should have provided us with a pseudofidelity variable regarding its use with patients [15], but due to a high level of missing data for this question, we could not include this as an outcome. It may be possible that there was a lack of fidelity to the implementation of the intervention. In this pragmatic trial, we were not able to be present at the consultations to assess this. A future mixed methods or qualitative study could provide this information and help us to better see the impact of the intervention. Finally, at the cluster level, the intervention may not be applicable in every setting, since home care services are organized differently from one jurisdiction to another [10]. At the individual level, however, the results of this study are generalizable to frail older adults and caregivers of frail older adults with cognitive impairment with similar characteristics facing housing decisions.

Conclusions

Adding IP-SDM training to passive dissemination of a decision guide for home care teams was not sufficient to induce frail older adults or caregivers of cognitively-impaired frail older adults to take a more active role in housing decisions. Baseline involvement in decision-making was already high, suggesting that home care teams are already practicing a form of collaborative decision-making with their clients. When home care teams are overworked and understaffed, providing them with high-quality practical tools may be the best way to support them in involving their clients in decision-making. Further research could explore more effective dissemination of decision guides, a new SDM focus on time quality instead of time quantity, and how to adapt SDM interventions to crisis situations (eg, pandemics), when staff are absent or turnover is especially high.

Acknowledgments

We thank Guy Lacroix, Sophie Desroches, Serge Dumont, Kimberly D Fraser, Stéphane Turcotte, Henriette Bourassa, Lise Roy, Geneviève Painchaud Guérard, Sylvie-Marianne Rhugenda, and Sabrina Guay-Bélanger for their involvement in the realization of the trial. We also thank the managers and home care teams of Centre intégré de santé et de services sociaux (CISSS) Abitibi-Témiscamingue, CISSS Montérégie-Ouest, Centre intégré universitaire de santé et de services sociaux (CIUSSS) Estrie, CISSS Laurentides, CISSS Laval, CISSS Outaouais, CISSS Montérégie-Est, CISSS Montérégie-Centre, and CIUSSS de l'Ouest-de-l'île-de-Montréal who participated in this study for the difficult and useful work they are doing every day for the

XSL•FO
RenderX

well-being of older adults and their family caregivers. We thank Dr. Henian Chen, professor of biostatistics at the College of Public Health, University of South Florida, USA, and Yuanyuan Lu, his PhD student in biostatistics. Finally, we thank Louisa Blair for her editorial support.

Conflicts of Interest

DS holds a University of Ottawa research chair in knowledge translation to patients. The authors have no further conflcits of interest to declare.

Multimedia Appendix 1
Characteristic of frail elders by allocated sequence.
[DOCX File , 31 KB - aging_v5i3e39386_app1.docx]

Multimedia Appendix 2
Characteristic of caregivers of cognitively-impaired frail elders by allocated sequence.
[DOCX File , 32 KB - aging_v5i3e39386_app2.docx]

Multimedia Appendix 3
Marginal frequencies of the primary outcome by period and cluster for frail elderly without cognitive impairment.
[DOCX File , 32 KB - aging_v5i3e39386_app3.docx]

Multimedia Appendix 4
Marginal frequencies of the primary outcome by period and cluster for caregivers of cognitively-impaired frail elderly.
[DOCX File , 32 KB - aging_v5i3e39386_app4.docx]

Multimedia Appendix 5
Intraclass correlation coefficient (ICC) and cluster autocorrelation coefficient for primary and secondary outcomes among frail elders without cognitive impairment.
[DOCX File , 29 KB - aging_v5i3e39386_app5.docx]

Multimedia Appendix 6
Intraclass correlation coefficient (ICC) and cluster autocorrelation coefficient for primary and secondary outcomes among caregivers of cognitively-impaired frail elders.
[DOCX File , 29 KB - aging_v5i3e39386_app6.docx]

Multimedia Appendix 7
Effect of the intervention on primary and secondary outcomes for frail elders without cognitive impairment (secondary analyses).
[DOCX File , 32 KB - aging_v5i3e39386_app7.docx]

Multimedia Appendix 8
Effect of the intervention on primary and secondary outcomes for caregivers of cognitively-impaired frail elders (secondary analyses).
[DOCX File , 32 KB - aging_v5i3e39386_app8.docx]

Multimedia Appendix 9
Effect of the intervention on primary and secondary outcomes for frail elders without cognitive impairment using a model based on autoregressive between-period correlation (sensitivity analyses).
[DOCX File , 32 KB - aging_v5i3e39386_app9.docx]

Multimedia Appendix 10
Effect of the intervention on primary and secondary outcomes for caregivers of cognitively-impaired frail elders using a model based on uniform between-period correlation (sensitivity analyses).
[DOCX File , 32 KB - aging_v5i3e39386_app10.docx]

Multimedia Appendix 11
CONSORT eHEALTH Checklist (V 1.6.1).

XSL•FO
RenderX

[PDF File (Adobe PDF File), 96 KB - aging_v5i3e39386_app11.pdf]

References

1. Jaul E, Barron J. Age-related diseases and clinical and public health implications for the 85 years old and over population. Front Public Health 2017 Dec 11;5:335 [FREE Full text] [doi: 10.3389/fpubh.2017.00335] [Medline: 29312916]
2. Murman D. The Impact of Age on Cognition. Semin Hear 2015 Aug;36(3):111-121 [FREE Full text] [doi: 10.1055/s-0035-1555115] [Medline: 27516712]
3. Roy N, Dubé R, Després C, Freitas A, Légaré F. Choosing between staying at home or moving: A systematic review of factors influencing housing decisions among frail older adults. PLoS One 2018;13(1):e0189266 [FREE Full text] [doi: 10.1371/journal.pone.0189266] [Medline: 29293511]
4. Garvelink MM, Emond J, Menear M, Brière N, Freitas A, Boland L, et al. Development of a decision guide to support the elderly in decision making about location of care: an iterative, user-centered design. Res Involv Engagem 2016;2:26 [FREE Full text] [doi: 10.1186/s40900-016-0040-0] [Medline: 29062524]
5. Bradshaw SA, Playford ED, Riazi A. Living well in care homes: a systematic review of qualitative studies. Age Ageing 2012 Jul;41(4):429-440. [doi: 10.1093/ageing/afs069] [Medline: 22678747]
6. McQueen P. The role of regret in medical decision-making. Ethic Theory Moral Prac 2017 Nov 4;20(5):1051-1065. [doi: 10.1007/s10677-017-9844-8]
7. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367 [FREE Full text] [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]
8. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Apr 12;4:CD001431 [FREE Full text] [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
9. Stacey D, Légaré F, Lewis KB. Patient decision aids to engage adults in treatment or screening decisions. JAMA 2017 Aug 15;318(7):657-658. [doi: 10.1001/jama.2017.10289] [Medline: 28810006]
10. Adekpedjou R, Stacey D, Brière N, Freitas A, Garvelink MM, Dogba MJ, et al. Engaging Caregivers in Health-Related Housing Decisions for Older Adults With Cognitive Impairment: A Cluster Randomized Trial. Gerontologist 2020 Jul 15;60(5):947-957 [FREE Full text] [doi: 10.1093/geront/gnz045] [Medline: 31095318]
11. Ampe S, Sevenants A, Smets T, Declercq A, Van Audenhove C. Advance care planning for nursing home residents with dementia: Influence of 'we DECide' on policy and practice. Patient Educ Couns 2017 Jan;100(1):139-146. [doi: 10.1016/j.pec.2016.08.010] [Medline: 27544017]
12. Govasli L, Solvoll B. Nurses' experiences of busyness in their daily work. Nurs Inq 2020 Jul;27(3):e12350. [doi: 10.1111/nin.12350] [Medline: 32133740]
13. Giguère A, Légaré F, Grimshaw J, Turcotte S, Fiander M, Grudniewicz A, et al. Printed educational materials: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev 2012 Oct 17;10:CD004398 [FREE Full text] [doi: 10.1002/14651858.CD004398.pub3] [Medline: 23076904]
14. Hemming K, Taljaard M, McKenzie JE, Hooper R, Copas A, Thompson JA, et al. Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration. BMJ 2018 Nov 09;363:k1614 [FREE Full text] [doi: 10.1136/bmj.k1614] [Medline: 30413417]
15. Légaré F, Brière N, Stacey D, Lacroix G, Desroches S, Dumont S, et al. Implementing shared decision-making in interprofessional home care teams (the IPSDM-SW study): protocol for a stepped wedge cluster randomised trial. BMJ Open 2016 Nov 24;6(11):e014023 [FREE Full text] [doi: 10.1136/bmjopen-2016-014023] [Medline: 27884857]
16. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 2015 Feb 06;350:h391 [FREE Full text] [doi: 10.1136/bmj.h391] [Medline: 25662947]
17. Hemming K, Taljaard M. Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice? Int J Epidemiol 2020 Jun 01;49(3):1043-1052 [FREE Full text] [doi: 10.1093/ije/dyaa077] [Medline: 32386407]
18. Ottawa Decision Support Tutorial. Ottawa Hospital Research Institute. URL: https://decisionaid.ohri.ca/ODST/index.php [accessed 2022-08-31]
19. Boland L, Légaré F, Carley M, Graham ID, O'Connor AM, Lawson ML, et al. Evaluation of a shared decision making educational program: The Ottawa Decision Support Tutorial. Patient Educ Couns 2019 Feb;102(2):324-331. [doi: 10.1016/j.pec.2018.09.008] [Medline: 30224274]
20. Stacey D, Brière N, Robitaille H, Fraser K, Desroches S, Légaré F. A systematic process for creating and appraising clinical vignettes to illustrate interprofessional shared decision making. J Interprof Care 2014 Sep;28(5):453-459. [doi: 10.3109/13561820.2014.911157] [Medline: 24766619]
21. Maltais P, Goulet F, Borduas F. Educational skills and knowledge needed and problems encountered by continuing medical education providers. J Contin Educ Health Prof 2000;20(2):91-96. [doi: 10.1002/chp.1340200205] [Medline: 11232225]

XSL•FO
RenderX

22. Copas AJ, Lewis JJ, Thompson JA, Davey C, Baio G, Hargreaves JR. Designing a stepped wedge trial: three main designs, carry-over effects and randomisation approaches. Trials 2015 Aug 17;16:352 [FREE Full text] [doi: 10.1186/s13063-015-0842-7] [Medline: 26279154]
23. Kim S, Bayer I, Gewurtz R, Larivière N, Letts L. Comparing web-based and in-person educational workshops for canadian occupational therapists and understanding their learning experiences: mixed methods study. JMIR Med Educ 2022 Jan 04;8(1):e31634. [doi: 10.2196/31634] [Medline: 34982719]
24. Rottenberg S, Williams A. Web-based delivery of the caregiving essentials course for informal caregivers of older adults in ontario: mixed methods evaluation study. JMIR Aging 2021 Jun 15;4(2):e25671 [FREE Full text] [doi: 10.2196/25671] [Medline: 34128815]
25. Cohen-Mansfield J, Muff A, Meschiany G, Lev-Ari S. Adequacy of web-based activities as a substitute for in-person activities for older persons during the covid-19 pandemic: survey study. J Med Internet Res 2021 Jan 22;23(1):e25848 [FREE Full text] [doi: 10.2196/25848] [Medline: 33439851]
26. Geiger F, Hacke C, Potthoff J, Scheibler F, Rueffer JU, Kuch C, et al. The effect of a scalable online training module for shared decision making based on flawed video examples - a randomized controlled trial. Patient Educ Couns 2021 Jul;104(7):1568-1574. [doi: 10.1016/j.pec.2020.11.033] [Medline: 33334633]
27. Degner LF, Sloan JA. Decision making during serious illness: what role do patients really want to play? J Clin Epidemiol 1992 Sep;45(9):941-950. [doi: 10.1016/0895-4356(92)90110-9] [Medline: 1432023]
28. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 1995;15(1):25-30. [doi: 10.1177/0272989X9501500105] [Medline: 7898294]
29. Legaré F, Graham ID, O'Connor AM, Dolan JG, Bélanger-Ducharme F. Prise de décision partagée : traduction et validation d'une échelle de confort décisionnel du médecin. Pédagogie Médicale 2008 Aug 01;4(4):216-222 [FREE Full text] [doi: 10.1051/pmed:2003031]
30. Brehaut JC, O'Connor AM, Wood TJ, Hack TF, Siminoff L, Gordon E, et al. Validation of a decision regret scale. Med Decis Making 2003;23(4):281-292. [doi: 10.1177/0272989X03256005] [Medline: 12926578]
31. Melbourne E, Sinclair K, Durand M, Légaré F, Elwyn G. Developing a dyadic OPTION scale to measure perceptions of shared decision making. Patient Educ Couns 2010 Feb;78(2):177-183. [doi: 10.1016/j.pec.2009.07.009] [Medline: 19647970]
32. Melbourne E, Roberts S, Durand M, Newcombe R, Légaré F, Elwyn G. Dyadic OPTION: Measuring perceptions of shared decision-making in practice. Patient Educ Couns 2011 Apr;83(1):55-57. [doi: 10.1016/j.pec.2010.04.019] [Medline: 20537837]
33. Bucquet D, Condon S, Ritchie K. The French version of the Nottingham Health Profile. A comparison of items weights with those of the source version. Soc Sci Med 1990;30(7):829-835. [doi: 10.1016/0277-9536(90)90207-9] [Medline: 2315749]
34. Sharples LD, Todd CJ, Caine N, Tait S. Measurement properties of the Nottingham Health Profile and Short Form 36 health status measures in a population sample of elderly people living at home: results from ELPHS. Br J Health Psychol 2000;5(3):217-233. [doi: 10.1348/135910700168874]
35. Faria CDCM, Teixeira-Salmela LF, Nascimento VB, Costa AP, Brito NDP, Rodrigues-De-Paula F. Comparisons between the Nottingham Health Profile and the Short Form-36 for assessing the quality of life of community-dwelling elderly. Rev Bras Fisioter 2011 Oct;15(5):399-405. [doi: 10.1590/s1413-35552011005000023] [Medline: 22002188]
36. Zarit S, Orr NK, Zarit JM. The hidden victims of Alzheimer's disease: Families under stress. New York, NY: NYU Press; 1985.
37. Seng BK, Luo N, Ng WY, Lim J, Chionh HL, Goh J, et al. Validity and reliability of the Zarit Burden Interview in assessing caregiving burden. Ann Acad Med Singap 2010 Oct;39(10):758-763. [Medline: 21063635]
38. Hébert R, Bravo G, Girouard D. Fidélité de la traduction française de trois instruments d'évaluation des aidants naturels de malades déments. Can J Aging 2010 Nov 29;12(3):324-337. [doi: 10.1017/s0714980800013726]
39. Sepucha K, Mulley AG. A perspective on the patient's role in treatment decisions. Med Care Res Rev 2009 Feb;66(1 Suppl):53S-74S. [doi: 10.1177/1077558708325511] [Medline: 19001081]
40. Elkin EB, Kim SH, Casper ES, Kissane DW, Schrag D. Desire for information and involvement in treatment decisions: elderly cancer patients' preferences and their physicians' perceptions. J Clin Oncol 2007 Nov 20;25(33):5275-5280. [doi: 10.1200/JCO.2007.11.1922] [Medline: 18024875]
41. Tariman JD, Berry DL, Cochrane B, Doorenbos A, Schepp KG. Physician, patient, and contextual factors affecting treatment decisions in older adults with cancer and models of decision making: a literature review. Oncol Nurs Forum 2012 Jan;39(1):E70-E83 [FREE Full text] [doi: 10.1188/12.ONF.E70-E83] [Medline: 22201670]
42. Légaré F, Brière N, Stacey D, Bourassa H, Desroches S, Dumont S, et al. Improving decision making on location of care with the frail elderly and their caregivers (the DOLCE study): study protocol for a cluster randomized controlled trial. Trials 2015 Feb 12;16:50 [FREE Full text] [doi: 10.1186/s13063-015-0567-7] [Medline: 25881122]
43. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials 2007 Feb;28(2):182-191. [doi: 10.1016/j.cct.2006.05.007] [Medline: 16829207]

XSL•FO
RenderX

44. Campbell MK, Fayers PM, Grimshaw JM. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. Clin Trials 2005;2(2):99-107. [doi: 10.1191/1740774505cn071oa] [Medline: 16279131]
45. Légaré F, Labrecque M, Cauchon M, Castel J, Turcotte S, Grimshaw J. Training family physicians in shared decision-making to reduce the overuse of antibiotics in acute respiratory infections: a cluster randomized trial. CMAJ 2012 Sep 18;184(13):E726-E734 [FREE Full text] [doi: 10.1503/cmaj.120568] [Medline: 22847969]
46. O-7.2 Act to Modify the Organization and Governance of the Health and Social Services Network, in Particular by Abolishing the Regional Agencies. Gouvernement du Québec. 2015. URL: https://www.legisquebec.gouv.qc.ca/en/document/cs/O-7.2 [accessed 2022-09-02]
47. Ivers NM, Halperin IJ, Barnsley J, Grimshaw JM, Shah BR, Tu K, et al. Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. Trials 2012 Aug 01;13:120 [FREE Full text] [doi: 10.1186/1745-6215-13-120] [Medline: 22853820]
48. VanderWeele TJ, Shpitser I. A new criterion for confounder selection. Biometrics 2011 Dec;67(4):1406-1413 [FREE Full text] [doi: 10.1111/j.1541-0420.2011.01619.x] [Medline: 21627630]
49. Hemming K, Taljaard M, Forbes A. Analysis of cluster randomised stepped wedge trials with repeated cross-sectional samples. Trials 2017 Mar 04;18(1):101 [FREE Full text] [doi: 10.1186/s13063-017-1833-7] [Medline: 28259174]
50. Girling AJ, Hemming K. Statistical efficiency and optimal design for stepped cluster studies under linear mixed effects models. Stat Med 2016 Jun 15;35(13):2149-2166 [FREE Full text] [doi: 10.1002/sim.6850] [Medline: 26748662]
51. Hooper R, Teerenstra S, de Hoop E, Eldridge S. Sample size calculation for stepped wedge and other longitudinal cluster randomised trials. Stat Med 2016 Nov 20;35(26):4718-4728. [doi: 10.1002/sim.7028] [Medline: 27350420]
52. Kasza J, Hemming K, Hooper R, Matthews J, Forbes A. Impact of non-uniform correlation structure on sample size and power in multiple-period cluster randomised trials. Stat Methods Med Res 2019 Mar;28(3):703-716. [doi: 10.1177/0962280217734981] [Medline: 29027505]
53. Austin PC. Estimating multilevel logistic regression models when the number of clusters is low: a comparison of different statistical software procedures. Int J Biostat 2010 Apr 22;6(1):Article 16 [FREE Full text] [doi: 10.2202/1557-4679.1195] [Medline: 20949128]
54. Hemming K, Ryan R, Gill P, Westerby P, Jolly K, Marshall T. Targeted case finding in the prevention of cardiovascular disease: a stepped wedge cluster randomised controlled trial. Br J Gen Pract 2016 Oct;66(651):e758-e767 [FREE Full text] [doi: 10.3399/bjgp16X686629] [Medline: 27528707]
55. Leontjevas R, Gerritsen DL, Smalbrugge M, Teerenstra S, Vernooij-Dassen MJ, Koopmans RT. A structural multidisciplinary approach to depression management in nursing-home residents: a multicentre, stepped-wedge cluster-randomised trial. Lancet 2013 Jun 29;381(9885):2255-2264. [doi: 10.1016/S0140-6736(13)60590-5] [Medline: 23643110]
56. Hannan PJ, Murray DM. Gauss or Bernoulli? A Monte Carlo comparison of the performance of the linear mixed-model and the logistic mixed-model analyses in simulated community trials with a dichotomous outcome variable at the individual level. Eval Rev 1996 Jun;20(3):338-352. [doi: 10.1177/0193841X9602000306] [Medline: 10182208]
57. Deber RB, Kraetschmer N, Urowitz S, Sharpe N. Do people want to be autonomous patients? Preferred roles in treatment decision-making in several patient populations. Health Expect 2007 Sep;10(3):248-258 [FREE Full text] [doi: 10.1111/j.1369-7625.2007.00441.x] [Medline: 17678513]
58. Pel-Littel RE, Snaterse M, Teppich NM, Buurman BM, van Etten-Jamaludin FS, van Weert JCM, et al. Barriers and facilitators for shared decision making in older patients with multiple chronic conditions: a systematic review. BMC Geriatr 2021 Feb 06;21(1):112 [FREE Full text] [doi: 10.1186/s12877-021-02050-y] [Medline: 33549059]
59. Légaré F, Stacey D, Brière N, Fraser K, Desroches S, Dumont S, et al. Healthcare providers' intentions to engage in an interprofessional approach to shared decision-making in home care programs: a mixed methods study. J Interprof Care 2013 May;27(3):214-222 [FREE Full text] [doi: 10.3109/13561820.2013.763777] [Medline: 23394265]
60. Fuller C, Michie S, Savage J, McAteer J, Besser S, Charlett A, et al. The Feedback Intervention Trial (FIT)--improving hand-hygiene compliance in UK healthcare workers: a stepped wedge cluster randomised controlled trial. PLoS One 2012;7(10):e41617 [FREE Full text] [doi: 10.1371/journal.pone.0041617] [Medline: 23110040]
61. Johnson MJ, May CR. Promoting professional behaviour change in healthcare: what interventions work, and why? A theory-led overview of systematic reviews. BMJ Open 2015 Sep 30;5(9):e008592 [FREE Full text] [doi: 10.1136/bmjopen-2015-008592] [Medline: 26423853]
62. Zomahoun HTV, Guénette L, Grégoire JP, Lauzier S, Lawani AM, Ferdynus C, et al. Effectiveness of motivational interviewing interventions on medication adherence in adults with chronic diseases: a systematic review and meta-analysis. Int J Epidemiol 2017 Apr 01;46(2):589-602. [doi: 10.1093/ije/dyw273] [Medline: 27864410]
63. Yahanda AT, Mozersky J. What's the role of time in shared decision making? AMA J Ethics 2020 May 01;22(5):E416-E422 [FREE Full text] [doi: 10.1001/amajethics.2020.416] [Medline: 32449658]
64. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015 May 08;350:h2147. [doi: 10.1136/bmj.h2147] [Medline: 25956159]
65. Patsopoulos NA. A pragmatic view on pragmatic trials. Dialogues Clin Neurosci 2011;13(2):217-224 [FREE Full text] [doi: 10.31887/DCNS.2011.13.2/npatsopoulos] [Medline: 21842619]

XSL•FO
RenderX

66. Lurie JD, Morgan TS. Pros and cons of pragmatic clinical trials. J Comp Eff Res 2013 Jan;2(1):53-58. [doi: 10.2217/cer.12.74] [Medline: 24236521]
67. Linton SL, Leifheit KM, McGinty EE, Barry CL, Pollack CE. Association between housing insecurity, psychological distress, and self-rated health among US adults during the COVID-19 pandemic. JAMA Netw Open 2021 Sep 01;4(9):e2127772 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.27772] [Medline: 34591109]
68. Ertmann RK, Nicolaisdottir DR, Kragstrup J, Siersma V, Overbeck G, Wilson P, et al. Selection bias in general practice research: analysis in a cohort of pregnant Danish women. Scand J Prim Health Care 2020 Dec;38(4):464-472 [FREE Full text] [doi: 10.1080/02813432.2020.1847827] [Medline: 33242291]

Abbreviations

BpICC: between-period intraclass correlation coefficient
CAC: cluster autocorrelation coefficient
GLMM: generalized linear mixed models
ICC: intraclass correlation
IP-SDM: interprofessional shared decision-making
RA: research assistant
SDM: shared decision-making
WpICC: within-period intraclass correlation coefficients

Edited by T Leung, J Wang; submitted 09.05.22; peer-reviewed by N Bier; comments to author 31.05.22; revised version received 06.06.22; accepted 27.06.22; published 20.09.22.

Please cite as:
Adisso ÉL, Taljaard M, Stacey D, Brière N, Zomahoun HTV, Durand PJ, Rivest LP, Légaré F
Shared Decision-Making Training for Home Care Teams to Engage Frail Older Adults and Caregivers in Housing Decisions: Stepped-Wedge Cluster Randomized Trial
JMIR Aging 2022;5(3):e39386
URL: https://aging.jmir.org/2022/3/e39386
doi:10.2196/39386
PMID:35759791

©Évèhouénou Lionel Adisso, Monica Taljaard, Dawn Stacey, Nathalie Brière, Hervé Tchala Vignon Zomahoun, Pierre Jacob Durand, Louis-Paul Rivest, France Légaré. Originally published in JMIR Aging (https://aging.jmir.org), 20.09.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX