Original Papers

The Impact of a Naturally Occurring Retirement Community Supportive Services Program on Older Adult Participants’ Social Networks: Semistructured Interview Study (e37617)
Christine Mills, Simone Parniak, Carri Hand, Colleen McGrath, Debbie Laliberte Rudman, Cassandra Chislett, Mariah Giberson, Lauren White, Vincent DePaul, Catherine Donnelly. ................................................................. 3

Enhancing Food Intake Tracking in Long-term Care With Automated Food Imaging and Nutrient Intake Tracking (AFINI-T) Technology: Validation and Feasibility Assessment (e37590)
Kaylen Pfisterer, Robert Amelard, Jennifer Boger, Heather Keller, Audrey Chung, Alexander Wong. ................................................................. 14

Older Wheelchair Users Recommend Age-Friendly Design Improvements to a Wheelchair Maintenance App: Mixed Methods Development Study (e39301)
Alyssa Boccardi, Fangzheng Wu, Jon Pearlman, Anand Mhatre. ................................................................. 32

The Factors Influencing Older Adults’ Decisions Surrounding Adoption of Technology: Quantitative Experimental Study (e39890)
Jerad Moxley, Joseph Sharit, Sara Czaja. ................................................................. 47

An Online Dual-Task Cognitive and Motor Exercise Program for Individuals With Parkinson Disease (PD3 Move Program): Acceptability Study (e40325)
Josefa Domingos, John Dean, Júlio Fernandes, Catarina Godinho. ................................................................. 64

Impact and Perceived Value of iGeriCare e-Learning Among Dementia Care Partners and Others: Pilot Evaluation Using the IAM4all Questionnaire (e40357)
Andrew Scott, Stephanie Ayers, Pierre Pluye, Roland Grad, Richard Sztramko, Sharon Marr, Alexandra Papaioannou, Sandra Clark, Patricia Gerantonis, Anthony Levinson. ................................................................. 98

Electronic Tracking Devices for People With Dementia: Content Analysis of Company Websites (e38865)
Jared Howes, Yvonne Denier, Chris Gastmans. ................................................................. 105

The Impact of COVID-19 on Older Adults’ Perceptions of Virtual Care: Qualitative Study (e38546)
Lama Abdallah, Paul Stoilee, Kimberly Lopez, Alexandra Whate, Jennifer Boger, Catherine Tong. ................................................................. 147

Usability and Acceptability of a Palliative Care Mobile Intervention for Older Adults With Heart Failure and Caregivers: Observational Study (e35592)
Jennifer Villalobos, Sheana Bull, Jennifer Portz. ................................................................. 158

Mobile App Prototype in Older Adults for Postfracture Acute Pain Management: User-Centered Design Approach (e37772)
Kevin Tran-Nguyen, Caroline Berger, Roxanne Bennett, Michelle Wall, Suzanne Morin, Fateme Rajabi yazdi. ................................................................. 167
Lessons and Reflections From an Extended Co-design Process Developing an mHealth App With and for Older Adults: Multiphase, Mixed Methods Study (e39189)
Catherine Tong, Alison Kernoghan, Kassandra Lemmon, Paige Fernandes, Jacobi Elliott, Veronica Sacco, Sheila Bodemer, Paul Stolee...  
1

Designing Mindfulness Conversational Agents for People With Early-Stage Dementia and Their Caregivers: Thematic Analysis of Expert and User Perspectives (e40360)
Cassandra Seah, Zheyuan Zhang, Sijin Sun, Esther Wiskerke, Sarah Daniels, Talya Porat, Rafael Calvo. .......................................................... 197

Lessons Learned From an Effectiveness Evaluation of Inlife, a Web-Based Social Support Intervention for Caregivers of People With Dementia: Randomized Controlled Trial (e38656)
Hannah Christie, Alieske Dam, Martin van Boxtel, Sebastian Köhler, Frans Verhey, Marjolein de Vugt. .......................................................... 210

Evolving Hybrid Partial Genetic Algorithm Classification Model for Cost-effective Frailty Screening: Investigative Study (e38464)
John Oates, Niusha Shaftiabady, Rachel Ambagtsheer, Justin Beilby, Chris Seiboth, Elsa Dent. .......................................................... 224

Reviews

Evidence and User Considerations of Home Health Monitoring for Older Adults: Scoping Review (e40079)
Andrew Chan, Rachel Cohen, Katherine-Marie Robinson, Devvrat Bhardwaj, Geoffrey Gregson, Jeffrey Jutai, Jason Millar, Adriana Rios Rincón, Atena Roshan Fekr. .......................................................... 75

Effectiveness of Technology Interventions in Addressing Social Isolation, Connectedness, and Loneliness in Older Adults: Systematic Umbrella Review (e40125)
Eric Balki, Niall Hayes, Carol Holland. .......................................................... 127
The Impact of a Naturally Occurring Retirement Community Supportive Services Program on Older Adult Participants’ Social Networks: Semistructured Interview Study

Christine Marie Mills¹, BASc, MPH; Simone Parniak², BHSc, MA; Carri Hand³, BScOT, MSc, PhD; Colleen McGrath³, BA, MSc, PhD; Debbie Laliberte Rudman¹, BScOT, MSc, PhD; Cassandra Chislett¹, BSc, MScOT; Mariah Giberson¹, BA, MScOT; Lauren White¹, MScOT; Vincent DePaul¹,², BHSc, PhD; Catherine Donnelly¹,², BScOT, MSc, PhD

¹School of Rehabilitation Therapy, Faculty of Health Sciences, Queen’s University, Kingston, ON, Canada
²Health Services and Policy Research Institute, Queen’s University, Kingston, ON, Canada
³School of Occupational Therapy, Faculty of Health Sciences, Western University, London, ON, Canada

Corresponding Author:
Christine Marie Mills, BASc, MPH
School of Rehabilitation Therapy
Faculty of Health Sciences
Queen’s University
Louise D Acton Building
31 George Street
Kingston, ON, K7P 0M4
Canada
Phone: 1 6139027722
Email: christine.mills@queensu.ca

Abstract

Background: Most older adults want to age in place, in their homes and communities. However, this can be challenging for many, frequently owing to lack of supports that allow for aging in place. Naturally occurring retirement community supportive services programs (NORC-SSPs) offer an approach to help older adults age in place. Although qualitative studies have examined the experiences of NORC-SSP participants, little is known about how participation in NORC-SSP programming affects participants’ social networks.

Objective: This study aimed to explore the experiences of 13 NORC-SSP residents who participated in Oasis Senior Supportive Living (Oasis) and how participating in NORC-SSP programming, specifically based on the Oasis model, influenced their social networks.

Methods: Participants were recruited, using convenience sampling, from 4 naturally occurring retirement communities (NORCs) in Ontario, Canada. All participants (13/13, 100%) had participated in Oasis programming. Semistructured qualitative interviews were conducted with participants. Social network theory informed the interview guide and thematic analysis.

Results: In total, 13 participants (n=12, 92% women and n=1, 8% men) were interviewed. These participants were from 4 different NORCs where Oasis had been implemented, comprising 2 midrise apartment buildings, 1 low-rise apartment building, and 1 mobile home community. Overall, 3 main themes were identified from the interviews with Oasis participants: expansion and deepening of social networks, Oasis activities (something to do, someone to do it with), and self-reported impact of Oasis on mental health and well-being (feeling and coping with life better). Participants noted that Oasis provided them with opportunities to meet new people and broaden their social networks, both within and outside their NORCs. They also indicated that Oasis provided them with meaningful ways to spend their time, including opportunities to socialize and try new activities. Participants stated that participating in Oasis helped to alleviate loneliness and improved their quality of life. They noted that Oasis provided them with a reason to get up in the morning. However, the experiences described by participants may not be reflective of all Oasis members. Those who had positive experiences may have been more likely to agree to be interviewed.

Conclusions: On the basis of the participants’ interviews, Oasis is an effective aging-in-place model that has been successfully implemented in low-rise apartment buildings, midrise apartment buildings, and mobile home communities. Participating in Oasis allowed participants to expand their social networks and improve their mental health and well-being. Therefore, NORCs may...
offer an ideal opportunity to build strong communities that provide deep, meaningful social connections that expand social networks. NORC-SSPs, such as Oasis, can support healthy aging and allow older adults to age in place.

(JMIR Aging 2022;5(4):e37617) doi:10.2196/37617

KEYWORDS

aging in place; naturally occurring retirement communities; social networks; social networking; social capital; aged

Introduction

Background

Older adults are the fastest-growing demographic group in Canada [1], and most of them want to age in place within their communities [2]. Aging in place means that an older adult’s health, access to services, and social support interact in ways that enable living safely and autonomously in their homes or their communities for as long as they desire [3]. Aging in place, although desired by most older adults, is challenging for many. Approximately one-fourth of older adults report feeling isolated [4,5], and between 20% and 34% indicate that they are lonely [6]. Loneliness and social isolation are associated with increased risk of mortality and morbidity, representing as strong a risk factor for premature death as smoking [7].

The World Health Organization defines healthy aging as “the process of developing and maintaining the functional ability that enables well-being in older age” [8]. Building and maintaining relationships are critical for health and well-being and enable older adults to continue to live and participate in their communities. Naturally occurring retirement communities (NORCs), unplanned communities (eg, apartment buildings) with a high proportion of older residents [9], are ideally positioned to support aging in place by integrating programs designed to build social connections and support healthy aging. By their very nature, NORCs are naturally existing high-density areas of older adults, which makes them a natural fit for older adult–focused programs and services, otherwise known as the creation of a NORC supportive services program (NORC-SSP).

A way in which NORCs support aging in place is through these NORC-SSPs. A NORC-SSP “is a community-level intervention in which older adults, building owners and managers, service providers, funders, and other community partners create a network of services and volunteer opportunities to promote aging in place among older adults who live in naturally occurring retirement communities, housing developments and residential areas not planned for older adults but in which large numbers of older adults reside” [10]. Thus, the goals of NORC-SSPs are to help older adults to age in place, in their communities, as most of them wish to do [2].

Oasis Senior Supportive Living (Oasis), is an example of a NORC-SSP. Oasis is an innovative aging-in-place model that integrates health and supportive services for older adults in a NORC in partnership with older adults, building owners, researchers, and community partners. The Oasis model was co-developed with the Kingston Council on Aging in 2011 by a group of older adults living in a Kingston, Ontario apartment building, to address the problem of social isolation and its associated health risks among its residents. Since 2019, Oasis programs have opened in 7 additional NORCs across Ontario, each providing a supportive, socially connected, living environment. Oasis empowers older adults to identify their needs and determine the services and activities that best meet these needs [11]. An on-site Oasis coordinator supports members to implement desired activities and services, including member-led activities such as craft circles, line dancing, potluck meals, or engaging community service providers to provide specific programming including exercise classes in communal meeting spaces. Oasis is novel in its emphasis on improving health through community development, strong social networks, and self-determination in older adult members.

NORC-SSPs were created with the intention of bringing programming to where older adults live to strengthen social connections and promote engagement in activities, with the ultimate goal of supporting older adults to age within their communities. Social network theory posits that social structures play an important role in determining individual behaviors [12] and offers a lens to delve deep into how social networks affect mental health and well-being within NORC-SSPs. Grounded in social network theory, Berkman et al [12] offer a conceptual framework that explores the influence of social networks (at the mesolevel) on individual behaviors (at the microlevel). These social networks and individual behaviors are seen to influence health outcomes, including mental health and well-being [12].

A recent scoping review of NORC literature found that much of the research seeks to define NORCs and describe activities and services meant to build strong social relationships [13]. So far, there has been less focus on the effect of NORCs and NORC-SSPs on social networks. Studies have found that participants living in NORC-SSPs have a greater sense of community [14-20], experience enhanced social support, and report reduced social isolation [11,15,17-19,21-25]. A handful of qualitative studies have been conducted with participants living in NORCs; one study sought to understand the factors that influence member participation in NORCs [26] and another study sought to gain perspectives of members living in a NORC as part of an evaluation to understand the implementation of programming in a NORC [27]. So far, no studies have sought to understand how and in what ways a NORC or NORC-SSP influences social networks and how these social networks in turn affect health and well-being.

Objectives

The purpose of this study was to (1) understand older adults’ experiences of participating in Oasis; (2) explore how participation in Oasis influenced their social networks at the mesolevel and microlevel, using the social network theory by Berkman et al [12]; and (3) understand how these social networks influenced participants’ health and well-being.
Methods

Approach
This study used a qualitative descriptive approach to examine the experiences of participants and their social networks (as described by Berkman et al [12]) within a NORC-SSP model, Oasis [28,29]. Qualitative description is grounded in the philosophical tenets of naturalistic inquiry [30] and is an approach to qualitative research that offers a summary and understanding of experiences and includes contextual factors that shape these experiences [31]. A qualitative descriptive study is guided by a theoretical framework, and the social network theory by Berkman et al [12] guided this study, including the development of the interview guide, identification of the variables and relationships to be explored, and analysis. For this study, we were interested in Oasis participants’ experiences, how their experiences affected their social networks, contexts in which the experiences occurred, and how participation in Oasis influenced their health and well-being, thus making qualitative description appropriate for this study and making the social network theory by Berkman et al [12] an appropriate theoretical framework.

Participants
Participants included a sample of older adults who participated in the Oasis program at 4 different locations in Ontario, Canada. The NORC contexts in which the Oasis program was implemented included 2 midrise apartment buildings, 1 low-rise apartment building, and 1 mobile home community. Recruitment continued until data saturation had occurred, that is, until interviewers repeatedly heard statements similar to those heard in previous interviews [32]. A total of 13 participants (n=12, 92% women and n=1, 8% men) representing the 4 Oasis communities were interviewed. The Oasis program had been implemented approximately 1 year before the interviews, thus ensuring that participants had experience with and engagement in the program.

Semistructured Interview Guide
Consistent with a qualitative descriptive approach that recommends theory-informed data collection, the social network theory by Berkman et al [12] supported the development of the interview guide (Multimedia Appendix 1). The interview guide sought to explore the macroenvironment that may have influenced social networks and included questions about sociostructural factors (such as building dynamics). Questions also examined the mesoenvironment and explored social networks, including who was in their social network, influence of Oasis on social networks, and psychosocial mechanisms of these networks including the types of supports they received, and Oasis programs they participated in. Finally, the interview guide included questions about potential health pathways, including questions about how Oasis affected daily routines and overall health and well-being.

Ethics Approval
Ethics approval and consent were obtained from the Queen’s University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board in Kingston, Ontario (REH-722-18).

Procedure
Participants were recruited by inviting older adults who were Oasis members to participate in the study. The study was conducted at the beginning of the COVID-19 pandemic, and owing to physical distancing protocols, all interviews were conducted via telephone during March 2020 and April 2020. Consent was obtained verbally. The interviews were audio-recorded and, subsequently, transcribed verbatim and analyzed.

Analysis
The aims of the analysis were to explore and understand experiences of the Oasis participants; therefore, thematic analysis was chosen as the analytic strategy. Thematic analysis, as described by Braun and Clarke [33], is a qualitative descriptive approach that is used to identify, analyze, and report patterns, called themes, within data. It is useful for “analyzing narrative materials of life stories” [34].

NVivo (version 12; QSR International) was used to conduct the thematic analysis of the transcripts. Social network theory [12] provided the conceptual framework for the development of the initial themes. Overall, 6 authors (CC, CD, MG, C Mills, SP, and LW) independently read one of the transcripts to develop a list of preliminary codes based on the conceptual framework. Then, this transcript and the codes each author had developed and applied were reviewed line by line as a group, and a preliminary code book was developed. Then, the 6 authors collated the codes into potential themes by grouping similar codes together. Then, 5 authors (CC, CD, MG, C Mills, and LW) were each assigned several transcripts to code; a second member of this coding team then reviewed how those transcripts had been coded. The codes were discussed until consensus was reached. Then, the same 6 authors met again as a group, reviewed one of the transcripts, and further refined the codes and themes until consensus was reached, and a master list of final codes and themes was developed. Then, each of these 6 authors returned to the transcripts they had been assigned to apply the master list to segments of text and to highlight quotes that provided strong illustrations of the themes. Then, 2 authors (C Mills and SP) reviewed the chosen quotes to find exemplars for each theme.

Results
Sample Description
In total, 13 participants from 4 different Oasis communities were interviewed. Of these 13 participants, 12 (92%) identified as women and 1 (8%) identified as a man. The represented Oasis communities consisted of 2 midrise apartment buildings, 1 low-rise apartment building, and 1 mobile home community. Overall, 3 broad themes were identified corresponding to the 3 levels in the theory by Berkman et al [12].
Social Networks (Mesolevel)—Expansion and Deepening of Oasis Member Social Networks

Participants’ Experiences

At the mesolevel, it was clear from the interviews that Oasis provided older adults the opportunity to expand their social networks. A participant shared the following:

[Oasis] gave us...an opportunity to meet people and live beyond our own walls. Expand us. Which was really a positive thing. Because you can always keep yourself busy at home but then you have no social network, you have no friends, you have no one to talk to. And you get small when you do that. [Participant 8]

Another participant (participant 12) stated, “It just makes you feel more like you’re more welcome to the building.”

Participants noted that Oasis provided them with opportunities to meet new people within their NORC and helped them to broaden their social network. A participant (participant 8) stated, “I was pretty social before this started but just in here there’s a lot more people I know...that’s just made my social group larger.”

Another participant (participant 6) shared, “I’ve lived here for four years, and I met more people when Oasis was going on than I had in the whole four years before that.”

Expanding their social networks meant that participants had more people they could call on for support. A participant shared the following:

It’s [social network] gotten a bit bigger. Before I only socialized with maybe 2, 3 people and since, Oasis has opened up a whole new group of friends that I can rely on and they can rely on me. [Participant 13]

New Friendships

In addition to the increase in social network size, participants spoke of new friendships that they had formed with other Oasis participants. Through the social components of Oasis, participants were able to connect with other older adults and form meaningful relationships that extended beyond the Oasis setting. A person shared the following:

I just gained so much from Oasis. I gained friends, I gained friendships, I gained companionship. I gained a lot of empathy because there was a lot of people who were a lot worse off than I am, that I was able to empathize with and show compassion for. [Participant 5]

Another participant stated the following:

We have some very tight friends we’ve made because of this. And that’s expanded our community. [Participant 8]

This participant later added the following:

[Oasis] kept us all getting to know each other even better and better. So, we’ve made some wonderful friends here. Wonderful. [Participant 8]

Deepening of Previous Friendships With Neighbors

Oasis also strengthened participants’ previous social networks. Participants emphasized that Oasis allowed them to deepen friendships that they had previously; this resulted in participants feeling more closely connected to residents of their NORC. A participant shared the following:

I think it strengthened it [my social network] here in the building and I saw those people that I mentioned more rather than less. Having a meal together once a week. Going to coffee hour...going to craft mornings occasionally. [Participant 3]

Another participant (participant 11) stated, “I developed deeper friendships with the people in the building.” This participant later added, “It [Oasis] just made me closer to people and makes better connections.”

Comfort in Knowing That Friends Are Nearby and Available to Provide Support If Needed

It appeared that participants appreciated the sense of comfort that came with having close connections and friendships with other residents of their NORC owing to participating in Oasis. Participants often spoke about how nice it was to know that they had people in their social network whom they could rely on when they needed various kinds of support. A participant (participant 3) stated, “I always had a support system but it’s larger now.”

Another participant shared the following:

It’s still just comforting, the only way I can describe it is just very comforting to know I can pick up the phone or I can take my walking poles and walk downstairs to one of my many friends and say, “I need a coffee and a hug” and that’s exactly what we get. [Participant 4]

Oasis participants began to feel similar to family. A participant shared the following:

There’s a gentleman here, a lovely gentleman, he’ll be [over 85], I believe. He had lived here for, I believe, 17 or 20 years and basically knew nobody. And Oasis is now what he refers to as his family. [Participant 7]

Positive Relationship With Oasis Coordinator

Oasis expanded older adults’ social networks, not only through interacting with other Oasis participants, but also through their relationship with the Oasis coordinator. Participants expressed positive feelings toward the Oasis coordinator and the significant role that they played. The support the coordinator provided to Oasis participants was appreciated. Throughout the interviews, it was clear that Oasis participants value the Oasis coordinator as part of their social network. A participant shared the following:

[The coordinator] is a lovely person, she did the exercises, and she was lots of fun...She calls me about once a week to see how I’m doing and that’s wonderful to think that somebody cares and says that
if you need anything or want to talk to somebody just give me a ring. [Participant 2]

Another participant echoed this view:

If they did have any problems they’d talk to [the coordinator] about things that was private between her and them, but they did know that if they needed help, they could go down and talk to her. [Participant 1]

Participants appreciated the work the coordinator put into activities: participant 11 stated, “The coordinator was excellent, she arranged all these ideas...[you] could talk to her about anything.”

The coordinator would try to get answers to participants’ questions:

Very helpful and very friendly and very informative...“oh, you want to know about that? I’ll look into that for you, I’ll see where I can get some answers for you.” [Participant 7]

This was echoed by another participant:

She was very good, and she was very friendly, and she’s very organized and helping everybody in lots of ways and helping get us information...If there was something we wondered about and she didn’t know, she would look into it for us and get back to us. She was just all around doing a wonderful job. [Participant 8]

The coordinator was a key part Oasis’ success:

What I would take away from this program is that it proved to me how valuable it is for a naturally occurring retirement community to be supported in this way, with an on-site coordinator. [Participant 3]

Community Partnerships

Oasis also provided participants with the opportunity to expand their social networks even further, to the broad community. Participants valued the guest speakers and community partners who were brought into Oasis. Oasis helped to connect older adults to their community by increasing their knowledge of services that are available. A participant said the following:

There was a speaker series and so we had, people from organizations such as Diabetes and the Hearing Society and, Canadian National Institute for the Blind...We had the Law Society from Queens, the Elder Law Clinic, the student who ran that came. Some of us got our wills redone who hadn’t redone them since we’d lost a spouse. [Participant 3]

This participant later added, “The Oasis program gave me access to guest speakers and organizations in an easy way.”

Oasis participants appreciated many of the topics discussed:

They had a speaker in once a month...they had so many different ones...I found that very interesting, and we were learning at the same time. [Participant 2]

This was echoed by another participant:

We have a lot of speakers come in about nutrition, about falling, balance, things like that, about pharmacists...They were really good, really informative. [Participant 11]

Oasis provided participants with the opportunity to expand and deepen their social networks. Participants indicated that other Oasis participants and the Oasis program coordinator provided them with support. Participating in Oasis also expanded their knowledge of community programs and services.

Individual Behaviors (Microlevel)—Oasis Activities: Something To Do, Someone To Do It With

Participants’ Experiences

As participants described their experiences with Oasis, it was evident that programming provided participants an opportunity to engage in meaningful activities with other people. Participants noted that they enjoyed doing fun activities in a warm social atmosphere. A participant shared the following:

You know, having coffee and basically laughing, which was I think the primary wonderful aspect of this whole thing. The laughter that was ensuing every time we went down there [to the common room], and the camaraderie was just absolutely heartwarming. [Participant 7]

Socialization

When asked to describe what they liked about the programming, most participants (12/13, 92%) described both activities and socialization opportunities the activities provided. For example, a participant (participant 13) shared, “I really looked forward to seeing the people, talking with the people and then doing the exercises.”

Similarly, another participant (participant 2) stated, “They had books with puzzles in and everything and we had coffee and we just socialized and talked, and laughed, and had fun.”

Talking and laughing were noted repeatedly by Oasis participants as core elements of the regularly scheduled programming. A participant stated the following:

There was socialization. Doing things that we had fun doing. We’d go down and play Bingo and we’d kill ourselves laughing about something stupid. [Participant 1]

New Activities

In addition to fun, Oasis programming provided participants with a wide variety of new activities to engage in. A participant described Oasis activities as “new adventures” and was appreciative that the coordinator was “always bringing in different things” (participant 8). Another participant (participant 12) stated, “I’ve learned about some things that I never knew before.”

Types of activities differed by site, but included various crafts, coloring, card and board games, line dancing, movie afternoons, painting, bingo, and social coffee hours, among others. Many participants (7/13, 54%) tried new activities, including different
types of crafts. A participant shared how much she enjoyed the new sewing class:

I love the sewing...that's new on my plate. Never sewed before! [Participant 10]

Participants noted that the variety provided by Oasis allowed them to choose activities that suited their interests. Another participant (participant 4) explained, “When I looked at the program, there was something for everyone.”

**Meaningful Ways to Spend Time**

Participants noted that Oasis activities provided them with a meaningful way to fill otherwise vacant time. They expressed the value of having activities to break up long days and get them out of their apartments. A participant (participant 7) shared, “Mondays I would usually just go down for a coffee for an hour or so, just enough to feel like you weren’t alone stuck in your house.”

As activities were organized and run by a combination of the coordinator, community partners, volunteers, and other Oasis participants, the calendars filled up quickly. A participant (participant 8) exclaimed, “...This is getting so busy that we have to pick and choose!”

When asked if she enjoyed being very busy with Oasis activities, a participant (participant 4) remarked, “Well yes, a lot of us did because some days, you know, the days are long.”

Participants also indicated that they liked the routine and looked forward to programming days. A participant (participant 2) shared, “I was always glad when Tuesday, Wednesday, and Thursday came because there were different things going on each day.”

Of note is the emphasis participants placed on the value of simply knowing that there were activities to look forward to in the future. The positive sense of anticipation these future activities provided was a mood booster when they were at home with nothing to do. A participant explained the following:

It's something to look forward to and I think that's a key. That when you're a senior that you're just not locked into your building, that there are things that you look forward to. [Participant 4]

Both future anticipation and fun had in the moment were motivating factors.

Oasis provided the participants with something to do and someone to do it with, which directly aligns with the program’s aim to promote healthy aging within NORCs.

**Health Outcomes—Impact of Oasis on Self-reported Mental Health and Well-being: Feeling and Coping With Life Better**

**Participants’ Experiences**

Oasis participants reported experiencing increased overall well-being. In addition, many participants (7/13, 54%) noted that having fun activities to do and look forward to improve their self-reported mental health. A participant shared the following:

Well of course I feel so much better because I find that I’m able to cope with life better than I used to...I guess if it hadn’t been for Oasis, I don’t know what I’d be doing or what I’d be like now. I would have probably jumped over this balcony or something...Joking [laughing]. [Participant 9]

**Well-being Before Oasis**

When asked about their self-perceptions of well-being before the existence of Oasis in their community, a participant (participant 4) explained, “On a scale of 1 to 10, there were days when I was just severely depressed and isolated.”

This participant later expanded on this feeling of depression by adding the following:

We’ve all had those days where I think why should I bother getting dressed, and brush my teeth, and have a shower? I’m not going see anybody. There would be weeks go by...where we would never see a soul, or talk to neighbors, or anything. [Participant 4]

Repeatedly, participants described their feelings before joining Oasis as “lonely,” “lonesome,” and “isolated.” After Oasis, their mental outlook improved. A participant shared the following:

It gave me something to get up for and something to look forward to everyday. And I didn’t even have that in my life before, it was a very, very depressing life. [Participant 5]

**Alleviating Loneliness**

There were participants who, when referencing their mental health struggles, discussed how the positive atmosphere of Oasis affected their feelings of loneliness and gave them an opportunity to re-engage. A participant shared the following:

I used to sit here. I was very lonely. Then once Oasis started it was a different story. It really brought me out and got me going again. [Participant 9]

This participant later added, “I don’t feel isolated anymore, and I don’t feel lonesome.”

This was echoed by several participants (11/13, 85%):

You don’t feel alone. You might live alone, but you don’t feel alone. [Participant 4]

You feel like you’re not all alone in the world. [Participant 3]

I didn’t feel so alone. I didn’t feel so isolated. [Participant 5]

**A Reason to Get Up**

Oasis clearly provides more than social support; it is a program that individuals heavily rely upon to help them cope with life and a variety of aging-related challenges. Participants stated that Oasis activities provided a reason to get up, get dressed, and proceed with the day when they otherwise may not have.

A participant shared the following:

That's where programs like Oasis make all the difference. They get people active...human beings
have to have a reason to get up in the morning. [Participant 8]

The same participant later shared the following:

People need that reason. If you don’t have a reason to get out of bed and get moving and you don’t have anyone to share what you’re doing or your thoughts with, it gets pretty depressing and pretty lonesome. [Participant 8]

Having “something to get up for” was reiterated by many participants (6/13, 46%). For instance, participant 5 stated, “Like I said it literally gave me a reason to get up.”

Another participant stated the following:

It was something to look forward to, something to get dressed for. I bought better clothes. [Participant 11]

Having a reason to get up and get dressed led to increased motivation to engage in other activities. A participant shared the following:

It was great, it was motivational, it was stimulating. It was enough, to say okay, you know, now that I’ve finished there, I’m going go out and do this because I’m already dressed. [Participant 4]

Oasis provided participants with a sense of purpose that was beneficial in helping them to maintain their mental health and well-being. When questioned about how Oasis affected their life, a participant shared the following:

It really gave [me]...a lift, you know, in my life, that I hadn’t felt for a long time...it really gave me a lift. [Participant 11]

Another participant stated the following:

It’s changed my everyday life...it’s just nice when you get up in the morning and you have your coffee...you have friends. [Participant 8]

When asked how Oasis affected their overall well-being, this participant described the impact by sharing, “I can’t say anything short of a hundred percent.”

Quality of Life

The main purpose of Oasis is to build healthy communities for older adults within NORCs that are supportive, foster independence, and improve overall quality of life, to support aging in place. Oasis participants reported improvement in their mental health, well-being, and overall quality of life. A participant, when asked about the impact Oasis had on her daily routine and her outlook on life, stated the following:

It has made me feel better, everyday...I think it does make you feel more like living. When you’re alone lots of time you’re thinking about “oh I’m going downhill,” you feel like you’re dying sometimes. And when you’re at Oasis it builds you up, and you feel more lively. [Participant 2]

Another participant (participant 3) said, “I have been happy to see the positive impact this program [Oasis] has had.”

Another participant (participant 4) commented, “I think my life is better for it with Oasis then it would have been without it.”

The powerful influence Oasis had on mental health and well-being was experienced by many participants (11/13, 85%). All the participants who were interviewed (13/13, 100%) commented on this important aspect of Oasis, highlighting the positive impact Oasis had on those who choose to participate in the program.

Discussion

Principal Findings

The purpose of this qualitative study was to explore the experiences of older adult participants who are attending the Oasis program at 4 unique NORCs in Ontario, Canada. To the best of our knowledge, this study offers the first exploration of experiences of older adults living in a NORC through the lens of the social network theory by Berkman et al [12], which describes how social networks influence health and well-being. Participants reported that their involvement in Oasis influenced their social networks, personal behaviors, and health. This study also adds to the NORC literature by examining how participation in Oasis, a NORC-SSP, influenced participants’ social networks, both within and beyond their NORCs.

Overall, 3 main themes were identified from the interviews with Oasis participants: expansion and deepening of social networks, Oasis activities (something to do, someone to do it with), and impact of Oasis on mental health and well-being (feeling and coping with life better).

At the mesolevel of the social network theory by Berkman et al [12], Oasis provided participants with an opportunity to expand and deepen their existing social networks. Participants reported that the number of individuals in their social networks increased, they made new friends, and previous friendships deepened. Participating Oasis members were comforted by the fact that friends were nearby and could be counted upon if they needed support. Oasis participants also discussed their positive relationship with the on-site Oasis coordinator. In addition, Oasis participants also spoke about the influence that Oasis had on their ability to make new connections with a variety of community organizations. To the best of our knowledge, this finding of the role of a NORC-SSPs in broadening of social networks beyond the walls of the NORC itself to professional services and organizations, has not been identified previously.

At the microlevel [12], participation in Oasis affected behavior. Oasis provided participants with opportunities to socialize and try new activities and meaningful ways to spend their time. Uniquely, Oasis participants reported trying and engaging in new activities.

Regarding the health outcomes of the social network theory by Berkman et al [12], participants reported the positive impact that Oasis had on their mental health and well-being, alleviating loneliness, and improving quality of life. Notably, and unique to this study, Oasis participants reported that Oasis activities provided them with a reason to get up every day and provided something to look forward to.
Comparison With Previous Studies
The 3 main themes found in this study are similar to those found in studies examining NORCs in the United States. For example, a study of NORCs in Cleveland, Ohio, found that, when asked what they liked best about their NORC, four themes emerged: interaction with neighbors and making friends, activities or services, resource coordinators, and choice or variety of activities or services [17].

At the mesolevel, Oasis participants reported making new friends and deepening existing friendships. Similarly, other studies of NORC-SSPs have also found that residents made friends with other individuals in their building and had more contact with their neighbors [17,19,23,24,35,36]. The social support provided by other Oasis participants is something that has also been experienced in other NORC-SSPs [19,23,36]. A positive relationship with the on-site Oasis coordinator was something mentioned by many Oasis participants (10/13, 77%). The importance of an on-site program coordinator in supporting NORC activities has also been identified in previous studies of structured NORC-SSPs [17,25,35] and, this highlights the importance of incorporating formal supports into NORC-SSPs, including paid staff, when appropriate. In their survey of 191 older adult participants of the Community Options NORC-SSP in Ohio, Anetzberger [17] indicated that participants repeatedly made unprompted references to the positive effect of the resource coordinator on their access to activities and resources and in empowering them in decision-making around programming.

At the microlevel, similar to several other NORC-SSPs, Oasis empowered its members, as they had the opportunities to choose what activities they wanted to engage in and felt they needed [10,17,37]. Other NORC studies have also reported that NORC-based activities provided residents with something to do [17,22,24,25]. For example, Anetzberger [17] found that NORC residents appreciated the variety of options available to them, including recreational and educational activities and monthly luncheons. In another study by Cohen-Mansfield et al [22], NORC residents engaged in leisure activities offered by a recreation service, such as trips, tours, and educational programs.

Participants reported that participating in Oasis positively influenced their health and well-being. Participating Oasis members reported decrease in isolation, which is something that has also been reported in other NORCs [10,22,23]. The decrease in depression reported by some Oasis participants has also occurred in other NORCs [22].

Strengths and Limitations
First, the results of this study are reflective of the experiences of Oasis participants, and therefore, may not be applicable to other NORC-SSPs. Second, Oasis participants who agreed to be interviewed may also be different from those who chose not to be interviewed, and this may have influenced the results. It is possible that participants with more positive experiences or those who were more involved in Oasis would be more willing to share their thoughts with interviewers. Third, data were gathered at only 1 point in time. It is possible that participants’ thoughts and experiences could change over time. However, participants reflected on their lives before Oasis and reported how things had changed since Oasis implementation. Fourth, we did not gather any demographic information beyond gender, and it is possible that ethnicity or age may have affected Oasis participants’ experiences. Fifth, of the 13 participants, 12 (92%) identified as women, leading to our inability to explore the presence of gender-specific experiences with the Oasis program. However, as most participants in Oasis identified as women, our results likely reflect the experiences of many Oasis participants.

Despite these limitations, this study has strengths. First, interviews were conducted with participants from 4 different Oasis sites, indicating that these positive experiences occurred in multiple NORCs in different locations with different program coordinators and reflecting different types of communities. Second, the questions asked provided an in-depth look at Oasis participants’ experiences, thus helping to deliver the rich data presented in this paper.

Future Directions
Future studies could expand on the thematic findings presented in this paper by further examining the contextual features of the various Oasis sites, such as the types of physical spaces available for activities or the demographic composition of the buildings, to deepen the understanding of how such features shape experiences of the program. Future studies will also examine how physical, social, and psychological factors change over time among Oasis participants.

Conclusions
On the basis of these findings, Oasis is an effective aging-in-place model because it has expanded and deepened participants’ social networks, increased participants’ engagement in activities, and positively affected participants’ mental health and well-being. Oasis provides a model that has potential to be implemented in other NORCs, as it has been successfully implemented in low-rise apartment buildings, midrise apartment buildings, and mobile home communities. Implementing Oasis programs in additional communities and buildings could address some of the challenges of aging in place and improve older adults’ health and well-being.

Acknowledgments
The authors thank the study participants and members of the Oasis Study Collaborative. The Oasis expansion was funded by the Baycrest Centre for Aging and Brain Health Innovation, Ontario Ministry of Health and Long-Term Care, and Ontario Ministry for Seniors and Accessibility. Each grant was specifically designated to support the expansion to a particular Oasis community.
Data Availability
The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions
CD, VD, SP, CH, C McGrath, and DLR were involved in the design and implementation of the Oasis expansion. CD, VD, and SP designed this qualitative study. CC, MG, and LW transcribed the interviews. CC, MG, C Mills, CD, SP, and LW performed the thematic analysis. CC, MG, and LW drafted the results. C Mills wrote the manuscript. All authors reviewed and edited the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Semistructured interview guide.

References


Abbreviations

**NORC:** naturally occurring retirement community

**NORC-SSP:** naturally occurring retirement community supportive services program

**Oasis:** Oasis Senior Supportive Living
Enhancing Food Intake Tracking in Long-term Care With Automated Food Imaging and Nutrient Intake Tracking (AFINI-T) Technology: Validation and Feasibility Assessment

KAYLEN PFISTERER, BSc, MSc, PhD; ROBERT AMELARD, BSc, MASc, PhD; JENNIFER BOGER, BSc, MASc, PhD; HEATHER KELLER, BASc, MSc, PhD; AUDREY CHUNG, BASc, MASc, PhD; ALEXANDER WONG, BASc, MASc, PhD

1 Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada
2 Waterloo AI Institute, Waterloo, ON, Canada
3 Schlegel-University of Waterloo Institute for Aging, Waterloo, ON, Canada
4 Centre for Digital Therapeutics, Techna Institute, University Health Network, Toronto, ON, Canada
5 KITE-Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
6 Department of Kinesiology and Health Sciences, University of Waterloo, Waterloo, ON, Canada

* these authors contributed equally

Corresponding Author:
KAYLEN PFISTERER, BSc, MSc, PhD
Department of Systems Design Engineering
University of Waterloo
200 University Ave W
Waterloo, ON, N2L 3G1
Canada
Phone: 1 519 888 4567
Email: kaylen.pfisterer@uhn.ca

Abstract

Background: Half of long-term care (LTC) residents are malnourished, leading to increased hospitalization, mortality, and morbidity, with low quality of life. Current tracking methods are subjective and time-consuming.

Objective: This paper presented the automated food imaging and nutrient intake tracking technology designed for LTC.

Methods: A needs assessment was conducted with 21 participating staff across 12 LTC and retirement homes. We created 2 simulated LTC intake data sets comprising modified (664/1039, 63.91% plates) and regular (375/1039, 36.09% plates) texture foods. Overhead red-green-blue-depth images of plated foods were acquired, and foods were segmented using a pretrained food segmentation network. We trained a novel convolutional autoencoder food feature extractor network using an augmented UNIMIB2016 food data set. A meal-specific food classifier was appended to the feature extractor and tested on our simulated LTC food intake data sets. Food intake (percentage) was estimated as the differential volume between classified full portion and leftover plates.

Results: The needs assessment yielded 13 nutrients of interest, requirement for objectivity and repeatability, and account for real-world environmental constraints. For 12 meal scenarios with up to 15 classes each, the top-1 classification accuracy was 88.9%, with mean intake error of −0.4 (SD 36.7) mL. Nutrient intake estimation by volume was strongly linearly correlated with nutrient estimates from mass ($r^2$=0.92-0.99), with good agreement between methods ($\sigma$=-2.7 to -0.01; 0 within each of the limits of agreement).

Conclusions: The automated food imaging and nutrient intake tracking approach is a deep learning–powered computational nutrient sensing system that appears to be feasible (validated accuracy against gold-standard weighed food method, positive end user engagement) and may provide a novel means for more accurate and objective tracking of LTC residents’ food intake to support and prevent malnutrition tracking strategies.

(JMIR Aging 2022;5(4):e37590) doi:10.2196/37590
KEYWORDS
long-term care; automated nutrient intake; convolutional neural network; food segmentation; food classification; depth imaging; deep learning; collaborative design; aging; food intake

Introduction

Background
Malnutrition leads to high morbidity [1] and low quality of life [2]. In the United States, malnutrition imparts >4 times high odds of hospitalization and an average of US $21,892 more in total charges per stay [3]. It is clear that nutritional status has multidomain effects with both fiscal and clinical ramifications and should be monitored. Older adults (aged ≥65 years) living in long-term care (LTC) homes are especially nutritionally vulnerable, in part owing to low food intake [4]. More specifically, in Canada, 54% of LTC residents are either malnourished or at risk for malnutrition [5]. This is higher than global estimates, ranging from 19% to 42% (37 studies; 17 countries) [6]. Additional independent risk factors for malnutrition are eating challenges and increased cognitive impairment [4,7], which describes between 47% to 90% of the Ontario LTC population [8,9]. Thus, tracking and preventing poor food intake is essential for supporting healthy aging.

However, there is a lack of objective and quantitative tracking methods for food and fluid intake, especially for centralized intake tracking by proxy (ie, multiple staff tracking a set of residents’ intakes). Registered dietitian (RD) referrals are triggered and nutritional support system effectiveness is monitored based on nutritional assessment best practices including unintentional weight loss and usual low intake of food [10]. Resident food and fluid intake charting completed by either personal support workers or nursing assistants captures intake across a meal via visual assessment within 25% incremental proportions at the end of the meal, but may be completed hours later owing to multiple competing priorities during mealtime. Therefore, owing to inconsistency and subjectivity in charting methods, approximately half of residents who would benefit from an intervention are missed [11,12].

Furthermore, there is a lack of trust in current methods because they are known to have poor accuracy and validity [13,14], thus limiting clinical utility. However, it raises awareness to some extent, regardless of whether the measurements are inaccurate (eg, food spills). Measuring food intake is a proxy for nutritional status; however, it provides a sense of why something may be going wrong (in combination with biomarkers). Better, more reliable measurements will enable more meaningful assessment of probing when, how, and why something may be going wrong to better inform intervention strategies, and care providers have expressed a desire to leverage high-quality data, provided they are reliable and trustworthy [15].

Objectives
Automated tools may provide a palatable solution that removes subjectivity and has higher accuracy than human assessors. This may also enable time-efficient measurement of food intake at the energy, macronutrient, and micronutrient levels [15]. More specifically, in LTC, it is desirable to have a high level of detail including intake breakdown for each item consumed (not averaged across a plate) [15]. To estimate food intake and nutrient consumption, 4 main questions must be answered: where is there food on a plate (segmentation), which foods are present (classification), how much food was consumed (preprandial and postprandial volume estimation), and what was the estimated food and nutrient intake? This study builds on previous studies exploring where food is and how much food was consumed at a bulk intake level by leveraging a specialized food segmentation method powered by deep learning for automated segmentation, moving from bulk food segmentation to nutritional estimation with a few additional steps modularized for systematic error assessment [16]. Here, we focused on which foods are present and how much food was consumed for enabling assessment of what was the estimated food intake at the nutrient level.

The purpose of this study was to describe the final stage of feasibility testing of the automated food imaging and nutrient intake tracking (AFINI-T) system comprising pixel-wise food classification and nutrient linking through intake prediction, for providing food and nutrient intake estimation with specific feasibility considerations for use in LTC. Our proposed AFINI-T technology measures food intake compared against gold-standard ground truth weighed food records, addresses automatic segmentation with integrated red-green-blue-depth (RGB-D) assessments, was evaluated in both regular texture foods (RTFs) and modified texture foods (MTFs), and describes the valence of the system within the user context.

Methods

This study used an iterative action research design, blending mixed methods needs assessment with technical implementation and experimental evaluation.

Ethics Approval
This study received ethics clearance from the University of Waterloo’s Office of Research Ethics Board (23124).

End User Data to Shape Technological Requirements—A Case Study
Insights motivating the technical approach described in this paper were gathered through interviews and workshop discussions with Schlegel Village team members during our previous user study, but not included in the paper [15]. Overall, 2 interviews (an RD nutrition research expert and an RD working in LTC) and discussion with experts during a workshop were conducted. The workshop included 21 participants representing 12 LTC and retirement homes who were recruited through self-enrollment, including an administrative assistant, chef, dining lead (similar to a dining room manager), director of recreation, dietary aides, neighborhood coordinator, recreation assistant, restorative care, senior nurse consultant, directors and assistant directors of food services, registered nurse, and personal support workers [15]. Participants identified potential
barriers to uptake including time and whether the level of detail is desired or seen as valuable. Qualitative results from interviews and workshops with end users illuminated the following user needs, which, guided by grounded theory [17], were translated into design requirements for application within the LTC context.

Experimental Procedure—AFINI-T’s Technical Approach

Data Collection

As described in the study by Pfisterer et al [16], data were collected in an industrial research kitchen at the Schlegel University of the Waterloo Research Institute for Aging’s Centre of Excellence for Innovation in Aging. This kitchen was modeled after industrial research kitchens found in LTC homes. RGB-D images were acquired using Intel RealSense (F200), with a depth resolution of 640×480 pixels. A sequence of 10 depth images was acquired for each plate and averaged to reduce pixel noise. An optical imaging cage was constructed to enable top-down image capture, as described in the study by Pfisterer et al [16]. The camera was connected to a computer for data acquisition, and plates were weighed at a nearby weigh station. Figure 1 shows examples of the data sets used for training the convolutional autoencoder and food classification network, which are described in detail in the following subsections.

Figure 1. Example images in the data sets used for training the convolutional autoencoder (ie, UNIMIB+ [UNIMIB2016 with additional green representation]) [18,19] and food classification training and testing on modified and regular texture foods. A: UNIMIB+; B: Modified texture foods; C: Regular texture foods.

RTF and MTF Data Sets

We used our RTF data set (9 foods; 9 classes; 375 images) and our MTF data set (47 foods; 93 classes; 664 images). Table 1 provides an overview of data set characteristics, and a summary of all food items imaged can be found in Table 2. Our RTF data set comprised 3 meal plates, each consisting of 3 foods imaged at every permutation of 25% simulated intake. Our MTF data set consisted of 134 food samples representing 47 foods, each consisting of a set of at least one purée and one minced texture food. Each sample was imaged 5 times by progressively removing food, with the exception of 4.5% (6/134) of the samples consisting of 4 each with 1 lost image.

For each food item, 1 full serving was defined by the nutritional label portion size (RTF data set) or the recipe-defined portion size received from the kitchen and was weighed to the nearest 1 g using an Ohaus Valor Scale.

For the RTF data set, in which a serving size was referenced using volume, that volume of food (eg, corn) was weighed, and the mass was used thereafter. As manufacturers supply nutritional information for minerals as percentage of daily value (assuming a 2000-calorie diet), for the RTF data set, minerals were reported similarly. For more details on conversion, refer to Table S1 in Multimedia Appendix 1 [20]. Mass in grams was used to define all serving sizes.

For the MTF data set, we expanded our original MTF data set [16] with additional examples (that did not include recipes) for further segmentation and volume estimation analysis. Nutritional analysis was conducted on a subset of 47.3% (314/664) of the images. As nutritional information was provided according to mass, we converted from mass to volume. To accomplish this, we calculated the food’s density to convert by using the full plate’s true volume (in mL) with its mass (in grams). This enabled the scaling of nutritional information using the RTF data set pipeline for validating these findings using mass; it was not required for the system to operate.
Table 1. Overview of data set characteristics. The UNIMIB+ data set was used for training and validation [18,19].

<table>
<thead>
<tr>
<th>Data set overview</th>
<th>UNIMIB+</th>
<th>RTF&lt;sup&gt;b&lt;/sup&gt;</th>
<th>MTF&lt;sup&gt;c&lt;/sup&gt;</th>
<th>RTF+MTF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of images</td>
<td>1214</td>
<td>375</td>
<td>664</td>
<td>1039</td>
</tr>
<tr>
<td>Number of samples</td>
<td>N/A&lt;sup&gt;d&lt;/sup&gt;</td>
<td>3</td>
<td>134</td>
<td>137</td>
</tr>
<tr>
<td>Number of classes</td>
<td>76</td>
<td>9</td>
<td>93</td>
<td>102</td>
</tr>
<tr>
<td>Number of foods represented</td>
<td>76</td>
<td>9</td>
<td>47</td>
<td>56</td>
</tr>
<tr>
<td>Number of foods with recipes</td>
<td>N/A</td>
<td>9</td>
<td>27</td>
<td>36</td>
</tr>
</tbody>
</table>

<sup>a</sup>UNIMIB+: UNIMIB2016 with additional green representation.
<sup>b</sup>RTF: regular texture food.
<sup>c</sup>MTF: modified texture food.
<sup>d</sup>N/A: not applicable.
Table 2. List of foods in the RTF\textsuperscript{a} and MTF\textsuperscript{b} data sets used for testing the AFINI-T\textsuperscript{c} system.

<table>
<thead>
<tr>
<th>Food component</th>
<th>RTF with recipes</th>
<th>MTF with recipes</th>
<th>Additional MTF with segmentations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grains</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cheese tortellini with tomato sauce</td>
<td>Bow-tie pasta with carbonara sauce</td>
<td>Basmati rice</td>
</tr>
<tr>
<td></td>
<td>Oatmeal</td>
<td>Macaroni salad</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Whole wheat toast</td>
<td>Vegetable rotini</td>
<td>—</td>
</tr>
<tr>
<td><strong>Vegetables and fruits</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Corn</td>
<td>Asian vegetables</td>
<td>Beet and onion salad</td>
</tr>
<tr>
<td></td>
<td>Mashed potatoes</td>
<td>Baked polenta with garlic</td>
<td>Cantaloupe chunks</td>
</tr>
<tr>
<td></td>
<td>Mixed greens salad</td>
<td>California vegetables</td>
<td>Green beans with pimento</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Greek salad</td>
<td>Grilled vegetable salad</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Mango and pineapple</td>
<td>Roasted cauliflower</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Red potato salad</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Sauteed spinach and kale</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Seasoned green peas</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Stewed rhubarb and berries</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Strawberries and bananas</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Sweet and sour cabbage</td>
<td>—</td>
</tr>
<tr>
<td><strong>Proteins</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Meat loaf</td>
<td>Baked basa</td>
<td>Bean and sausage strata</td>
</tr>
<tr>
<td></td>
<td>Scrambled egg</td>
<td>Braised beef liver and onions</td>
<td>Grilled lemon and garlic chicken</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Braised lamb shanks</td>
<td>Pork tourtiere</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Hot dog wiener</td>
<td>Roast beef with miracle whip</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Orange ginger chicken</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Salisbury steak and gravy</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Teriyaki meatballs</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Tuna salad</td>
<td>—</td>
</tr>
<tr>
<td><strong>Mixed dishes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Oatmeal cookie</td>
<td>Barley beef soup</td>
<td>Black bean soup</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Blueberry coffee crumble cake</td>
<td>Broken glass parfait (mixed gelatin)</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Eggplant parmigiana</td>
<td>Butternut squash soup</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>English trifle</td>
<td>Cranberry spice oatmeal cookie</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>Lemon chicken orzo soup</td>
<td>Lemon meringue pie</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>—</td>
<td>Peach jello</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>—</td>
<td>Pear crumble cake</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>—</td>
<td>Roast beef with miracle whip on whole wheat</td>
</tr>
<tr>
<td></td>
<td>—</td>
<td>—</td>
<td>Turkey burger on wheat bun</td>
</tr>
</tbody>
</table>

\textsuperscript{a}RTF: regular texture food.  
\textsuperscript{b}MTF: modified texture food.  
\textsuperscript{c}AFINI-T: automated food imaging and nutrient intake tracking.  
\textsuperscript{d}There were varying numbers of items in the data sets.
Training Data Set

We expanded the UNIMIB2016 data set (1027 tray images; 73 classes) [19] with additional examples from the FoodX-251 data set [18] to train the convolutional autoencoder (described in detail in the Automation With a Convolution Autoencoder section). We discovered that UNIMIB2016 had an underrepresentation of green foods compared with what is served in LTC, which affected the autoencoder’s ability to differentiate among all colors and textures. To address this difference in the canteen images from the original UNIMIB2016, we augmented the training data set by adding 91 examples of lettuce, 91 examples of peas, and 89 examples of spinach from the FoodX-251 food data set [18]. Plates with plastic packaging (84/1027, 8.17%) were removed, as they confounded food feature learning and were not representative of LTC plates. We refer to this as the UNIMIB+ (UNIMIB2016 with additional green representation) data set (1214 images; 76 classes). Figure 2 shows the effect of this underrepresentation of green by its inability to reconstruct a vibrant hue across the autoencoder’s decoder output trained solely on the UNIMIB2016 data set for validation examples. The autoencoder was able to converge to low validation loss on the UNIMIB+ data set. Empirically, this resulted in greens appearing greener, reds appearing redder, and yellows and whites appearing less murky, as shown in the UNIMIB+ examples compared with the UNIMIB2016 in Figure 2. This suggests that the addition of the green samples enabled the autoencoder to learn good food representations; encode features more deeply; and align more closely with how a human would perceive the foods, which is a crucial point for the LTC application.

Figure 2. Effect of underrepresentation of green foods in the UNIMIB2016 database on decoder output on segmented food from plates. The decoder output from the autoencoder trained on the UNIMIB+ (UNIMIB2016 with additional green representation) data set in the bottom appears less murky and more vibrant, with truer perceived greens than the UNIMIB2016 counterpart in the middle.

Computational Methods

The following sections describe how the segmentation strategy was refined compared with our initial work [16], the general food or no food classification approach, followed by system automation using a convolutional autoencoder. Figure 3 shows the processing pipeline from image acquisition to classified food pixels.

Figure 3. System diagram showing the processing pipeline from image acquisition to food classification. EDFN-D: depth-refined encoder-decoder food network; RGB-D: red-green-blue-depth; UNIMIB+: UNIMIB2016 with additional green representation.
Refined Segmentation Strategy

Modifications to the training process were made to enhance network performance. We introduced early stopping criteria to halt training early to avoid overfitting, yielding a network that was trained over fewer epochs than one that is overtrained and outputting a pixel-level image mask as food or no food with calibrated depth [16]. Volume consumed was mapped onto nutritional information for intake approximation. These nutrient-level intake estimates were validated against the ground truth nutritional information obtained through the weighed food method.

General Classification Approach

Here, the UNIMIB+ data were used to train the autoencoder. Using the autoencoder’s trained weights, the last layer of the autoencoder (120x160x3) was spliced to use the feature map as a latent feature extractor for classification (refer to Figures 2 and 4 for system diagram and network architecture). This approach was modeled based on our previous study on classification for predicting relative nutritional density of a dilution series of commercially prepared purées [21], because MTF comprises 63.91% (664/1039) of our testing data set and 47% of the LTC population receives MTF [22].

Figure 4. Convolutional autoencoder network for learned feature representation and in the context of classification. (A) The architecture for learning feature representation: an input image is given and the output is a reconstruction of that image. Training minimized the error between input and output images; we used mean squared error loss with Adam optimizer, learning rate of 0.0001, and batch size of 32. The early stop criteria used were change of loss of <0.0001 and patience of 5 epochs. The autoencoder was spliced; weights were frozen; and only a classification layer for nc classes was trained for classification, where nc is the number of food items for meal c. We used categorical cross-entropy (ignoring background pixels) loss, with Adam optimizer and learning rate of 0.1. The early stop criteria used were a change of loss of <1x10^-5 and patience of 5 epochs. We used 70%:30% train to validation split of augmented data. The data were augmented by generating 300 images from the full set of plates and applying random flips, rotations, and increased or decreased contrast. The outputs are distinct classes, which were mapped onto the meal-specific classifier (in this example, as ravioli [blue], salad [green], and oatmeal cookie [yellow]). ReLU: rectified linear unit; RGB: red-green-blue.

Automation With a Convolutional Autoencoder

We report nutrient intake accuracy using the automated system (ie, the automated classification case) to enhance pragmatic feasibility (ie, reduced user input). For this automated approach, we developed a semantic segmentation network with a convolutional autoencoder feature extractor for classification of foods, which was roughly inspired by a highly successful convolutional neural network (CNN), the Visual Geometry Group network [23], in Tensorflow 2.3.0. For a given meal or time of day, we fed the masked output from the depth-refined encoder-decoder food network (EDFN; food or no food detector, as described in the study by Pfisterer et al [16]) into the convolutional autoencoder. CNNs encode spatial information, and given how food has differing degrees of cohesion, we felt that the context of spatial information will be an asset. In addition, we sought to extract latent features via a method requiring a round of training offline. For classification, a small classification layer was appended and trained for each meal using a priori information about the meal items offered. Loss for the autoencoder network was computed as pixel-wise mean squared error between the input and reconstructed output; therefore, they did not require labeled training data.

We trained an autoencoder to be a feature extractor using the UNIMIB+ data set consisting of 1214 images. Data were divided into 70% training and 30% validation. Training was performed using the Adam optimizer with batch size of 32, mean squared error loss, and early stopping (<0.0001 validation loss change) with 5-epoch patience. Only food pixels were used in the loss calculation using the ground truth masks. After training, the convolutional autoencoder network was spliced before the final 1x1 convolution block to produce original resolution 16-channel latent feature vectors. The weights of this network were frozen and used as a feature extractor for the classification training.

Given that there are many food options and as new meals are planned, we needed a flexible modular approach, which also enables us to use only 1 labeled example per item; the AFINI-T
method uses only 1 full reference portion to classify foods and infer intake. For nutritional intake estimation, we leveraged the homes’ known nutritional information from menu planning software (or supplied by the manufacturer) to link proportional nutrient intake. We assumed that recipes were followed exactly. Denoting the number of menu items for meal \( m \) as \( n_c \), the classification network for meal \( c \) was built by appending \( n_c \times 1 \times 1 \) convolution kernels onto the feature extractor network. The meal full portion training data were constructed by augmenting the full set of plates by applying random flips, rotations, and increased or decreased contrast, yielding 300 augmented instances of the meal. We used 1 reference image (the full portion image) to learn what each class looked like and then mapped subsequent instances onto these prelabeled classes by grouping all the full plates of food for a given meal into the training set. The data were divided into 70% training and 30% validation. Training was performed using the Adam optimizer with batch size of 32, categorical cross-entropy loss, and early stopping (\( <1 \times 10^{-5} \) validation loss change) with 5-epoch patience. Only food pixels were used in the loss calculation using the ground truth masks. Finally, we applied ground truth labels to the full portion plate to link the proper proportional intake at the nutrient level and assess the accuracy of the intake estimates compared with gold-standard weighed food approach.

**Nutrient Intake Association**

This step comprised three general stages: (1) determine the relative consumption of each food item compared with a full reference portion, using food volume estimation from the depth maps; (2) compare relative consumption with nutritional information, to infer nutritional intake for each item; and (3) sum the inferred nutritional intake for each item across a plate for estimation of total nutrition consumed during a meal (for MTF, this was across the plate of one food item).

**Statistical Analyses**

**System Accuracy**

Segmentation accuracy was assessed using intersection over union (IOU). Classification accuracy was described using top-1 accuracy and summarized using per-meal classifiers. Bulk intake accuracy (ie, class-agnostic, overall food volume intake) was assessed using mean absolute error (mL) and 3D, % intake error, described in the study by Pfisterer et al [16] in which intake error was calculated for volume (3D) data relative to the full portion. All values are reported as mean (SD). Nutrient intake accuracy was assessed using the fully automated classification approach (ie, without updating misclassified regions) to evaluate nutrition intake accuracy and is reported as mean (SD) and percentage error.

**Validating Nutrient Intake Estimation Against Weighed Food Records**

All data were analyzed using MATLAB 2020b (MathWorks). Linear regression was used to determine the goodness of fit through the degree of correlation with \( r^2 \) to summarize the extent to which nutritional intake information from weighed food mass is related to estimated nutritional information from food volume. Bland-Altman analysis was used to describe the level of agreement between nutritional intake information from weighed food mass compared with intake volume using mean average (\( G \)) and bias (\( \mu \)) between methods [24]. Several nutrients of concern in the RTF data set were reported in percentage daily value (ie, calcium, iron, vitamin B6, vitamin C, and zinc). We converted these values to absolute values to match the MTF data set using the 2005 Health Canada reference values for elements and vitamins. Where there was a difference across age, we used the reference for age >70 years; where there was a difference in requirement by sex, we used the average value.

**Results**

**Overview**

This study focused on the characterization of changes in volume at the whole plate level for bulk intake estimation, reporting degree of consumption (ie, proportion of food consumed) and nutritional intake estimation using a nutritional lookup table at the food item and whole plate level. Specific needs informed by workshop and interview responses included the following:

1. The system shall consider evidence-based and practice-relevant priority nutrients (output: 13 nutrients of interest—macronutrients: calories, carbohydrates, fats, fiber, and protein and micronutrients: calcium, iron, sodium, vitamin B6, vitamin C, vitamin D, vitamin K, and zinc).
2. The system shall support current workflow in which the dietitian is the gatekeeper:
   - The system shall facilitate automated, objective, intake estimates.
   - The system shall facilitate dietitian referrals by providing repeatable nutrient-specific intake insights.
3. The system shall work independently of internet connection.
4. The system shall incorporate real-world constraints and parameters:
   - The system shall include a salient feature extractor that can be trained in advance and supports real-time use.
   - The system shall use a classification method that is light in weight for mobile app use.
   - The system shall include an easily updatable classifier to account for a priori menu plans considering the time of day and therapeutic diet.

The following quantitative results provide an overview of the AFINI-T system’s food and nutrition intake estimation system including segmentation, classification, volume estimation, bulk intake, and nutrient intake accuracies.

**Segmentation Accuracy**

Table 3 provides an overview of segmentation accuracy. Generally, results represent 2 types of meal scenarios: multiple RTF data set on a plate and single MTF data set on a plate. The RTF data set had 9 unique foods across 375 simulated intake plates. The MTF data set foods were prepared by the LTC kitchen and included 93 unique foods including both purées and minced foods across 664 simulated intake plates. Across the RTF and MTF data sets, there are 102 classes represented in 1039 simulated intake plate images. Segmentation accuracy
was good with an average IOU of 0.879 across the RTF and MTF data sets (Table 3). Segmentation accuracy ranged from 0.823 for the MTF data set at lunch to 0.944 for the RTF data set for breakfast. From the perspective of IOU, the MTF data set was more poorly segmented by the depth-refined EDFN; however, consistent with the study by Pfisterer et al [16], the degree of visual-volume discordance was high for modified texture diets and is discussed further in the Volume Estimation Accuracy section.

Table 3. Average segmentation and classification accuracies within and across data sets.

<table>
<thead>
<tr>
<th>Data set and meal</th>
<th>Classes (N=102), n</th>
<th>Images (N=1039), n</th>
<th>Segmentation accuracy (IOU), mean (SD)</th>
<th>Classification accuracy (top 1), %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RTF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breakfast</td>
<td>3</td>
<td>125</td>
<td>0.944 (0.019)</td>
<td>93.5</td>
</tr>
<tr>
<td>Lunch</td>
<td>3</td>
<td>125</td>
<td>0.919 (0.033)</td>
<td>93.5</td>
</tr>
<tr>
<td>Dinner</td>
<td>3</td>
<td>125</td>
<td>0.928 (0.019)</td>
<td>95.1</td>
</tr>
<tr>
<td><strong>MTF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 1—lunch</td>
<td>5</td>
<td>25</td>
<td>0.841 (0.123)</td>
<td>89</td>
</tr>
<tr>
<td>Day 1—dinner</td>
<td>15</td>
<td>90</td>
<td>0.823 (0.099)</td>
<td>70.2</td>
</tr>
<tr>
<td>Day 2—lunch</td>
<td>12</td>
<td>74</td>
<td>0.863 (0.118)</td>
<td>70.6</td>
</tr>
<tr>
<td>Day 2—dinner</td>
<td>12</td>
<td>90</td>
<td>0.840 (0.122)</td>
<td>64.9</td>
</tr>
<tr>
<td>Day 3—lunch</td>
<td>10</td>
<td>85</td>
<td>0.834 (0.132)</td>
<td>80.4</td>
</tr>
<tr>
<td>Day 3—dinner</td>
<td>15</td>
<td>109</td>
<td>0.859 (0.100)</td>
<td>70.4</td>
</tr>
<tr>
<td>Day 4—lunch</td>
<td>9</td>
<td>60</td>
<td>0.871 (0.113)</td>
<td>72.2</td>
</tr>
<tr>
<td>Day 4—dinner</td>
<td>10</td>
<td>90</td>
<td>0.837 (0.107)</td>
<td>67.8</td>
</tr>
<tr>
<td>Day 5—lunch</td>
<td>5</td>
<td>41</td>
<td>0.881 (0.117)</td>
<td>87.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data set and meal</th>
<th>Classes (N=102), n</th>
<th>Images (N=1039), n</th>
<th>Segmentation accuracy (IOU), mean (SD)</th>
<th>Classification accuracy (top 1), %</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RTF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breakfast</td>
<td>3</td>
<td>125</td>
<td>0.944 (0.019)</td>
<td>93.5</td>
</tr>
<tr>
<td>Lunch</td>
<td>3</td>
<td>125</td>
<td>0.919 (0.033)</td>
<td>93.5</td>
</tr>
<tr>
<td>Dinner</td>
<td>3</td>
<td>125</td>
<td>0.928 (0.019)</td>
<td>95.1</td>
</tr>
<tr>
<td><strong>MTF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 1—lunch</td>
<td>5</td>
<td>25</td>
<td>0.841 (0.123)</td>
<td>89</td>
</tr>
<tr>
<td>Day 1—dinner</td>
<td>15</td>
<td>90</td>
<td>0.823 (0.099)</td>
<td>70.2</td>
</tr>
<tr>
<td>Day 2—lunch</td>
<td>12</td>
<td>74</td>
<td>0.863 (0.118)</td>
<td>70.6</td>
</tr>
<tr>
<td>Day 2—dinner</td>
<td>12</td>
<td>90</td>
<td>0.840 (0.122)</td>
<td>64.9</td>
</tr>
<tr>
<td>Day 3—lunch</td>
<td>10</td>
<td>85</td>
<td>0.834 (0.132)</td>
<td>80.4</td>
</tr>
<tr>
<td>Day 3—dinner</td>
<td>15</td>
<td>109</td>
<td>0.859 (0.100)</td>
<td>70.4</td>
</tr>
<tr>
<td>Day 4—lunch</td>
<td>9</td>
<td>60</td>
<td>0.871 (0.113)</td>
<td>72.2</td>
</tr>
<tr>
<td>Day 4—dinner</td>
<td>10</td>
<td>90</td>
<td>0.837 (0.107)</td>
<td>67.8</td>
</tr>
<tr>
<td>Day 5—lunch</td>
<td>5</td>
<td>41</td>
<td>0.881 (0.117)</td>
<td>87.8</td>
</tr>
</tbody>
</table>

*a* There were no samples for day 5—dinner.

*b* IOU: intersection over union.

*c* RTF: regular texture food.

*d* MTF: modified texture food.

**Classification Accuracy**
As shown in Table 3, classification accuracy was high for the RTF data set, with top-1 accuracy (ie, the most likely class) ranging from 93.5% for breakfast and lunch to 95.1% for dinner. However, the RTF data set had only 3 classes per meal; therefore, it was a less challenging classification problem compared with a great number of classes to differentiate among, especially when considering the MTF data set had less texture variance. In contrast, the MTF data set top-1 accuracy ranged from 64.9% on day 2—dinner with 12 classes to 89% on day 1—lunch with 15 classes.

**Volume Estimation Accuracy**
Low-density foods pose challenges to depth scanning systems. Here, volume estimation was within tolerance with food volume error of 2.5 (SD 9.2) mL, and low-density foods (eg, salad) have the largest food volume error seen for RTF: lunch of −10.1 (SD 22.2) mL. A similar issue of low-density foods is seen through the 3D, % absolute error intake of 14.4% (SD 13.1%), which we suspect is owing to the air pocket below some of the pieces of toast that are placed at a tangential angle to the plate or when 2 pieces are stacked with overhang, as shown in Figure 5. This can be considered as one of the classic examples of the occlusion conundrum with the imaging limitation of collection from an overhead view. This is an example of where segmentation can be performed perfectly, but will translate to volume estimation errors.
Figure 5. The occlusion conundrum, as demonstrated by stacked toast with an overhang. As volumetric food estimation is based on pixel-wise classification, the pixels of the overhang are assumed to contain toast. This is a limitation to overhead imaging and provides a simplified example of low-density foods (e.g., salad) as does rigid toast placed as an inclined plane. This is seen in the depth images; bright parts denote pixels close to the camera (i.e., high food pixels). We see a gradient from low to high near the tip with a similar, but less obvious, trend in the third depth image. The depth map range was adjusted to exemplify the toast height.

Bulk Intake Accuracy

Table 4 summarizes the bulk intake accuracy within and across data sets. Compared with the study by Pfisterer et al [16], for this iteration, we incorporated more representation of green in the UNIMIB+ data set for training and validation and introduced a more optimal stop criteria for training for segmentation. In the study by Pfisterer et al [16], we saw that the mean absolute volume error was 18 (SD 50) mL for RTF and 2.3 (SD 3.2) mL for MTF and mean volume intake error was 130.2 (SD 154.8) mL and 0.8 (SD 3.6) mL for RTFs and MTFs, respectively. Here, accuracy is higher with mean absolute food volume error of 6.6 (SD 13.6) mL for RTFs and 2.1 (SD 3.1) mL for MTFs. Similarly, the bulk intake accuracy was higher, with mean absolute intake error greatly reduced for the RTFs (39.9, SD 39.9 mL), but slightly higher for MTFs (6, SD 5.6 mL). The higher degree of visual-volume discordance for MTFs compared with RTFs is again corroborated in Table 3, with mean food volume error of 3.8 (SD 8.8) mL and higher mean volume error for the RTF data set (6.6, SD 13.6 mL) than for the MTF data set (2.1, SD 3.1 mL).
Table 4. Bulk intake accuracy within and across data sets\(^a\).

<table>
<thead>
<tr>
<th>Data set and meal</th>
<th>Classes (N=102), n</th>
<th>Images (N=1039), n</th>
<th>Food volume error</th>
<th>Bulk intake accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Absolute error (food volume; mL), mean (SD)</td>
<td>Absolute error (intake; mL), mean (SD)</td>
</tr>
<tr>
<td><strong>RTF(^b)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breakfast</td>
<td>9</td>
<td>375</td>
<td>6.6 (13.6)</td>
<td>39.9 (39.9)</td>
</tr>
<tr>
<td>Lunch</td>
<td>3</td>
<td>125</td>
<td>3 (4.1)</td>
<td>17 (14.3)</td>
</tr>
<tr>
<td>Dinner</td>
<td>3</td>
<td>125</td>
<td>11 (21.7)</td>
<td>76.1 (48.5)</td>
</tr>
<tr>
<td><strong>MTF(^c)</strong></td>
<td>93</td>
<td>664</td>
<td>2.1 (3.1)</td>
<td>6 (5.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 1—lunch</td>
<td>5</td>
<td>25</td>
<td>1 (1.1)</td>
<td>3.4 (3.3)</td>
</tr>
<tr>
<td>Day 1—dinner</td>
<td>15</td>
<td>90</td>
<td>1.9 (2.9)</td>
<td>4.1 (3.7)</td>
</tr>
<tr>
<td>Day 2—lunch</td>
<td>12</td>
<td>74</td>
<td>2.2 (3.3)</td>
<td>7.4 (7.3)</td>
</tr>
<tr>
<td>Day 2—dinner</td>
<td>12</td>
<td>90</td>
<td>1.2 (1)</td>
<td>4.6 (4.3)</td>
</tr>
<tr>
<td>Day 3—lunch</td>
<td>10</td>
<td>85</td>
<td>3.8 (5.1)</td>
<td>7.6 (6.3)</td>
</tr>
<tr>
<td>Day 3—dinner</td>
<td>15</td>
<td>109</td>
<td>1.9 (2)</td>
<td>5.5 (3.8)</td>
</tr>
<tr>
<td>Day 4—lunch</td>
<td>9</td>
<td>60</td>
<td>1.5 (2.5)</td>
<td>5.6 (7.5)</td>
</tr>
<tr>
<td>Day 4—dinner</td>
<td>10</td>
<td>90</td>
<td>2.1 (1.9)</td>
<td>6.5 (4.8)</td>
</tr>
<tr>
<td>Day 5—lunch</td>
<td>5</td>
<td>41</td>
<td>3.4 (5.2)</td>
<td>9.5 (6.7)</td>
</tr>
<tr>
<td><strong>RTF+MTF</strong></td>
<td>102</td>
<td>1039</td>
<td>3.8 (8.8)</td>
<td>19.9 (30.8)</td>
</tr>
</tbody>
</table>

\(^a\)There were no samples for day 5—dinner; food volume error is equivalent to mean error bias; error (intake) is equivalent to volume intake error; and 3D, % intake error is the same as in the study by Pfisterer et al [16].

\(^b\)RTF: regular texture food.

\(^c\)MTF: modified texture food.

Validating Nutrient Intake From Volume With Mass

In Figure 6, the MTF plates (blue) tended to be of lesser mass than the RTF plates (red), largely owing to the nature of RTF. RTFs represent available food choices from the LTC home, but they were prepared by a supermarket, which may not be consistent with LTC serving sizes. MTF were offered and prepared by the LTC home. This translates to a clustering effect of MTF foods at lower values of nutrients with RTF foods toward higher values of nutrients. We also observed a banding effect on fiber for the RTF data set owing to how mass was controlled for matching 25% portion increments and given the relatively few foods that contained fiber in the RTF data set. Regarding the spread of nutrient distributions, there is also much higher variance for the MTF data set for larger amounts of a nutrient (e.g., protein, fat, and iron), with tighter variances observed on smaller portion sizes.

On the basis of the coefficients of determination shown in Figure 6, nutrient estimates by volume were tightly linearly correlated with nutrient estimates from mass, with \(r^2\) values ranging from 0.92 for fat to 0.99 for vitamin C and vitamin K. This was true for all nutrients of interest (refer to Tables S2-S4 and Figure S1 in Multimedia Appendix 1) for a comprehensive assessment. On the basis of the Bland-Altman plots, not only were they tightly correlated but there was also good agreement between methods, as evidenced by small bias (\(|\mu| \leq 2.7\)) and 0 contained within the limits of agreement. Ideally, the bias distributions will be centered around the y-intercept (i.e., \(\mu=0\)). This was the case with \(\mu\) ranging from a minimum of -0.01 for vitamin B6 (mg), zinc (mg), and fat (g) to a maximum of -2.7 for calories (kcal). Taken together, these results suggest that nutrient estimation using the AFINI-T system appears to be valid. Estimates were well aligned with the gold-standard weighed food method, with the advantage of only single image acquisition and no need for weighing plates.
Figure 6. Correlation and agreement between mass and volume estimates for determining nutritional intake at the whole plate level across all imaged samples. Left panel depicts the goodness of fit with linear regression and coefficient of determination ($r^2$), and right panel depicts the degree of agreement between measures and bias from the Bland-Altman method. Correlation and agreement between mass and volume estimates of macronutrients are shown in the figure: (A) calories, (B) protein, and (C) fiber. In total, 3 nutrients of interest are shown here for brevity. RMSE: root mean square error.

A. Calorie estimation: AFINI-T vs mass

B. Protein estimation: AFINI-T vs mass

C. Fiber estimation: AFINI-T vs mass

Benchmarking the AFINI-T Approach With Current Practice and Requirements

Now, let us consider the feasibility of theoretical portability and completion task time by comparing the end-to-end AFINI-T system with the current workflow. A requirement identified in the study by Pfisterer et al [15] was for the system to run on a portable tablet with inconsistent Wi-Fi. By design, methodology and models were selected to support portability. For example, having selected an approach to support offline training in the EDFN and autoencoder, only the final model residing on the device, which does not require Wi-Fi. The autoencoder, which requires only a single training session for global feature extraction, encompasses 84,176 parameters. The per-meal classification layer requires an additional $15n_c$ trainable features, where $n_c$ is the number of food classes for meal $c$ (Table 3). The EDFN food detection network requires 13.7 million parameters, but does not require fine tuning and can be used globally across meals.

The second benchmark is regarding theoretical task completion time. In terms of benchmarking theoretical task completion time, we can compare with results from the study by Pfisterer et al [15]. When assuming a very conservative estimate including food handling of 10 seconds per image for acquisition, the time for preprocessing (eg, plate finding) takes approximately 2.5 seconds per image, with segmentation taking 0.7 seconds per image and classification taking 0.05 seconds per image (Dell XPS 15 9570; i7-8750H 2.20 GHz 6-core central processing unit; Nvidia GeForce GTX 1050 Ti). As shown in Table 5, even based on these conservative estimates, the theoretical completion time using AFINI-T meets the low end of task completion times (9 minutes 45 seconds vs a mode of 10-14 minutes of completion time for charting 1 meal). Here,
we have assumed separate imaging for each of the appetizers, mains, and desserts for each resident. Instead, if we consider acquisition as only acquiring the image (estimated time 1 second), this drops to 2 minutes 34 seconds. The true completion time will likely be between these upper and lower bounds, but the key point is that AFINI-T is platformed to take less time than the current methodology and with the added benefit of being objective and capturing data at a resident-centric level. Instead of a resident’s intake being binned into the 25% bin across the average foods served that day, AFINI-T captures details at the mL level and tracks personalized items ordered on a resident-by-resident basis.

Table 5. Summary of length of time required to complete food and fluid intake charting for 1 neighborhood (unit) comprising 16 residents, compared with theoretical AFINI-T processing.

<table>
<thead>
<tr>
<th>Type</th>
<th>Mode time (minutes)</th>
<th>Responses, n (%)</th>
<th>Time range (minutes)</th>
<th>AFINI-T estimate (1-second acquisition)</th>
<th>AFINI-T estimate (10-second acquisition)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food (per meal)</td>
<td>10 to 14</td>
<td>3 (33)</td>
<td>&lt;10 to &gt;25</td>
<td>2 minutes 34 seconds</td>
<td>9 minutes 45 seconds</td>
</tr>
<tr>
<td>Fluid (per meal)</td>
<td>10 to 14</td>
<td>4 (40)</td>
<td>&lt;10 to 25</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Snack (per snack)</td>
<td>&lt;10</td>
<td>5 (55)</td>
<td>&lt;10 to 19</td>
<td>52 seconds</td>
<td>3 minutes 15 seconds</td>
</tr>
</tbody>
</table>

aAFINI-T: automated food imaging and nutrient intake tracking.
bSample size, n=9.
cSample size, n=10.
dN/A: not applicable.

Discussion

Principal Findings

The AFINI-T method for estimating food intake is in strong agreement and tightly correlated with true intake. Especially in the case of larger intake portions, the AFINI-T method yielded accuracy of nutrient content with <5% error. For context, comparison with current visual assessment methods indicate errors in portion size 56% of the time for immediate identification and as low as 62% for delayed recording and stating that current methods’ error is too high for accurately identifying at-risk residents [25]. Interpretation of the acceptability of the precision and accuracy of the system requires further input from users, ideally through pragmatic trials. If warranted, improvements will require a degree of human input or expanded models. This may be in the form showing output classification masks so that misclassified segments can be reclassified as appropriate. Alternatively, it can be used to seed food item regions to tightly constrain food regions and then apply region growing to more accurately segment where there are food segments. This approach is consistent with what was integrated into the collaborative co-design prototype development outlined in the study by Pfisterer et al [15]. Although not fully automatic, collaborative segmentation through machine learning estimation that is checked and corrected, if necessary, by a human using a simple and intuitive interface will likely be an improvement on current food charting methods, particularly regarding accuracy and time. Timed comparison trials will be required to confirm this.

For the current AFINI-T approach, we show that segmentation of only 1 reference image is required and that even when some pixels are misclassified, there is reasonable robustness in nutrient intake accuracy. These misclassifications tended to occur near the edges of a food segment regardless of data set, which may be from a less uniform representation near the edges either because of higher crumbliness (eg, meat loaf crumbs) or owing to the convolutional kernel extending into the empty space (ie, the plate), making it easier to classify a pixel as food when there are food pixels surrounding it. These misclassification errors at the edges do not appear to translate to large intake errors. This fully automated classification strategy may be deemed feasible given the time savings. It is also consistent with the co-designed user interface and workflow we reported in our previous study, where users described acceptability for clicking on a large food region and defining its contents from a drop-down list [15] which can be prepopulated based on the menu items of the day.

In the case of frequent nutrient database missing values (eg, vitamin D [26]), there is reliance on complex imputations for estimates [27]. Additional discussions with end users and nutrition experts are warranted to evaluate the utility and appropriateness of reporting these values, the margin of error that is deemed acceptable for supporting trust in the system, and other considerations given the quality of data included in the underlying nutritional databases.

Comparison With Previous Studies

It is challenging to assess how AFINI-T compares with the literature because there are no food intake data sets on which benchmark tests can be conducted. Additional considerations affecting the ability to compare include the number of included classes, inconsistencies in accuracy reporting (eg, top-1 vs top-4 accuracy), and the complexity of the classification problem (eg, whole raw foods vs prepared meals modified texture versions of those prepared foods). Although direct comparison between the AFINI-T system and other automated methods for assessing LTC intake data is not possible because the AFINI-T system is the first to measure food intake and consider MTFs, these results suggest that AFINI-T’s deep neural network approach is among the highest performing approaches with a top-1 accuracy of 88.9%. Furthermore, the type of data represented in the MTF and RTF data sets for LTC contain more complex food scenarios, as they are prepared foods (RTF: 93.9% accuracy; MTF: 73.7% accuracy), and the accuracy we report is top 1,
which means that the AFINI-T approach may outperform the others.

Some accuracy for classification methods based on handcrafted features has been reported in the literature: 85% accuracy for 15 types of produce with minimum distance classifier [28], 88.2% accuracy for 18 classes of whole foods (entire pineapple) [29], 95% using top-4 accuracy for supermarket vegetable identification [30], 96.5% accuracy for 10 vegetables using a neural network with color and texture features [31], and 99% using top-2 accuracy for some fruits and vegetables by fusing 3 types of features (including Unser’s features) [32]. Regarding the trend for learned features, a deep learning approach has had comparatively slow adoption in the field of food imaging, with uptake occurring only recently [33-35].

Regarding accuracy reporting for segmentation and classification, these accuracies tend to not be mentioned [35-39] or are stated as beyond the scope of the present version of their system[37]. This is further confounded when segmentation and classification accuracies are combined instead of considering them as 2 subprocesses. Classification accuracies using deep learning vary from as high as 100% (11 classes) [34] to 82.5% (15 classes) [35]. Alternative methods used for classification were AdaBoost [37], K Nearest Neighbors [40], and support vector machines [36,39], with reported classification accuracies of 68.3% (50 classes) [36] to 99.1% (6 classes) [39]. At the inference level, few papers report percentage error at the nutrient level and tend to focus on calorie estimation or nonstandardized metrics: calorie estimation error of 0.09% (mean absolute error) on 6 categories using random forests and support vector machines [39] and 0.25% (mean SE) on 11 categories of entire foods (eg, green pepper) using a CNN [34]. Others have reported 80% of calorie estimates falling within 40% error (35% within 20% error) on 15 classes using a multitask CNN, with a maximum correlation coefficient of 0.81 ($r^2=0.64$ equivalent) and top-1 accuracy of 82.48% [35]. Previous study [35] also reports a comparison with the study by Miyazaki et al [37], with 79% of calorie estimates falling within 40% error (35% within 20% error) using handcrafted features, with a correlation coefficient of 0.32 ($r^2=0.10$ equivalent).

For comparison, the AFINI-T system demonstrated an error of 2.4% across 13 nutrients in 56 categories (102 classes) of food with minimum $r^2$ value of 0.92 (0.94 for calories). The average top-1 accuracy was 88.9%, ranging from 95.1% for 3 classes (RTF: dinner) to 70.4% and 89% for 15 class meals (MTF: day 3–dinner and MTF: day 1–lunch, respectively). On the basis of these comparisons, this study performs among the best reported results done in this area, these results represent a novel contribution both from the technical implementation and real-world implementation perspectives. Additional benefits of the AFINI-T system include its ability to measure a specific resident’s intake (as opposed to the proportion consumed across the average of all foods offered), with performance at least matching other approaches. Compared with the current visual assessment methods, it is easy to use, is fast to acquire and process, removes subjectivity, provides repeatable estimates, and can be tracked to the nutrient level to provide a comprehensive profile of each resident-specific intake in a quantitative way. This translates into high-quality data that can be used to inform resident preferences and streamline referrals to RDs, along with a data-driven approach for monitoring and evaluating nutritional interventions.

### limitations

First, ground truth volume was assumed to be equivalent to the RGB-D camera assessment. Although we collected ground truth weighed food records, as this study aimed to assess overall feasibility from an accuracy perspective through the lens of end users, we did not account for ground truth volume. Therefore, we were working under the assumption that AFINI-T’s volume assessment was accurate. Volume validation against gold-standard ground truth (eg, water displacement) is needed to corroborate the accuracy (although in actuality, there is some evidence suggesting there is <3% volume error of the RealSense [41]). This is an important consideration for more thoroughly quantifying error at each stage. Given the state of the literature on how error is typically reported (if it is at all), this paper provides evidence of the feasibility of more transparent technology for supporting trust in the system.

Second, although the plated foods are representative of LTC offerings, intake was physically simulated through incremental plating in the research kitchen by the researchers. Further studies need to be conducted to evaluate the imaging technology in real-world LTC resident food intake.

### Future Directions

Future directions include adding an additional stage for automatic food type classification as specific foods rather than arbitrary classes with associated nutritional values (ie, mashed potatoes are classified as mashed potatoes after the initial segmentation step). A human-in-the-loop version, where there is the opportunity to correct all misclassified regions (ie, the best-case scenario), can further improve results, albeit at the expense of manual hands-on time and effort, which needs to be minimized. In addition, improving the algorithms to handle more complex food types (eg, salads or soups in which the food comprises multiple components) and more complex plates of food to address food mixing as seen with mashed potatoes will improve AFINI-T’s ability to assess plates in the wild.

As observed in Figure 6, with small intake amounts and therefore small relative portion differences, the error was large. To improve on this, in the future, we must consider from where this error arose. A contributor may be the depth map variance. Results indicate that nutritional intake estimates had great variation at low levels of intake (large spread at low intake levels). This may correspond to the amount of variation in estimation at low levels of intake or large quantities of food left on the plate. We speculate that this is because of compounding of small discrepancies in depth maps, which get propagated to volume and then to nutritional intake. Future studies will address this issue by incorporating depth map variance as a feature to describe the food item. For example, for a green salad, we expect a higher variance in the depth map because it is a nondense food item. In contrast, meat loaf or slab cake will have very low depth
map variance across the food item, as these items are more similar to a block. Exploring automated 3D segmentation may also be intriguing, in which depth information can be stacked onto color channels and thus incorporated into salient feature extraction. Similar approaches have shown promise in recent advances in agriculture [42-44], construction [45], robotics, and automation [46].

From a translational standpoint, AFINI-T is platformed to provide actionable data-driven insights that can help to inform menu planning by dietitians and director of food services. For example, it can be used to develop recipes that are more nutrient-dense and complement the nutrients in recent past meals. Creating nutrient-dense meals while minimizing cost is a priority in LTC, as there is a fixed allocation of food cost per resident. The raw food allocation in Ontario was CAD $9.54 (US $6.82) per resident per day in 2020 [47]. Until recently, there was a disconnect between the perceived requirement to serve full portion to meet nutritional requirements (ie, the portion size that was costed to provide adequate nutrition); however, because of limited budget, the foods that were served were relatively inexpensive and the quantity required was unsuitable. This resulted in high degree of food waste [48,49], increasing the risk for malnutrition owing to less consumed nutrients than the planned nutrient consumption [50]. AFINI-T can also be used as a tool for developing more nutrient-dense recipes in which certain ingredients can be replaced with others. For example, replacing half of the ground beef in a chili recipe with lentils to decrease saturated fat and cholesterol and increase fiber. Data on which foods are consumed can inform how to design recipes to be smart, more expensive, and more nutrient-dense, with the expectation of less waste and more portion consumption, especially when paired with software such as Food Processor for designing recipes. Although these types of strategies were not part of this study, they are direct motivation for this project and have great potential to affect and disrupt the way we assess nutrition management and beyond when they are explored as part of future pragmatic trials.

Conclusions
AFINI-T is a feasible deep learning–powered computational nutrient sensing system that provides an automated, objective, and efficient alternative for food intake tracking, which provides food intake estimates. Novel contributions of this approach include a novel system with decoupled segmentation, classification, and nutrient estimation for monitoring error propagation and a convolutional autoencoder network for classifying regular texture and MTFs with top-1 accuracy of 88.9%, with mean intake error of 0.4 (SD 36.7) mL, and nutritional intake accuracy with strong agreement with gold-standard weighed food method and good agreement between methods ($r^2$ ranges from 0.92 to 0.99; $\sigma$ ranges from −2.7 to −0.01; 0 within the limits of agreement) across 13 nutrients of interest to LTC. Translation of AFINI-T may provide a novel means for more accurate and objective tracking of LTC resident food intake, thus providing new resident-specific insights for supporting well-being and preventing malnutrition. AFINI-T’s data-driven insights may streamline and prioritize dietitian referrals for supporting nutritional intervention efficacy. This may enhance the sensitivity of identifying at-risk residents and enable more holistic monitoring for malnutrition reduction.

Acknowledgments
This study was funded by the National Science and Engineering Research Council of Canada (PGSD3-489611-2016 and PDF-503038-2017), Canadian Institutes of Health Research (202012MFE-459277), Ontario Graduate Scholarship, and Canada Research Chairs program.

Data Availability
The data sets can be made available upon reasonable request.

Authors’ Contributions
KP and RA contributed equally to this study. KP conceptualized the system; RA and AW provided additional contributions to system design. KP was the main contributor to experimental design and contributed to algorithmic design. RA was the main contributor to algorithmic design, with additional contributions from KP and AW. AC provided additional support on data collection and preliminary analyses during the initial conceptualization of this project. KP was the main contributor to data analyses; KP and RA conducted data analyses. HK provided clinical nutrition direction and perspective. JB oversaw the user study elements. AW oversaw the project as a whole. KP was the main contributor to writing the manuscript, with additional contributions from RA. All authors reviewed the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
This multimedia appendix contains two parts: S1—standardizing nutrient values and S2—nutrient intake accuracies. S1 consists of Table S1, describing the workflow in converting percentage daily values to absolute measurements using Health Canada’s...
dietary reference intake values [20]. S2 contains supplementary tables on nutrient intake accuracies (Tables S2-S4), providing a comprehensive overview of nutrient intake accuracy within and across our long-term care datasets. S2 additionally contains Figures S1a to S1m, showing the correlation and agreement between mass and volume estimates for determining nutritional intake at the whole plate level across all imaged samples. The left panel depicts the goodness of fit with linear regression and coefficient of determination ($r^2$), and the right panel depicts the degree of agreement between measures and bias from the Bland-Altman method.

References


47. Ontario LCA. The role of long-term care. OLTCA. URL: https://www.oltca.com/oltca/OLTCA/Public/LongTermCare/FactsFigures.aspx [accessed 2022-03-08]


Abbreviations

AFINI-T: automated food imaging and nutrient intake tracking
CNN: convolutional neural network
EDFN: encoder-decoder food network
IOU: intersection over union
LTC: long-term care
MTF: modified texture food
RD: registered dietician
RGB-D: red-green-blue-depth
RTF: regular texture food
UNIMIB+: UNIMIB2016 with additional green representation

©Kaylen Pfisterer, Robert Amelard, Jennifer Boger, Heather Keller, Audrey Chung, Alexander Wong. Originally published in JMIR Aging (https://aging.jmir.org), 17.11.2022. 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.
Older Wheelchair Users Recommend Age-Friendly Design Improvements to a Wheelchair Maintenance App: Mixed Methods Development Study

Alyssa Boccardi¹, OTD; Fangzheng Wu¹, BS; Jon Pearlman¹, PhD; Anand Mhatre², PhD

¹Department of Rehabilitation Science and Technology, University of Pittsburgh, Pittsburgh, PA, United States
²Occupational Therapy Division, Ohio State University Wexner Medical Center, Columbus, OH, United States

Corresponding Author:
Anand Mhatre, PhD
Occupational Therapy Division
Ohio State University Wexner Medical Center
453 W 10th Ave
Columbus, OH, 43210
United States
Phone: 1 (614) 688 2081
Email: anand.mhatre@osumc.edu

Abstract

Background: Wheelchair part failures have doubled over the past decade. Preventative wheelchair maintenance reduces wheelchair failures and prevents user consequences. We are developing a smartphone app called WheelTrak, which alerts users when maintenance is required, to encourage maintenance practices and compliance.

Objective: This mixed methods study aimed to develop a wheelchair maintenance app using broad stakeholder advice and investigate older adults’ interaction experience with the app and their perceived barriers to and facilitators of maintenance.

Methods: Interviews were conducted with stakeholders, including mobility device users, to generate needs statements and app specifications. The app was designed in 2 stages. Stage 1 involved the development of the app according to the specifications and evaluation of the app interface by lead users. Stage 2 included the revision of the app screens and manual functionality testing. Usability testing and semistructured interviews were conducted with older wheelchair and scooter users. The System Usability Scale was used to measure app usability.

Results: Interviews with power and manual wheelchair users (37/57, 65%), wheelchair service providers (15/57, 26%), manufacturers (2/57, 4%), seating and mobility researchers (1/57, 2%), and insurance plan providers (2/57, 4%) informed the needs and specifications of the app technology. The 2-stage development process delivered a fully functional app that met the design specifications. In total, 12 older adults (mean age 74.2, SD 9.1 years; n=10, 83% women; and n=2, 17% men) participated in the usability testing study. Of the 12 participants, 9 (75%) agreed to use WheelTrak for preventative maintenance. WheelTrak scored an average System Usability Scale score of 60.25 (SD 16). Four overarching themes were identified: WheelTrak app improvements, barriers to maintenance, consequences related to mobility device failure, and smart technology use and acceptance. Older adults preferred the simplicity, readability, personalization, and availability of educational resources in the app. Barriers to maintenance pertained to health issues and lack of maintenance knowledge among older adults. Facilitators of maintenance included notification for maintenance, app connectivity with the service provider, reporting of device failure, and the presence of a caregiver for maintenance.

Conclusions: This study highlighted age-friendly design improvements to the app, making it easy to be used and adopted by older wheelchair users. The WheelTrak app has close to average system usability. Additional usability testing will be conducted following app revision in the future.

(JMIR Aging 2022;5(4):e39301) doi:10.2196/39301

KEYWORDS
aging; older adults; maintenance; mobile phone; repair; smartphone; wheelchair
**Introduction**

**The Calamity of Wheelchair Failures**

According to the World Health Organization, older adults will be using ≥2 assistive devices by 2030 to overcome barriers and experience full and equal participation in society [1]. Wheelchairs are assistive devices that serve as a primary means of mobility and independence for older adults and are linked to improved well-being and delayed need for long-term care. Unfortunately, although wheelchairs play a significant role in the lives of older adults and people with locomotor disabilities, they are known to break down frequently. Cross-sectional study findings over the past 2 decades show that approximately 45% to 88% of wheelchair users experience wheelchair part failures every 6 months owing to failures of wheelchair casters, rear wheels, brakes, frames, and seating systems [2-7]. Comparatively, in low- and middle-income countries, where rehabilitation services are scarce and outdoor environments are adverse, part failures occur every 2 to 3 months [8-13]. A high incidence of part failures among older wheelchair users was found in a study conducted in El Salvador [14,15]. Approximately 57% of the older adults have experienced part failures in the past 3 months of wheelchair use owing to high-risk failures of critical wheelchair parts such as casters and wheels. Approximately 75% to 95% of the older participants rated their wheelchairs as unsatisfactory and in unsafe working condition, which can contribute to part failure [14,15]. One-third of wheelchair breakdowns result in adverse events, including injuries, pain, depression, and hospitalization [2,4,6,12,13]. Overall, wheelchair part failures negatively affect the lives of older wheelchair users globally, thus increasing public health and personal burden.

**Need for Wheelchair Maintenance**

Community-based and secondary data analysis studies [7,16,17] and the World Health Organization’s guidelines [18] recommend preventative wheelchair maintenance to avoid failures that can lead to breakdowns, which make the wheelchair dysfunctional. A randomized controlled trial with 216 manual wheelchair users found that active checkups and maintenance in 12 months led to no wheelchair accidents in the treatment group [17]. The number of accidents in the control group remained the same. Despite this evidence, preventative maintenance is rarely conducted. This can be attributed to unfavorable health care policies; lack of user training, knowledge, and capability; and lack of tools for repair, among several other reasons. Researchers have developed resources such as training programs and maintenance checklists that include consensus-based, generic maintenance schedules for inspection and cleaning of wheelchair parts to support maintenance practices [17]. However, the checklists cannot monitor wheelchair use in the community and predict wheelchair failures to inform maintenance. In other industries, maintenance schedules are dependent on product use. For instance, in automobiles, the odometer indicates oil change based on the distance traveled by the vehicle [19]. In aircraft and heavy equipment industries, vibration-based condition monitoring systems are used to generate alerts for part replacement and preventative maintenance events and to prevent equipment damage and downtime [20,21]. Unfortunately, no such tools exist to monitor wheelchair use and wear down, determine the probability of high-risk failure, and alert users and wheelchair providers about maintenance and part replacement events.

**Availability of New Technology**

The widespread availability of low-cost activity monitoring tools, such as sensors and smartphones, offer an opportunity to track real-time wheelchair use characteristics and guide maintenance. Smartphones are widely available around the world. For instance, in low- and middle-income countries—approximately one-third of older individuals (aged >60 years) and people with disabilities have a smartphone, and that number continues to increase [22-24]. Therefore, we are developing a mobile health (mHealth) technology called WheelTrak to enable use-based maintenance practices for wheelchair users and stakeholders involved in wheelchair repair. The concept comprises a smartphone app connected via Bluetooth to a low-power sensor unit that attaches to the wheelchair and collects road shock data when the wheelchair is in motion. Shocks experienced over time will be benchmarked against a wheelchair wear index (WWI) that can predict the occurrence of critical wheelchair failure. When maintenance is required according to WWI, users can be notified through the WheelTrak smartphone app.

This paper describes the staged design process and usability testing of the WheelTrak smartphone app. The study aimed to evaluate the usability of the WheelTrak app for preventative maintenance and understand the barriers to and facilitators of maintenance, which can inform the future development of the technology.

**Methods**

A systematic design procedure proposed by Ulrich and Eppinger [25] was used to gather raw data and develop needs statements for the WheelTrak maintenance technology. The design and testing process followed in this study is shown in Figure 1.
Needs Identification

Multiple stakeholders in the wheelchair industry, who are associated with repairs and maintenance, including wheelchair users, were interviewed as a part of the technology transfer programs at the University of Pittsburgh. The interviews were conducted at a conference, in person at a wheelchair clinic, or via phone. Initial stakeholders were affiliated with the University of Pittsburgh’s Rehabilitation Science and Technology Continuing Education program. Then, stakeholders were recruited using the snowball sampling method or word of mouth at the conference or approached via social media platforms.

Customer discovery style interviews were conducted. Users were asked about their daily life activities and journey as a user of a mobility device. Likes and dislikes about the wheelchair and its use settings were collected. Instances where the user experienced inconveniences or consequences owing to wheelchair part failure or repair were discussed. Furthermore, users were probed regarding their understanding and awareness of preventative maintenance and related training. Opinions on wheelchair failures and repair services from wheelchair providers were collected. Users were asked about the features they would like to see in an mHealth preventative maintenance technology.

Wheelchair service providers were asked about their business operations and challenges, criteria for product selection and prescription, repair experiences of technicians, reimbursement versus costs, cascading effects of repairs on their operations, and experiences with clients and insurance. Providers were queried about the development and integration of new technology in their day-to-day repair-related operations. Wheelchair manufacturers and insurance plan providers were presented generic data on wheelchair part failures and asked to share their viewpoints on the existing state of affairs regarding repairs and maintenance. They were questioned about benefits of and risks with new maintenance-related technology and its integration into service. After each interview, the interviewee was asked for referrals for additional interviews. No identifiable information was collected during the interviews.

Raw data from interviewees were documented using handwritten notes, either by the author (AM) or 2 other researchers (refer to the Acknowledgments section). Notes included facts, insights, and quotes stated by the interviewees. Pain points for each stakeholder were extracted from the notes manually. These pain points were considered for generation of needs statements. For the providers, who articulated the needs well, direct quotes were converted into needs statements. Needs statements for the WheelTrak technology (app, sensor, and web-based platform) were generated by the author (AM) and a researcher (refer to the Acknowledgments section).

App Specification Development

On the basis of needs statements related to the maintenance app and benchmarking to existing health apps, functional requirements or specifications were generated for the WheelTrak app by authors (AM and FW) and another researcher (refer to the Acknowledgments section). These specifications focused on a low-fidelity version of the app or the minimum viable prototype for usability testing.

App Design—Stage 1

Using app specifications, screen wireframes were brainstormed by the author (AM) and other researchers (refer to the Acknowledgments section) and hand drawn subsequently. These screens were drafted in Adobe Illustrator (Adobe Inc) by a researcher and uploaded to a rapid prototyping platform. The platform established a sequence among the screens. Active, ultralight wheelchair users, who can be considered as lead users, tested the app prototype and provided suggestions.

App Design—Stage 2

Feedback from the users was implemented by revising the app screens in Adobe Illustrator. Then, the screens were migrated to Android Studio (Google). A low-fidelity WheelTrak app prototype was deployed on a smartphone for manually testing the appearance and functionality of app screens. Following several design iterations performed by the author (FW), the app screens and incorporated features addressed the app specifications.

Usability Testing

Ethics Approval

An institutional review board application (STUDY20100451) was reviewed and approved by the Human Research Protection Office committee at the University of Pittsburgh. Potential participants were contacted via phone, and a script approved by the institutional review board regarding study introduction
was followed to seek verbal consent from interested older wheelchair users.

**Participant Recruitment**

Inclusion criteria for participants were as follows: (1) aged ≥60 years and (2) had a manual wheelchair, power wheelchair, or scooter. Participants were recruited through the University of Pittsburgh Pepper Center Registry. Semistructured interviews were conducted with older wheelchair and scooter users to perform usability testing with the WheelTrak app.

**Usability Testing Procedure**

Before the interview was conducted, recruitment, screening, interview availability, and location were determined. We conducted the interviews using a set of questions (Multimedia Appendix 1) for older adults based on early wheelchair user interaction experiences related to WheelTrak development. During the interviews, participants were asked about their experiences with wheelchair failure, repair and maintenance, barriers to and facilitators of maintenance, and technology use. The app designed in stage 2 was downloaded in an Android phone. The interviewer demonstrated selective WheelTrak screens and allowed the participant to use the app afterward. Participants were asked questions regarding their impressions about the app. They were probed regarding whether the app would fit into their lifestyle and how often they would use it. Open-ended questions on the use of app features—wheelchair use data, failure reporting, maintenance notification, and awards—were asked. Finally, participants completed the System Usability Scale (SUS), which is a self-report, validated instrument widely used to assess users’ satisfaction with use, internal beliefs, motivation, attitudes, and intentions toward technology [26].

**Data Analysis**

Interviews were audio recorded and then transcribed verbatim. Transcripts were deidentified. A systematic approach to qualitative thematic analysis was used to analyze interview data and identify and develop codes and themes using NVivo (version 12 Plus; QSR International). First, all interviews were read by a researcher (AB) to gain an overview of the content. Open, axial, and selective coding strategies were used, enabling the researchers to interact with, compare, and reduce the data constantly. These strategies create a dynamic and nonlinear process that enables themes to be identified, coded, and interpreted [27]. Open coding is the first level of coding used to identify common concepts and themes expressed through interviews [28]. These ideas were given a descriptive label or code. The second level of coding used was axial coding, which was used to further refine, align, combine, and categorize existing themes that share similar ideas [27]. Finally, selective coding (third level of coding) was used to condense categories identified during axial coding, to discover overarching and main themes from participant interviews [27]. Two researchers (AB and AM) collaborated and created the coded interview results.

To evaluate the usability of WheelTrak quantitatively, the average of SUS total scores of all participants was computed. An SUS score of 80 out of 100 indicates that users are impressed with the app and would recommend it to others. An SUS score of approximately 68 is an average usability rating that indicates scope for improvement, and a score <51 indicates lack of usability and need for improvement [26].

**Results**

**Needs Identification Results**

Stakeholders listed in Table 1 were interviewed during the Randall Family Big Idea Competition (February 2019 to April 2019) and the School of Health and Rehabilitation Sciences Innovation Challenge (January 2020 to April 2020). Overall, stakeholders appreciated the development of preventative maintenance technology.
<table>
<thead>
<tr>
<th>Stakeholder</th>
<th>Participants, n (%)</th>
<th>Pain points</th>
<th>Notable quotes</th>
<th>Needs statements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Power and manual wheelchair users</td>
<td>37 (65)</td>
<td>• Failures degrade quality of life</td>
<td>• “The app is totally needed even for smaller repairs on my own chair.”</td>
<td>• WheelTrak sends maintenance alerts before device failure.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Repairs and expensive bills from providers</td>
<td>• “I could fix my own brakes instead of waiting for a provider.”</td>
<td>• WheelTrak is user-friendly.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Repair waiting times exceeding 6-7 months</td>
<td>• “You learn to become your own mechanic. It’s easier to do it (repairs) for yourself.”</td>
<td>• WheelTrak connects end user to the provider for reporting failures and scheduling repair.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Tired of waiting on phone to talk to someone and getting the runaround</td>
<td>• “Often the providers have ordered the wrong parts. I have to call providers to remind them to order parts.”</td>
<td>• WheelTrak allows user to order parts from vendors for small repairs.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Never received any wheelchair training</td>
<td>• “If your chair breaks, you are out of the chair, it is frustrating.”</td>
<td>• WheelTrak displays wheelchair use data.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Small failures lead to big consequences</td>
<td>• “Maintenance user manual is a piece of paper, nobody sees it.”</td>
<td>• WheelTrak is available through a subscription service.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Need to keep track of wheelchair use</td>
<td>• “Insurance won’t cover loose brakes which can cause falls or tips.”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Users have to take wheelchair to the provider many times</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Time off work is common</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Maintenance is not performed despite training</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wheelchair providers</td>
<td>15 (26)</td>
<td>• Financial losses from wheelchair repairs</td>
<td>• “Repairs are the bane of our existence. This technology can make repairs easy,”</td>
<td>• WheelTrak performs repair coordination and scheduling.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• “Insurance does not look at long term savings;”</td>
<td>• WheelTrak clubs repairs in a distant area and reduces technician trips.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• “…Seems like they (insurance) purposely drag their feet and stall these things.”</td>
<td></td>
</tr>
<tr>
<td>Wheelchair manufacturer</td>
<td>2 (4)</td>
<td>• Loss of reputation</td>
<td>• “We will find how our products are performing.”</td>
<td>• WheelTrak acts as a wheelchair add-on technology.</td>
</tr>
<tr>
<td>Seating and mobility researcher</td>
<td>1 (2)</td>
<td>• No immediate consequence</td>
<td>• “This will bring peace to the wheelchair sector; suppliers won’t have to go through the hassle of convincing the insurance through paperwork.”</td>
<td>• WheelTrak can report health and mobility outcomes.</td>
</tr>
<tr>
<td>Insurance plan provider</td>
<td>2 (4)</td>
<td>• No consequence, but concerned about the risk to the patient population</td>
<td>• “…We can more accurately measure causes for a failure, which is fantastic, and would help with claims justification.”</td>
<td>• None reported.</td>
</tr>
</tbody>
</table>

**App Specification Development**

The app specifications were generated explicitly in this study for targeted use by wheelchair or mobility device users (Textbox 1).
Textbox 1. App specifications for targeted use by wheelchair or mobility device users.

- User account setup
  - The user should be able to register and log in to the app using an email or their Google or Facebook accounts.
  - The user should be able to retrieve and edit their information.
- Communicate wheelchair use data
  - The app displays daily and weekly travel distance, speed, and impacts.
- Wheelchair information
  - The app collects and displays data on wheelchair purchase, model, and manufacturer.
- Wheelchair failure reporting
  - The app records wheelchair failure for communicating repair to the wheelchair provider.
- Push notification
  - The app sends a notification when maintenance is required.
- Health data collection
  - The app collects mobility outcomes data every 3 months using the validated Functional Mobility Assessment tool [29]. This feature was requested by the sponsors of the study. It was hypothesized that wheelchair condition is related to functional mobility outcomes and health.
- Reward user
  - The user scores an award for attaining maintenance milestones.
- Platform and device
  - The app can be used on Android phone and tablet.
- Data storage platform
  - The app uses Firebase (Google).
- Connectivity
  - The app uses Bluetooth Low Energy 4.0.

Stage 1 App Design Results
Dummy prototype screens based on specifications and brainstormed wireframes were deployed on Marvel app display engine (Figure 2). In total, 5% (2/37) of the adult wheelchair users navigated through the sequence of screens and were excited at the prospect of having a wheelchair maintenance app for reminders. Both users liked the color contrast, layout, and app features. A user requested for increment in text size and optimization of the display on use factors such as distance, speed, and impacts. Another user requested the inclusion of a parts store, which was beyond the scope of this study.
Figure 2. WheelTrak app account setup (A), wheelchair information (B), main screen (C), and maintenance notification (D).

Stage 2 App Design Results

Apple Health and Fitbit were installed in the phone, and essential health app features, including data display and user account management, were reviewed [30, 31]. The feedback from stage 1 and other app reviews informed the development of next version of the screens. The revised screens (Figure 3) were built and deployed on an Android smartphone for iterative specification testing. The app screens and features met the design specifications. The WheelTrak app’s main screen displays wheelchair use data. The records tab at the bottom left of the main screen shows daily and weekly use. Wheelchair information collection, failure reporting, health data collection, and scoring of awards are displayed on separate app screens and can be navigated through icons at the bottom of the main screen. User account setup and sensor connection is performed when the user logs in for the first time. Data retrieval from and deposition on the cloud occurs in the background. No glitches or crashes were encountered in the app version used for usability testing.

Figure 3. WheelTrak login screen (A), main screen (B), maintenance notification (C), and failure report screen (D).

Usability Testing Results

In total, 12 older adult wheelchair and scooter users participated in usability testing. Overall, 25% (3/12) of the interviews were conducted via phone and Zoom Meetings (Zoom Video Communications), a web-based meeting platform [32]. The remaining interviews (9/12, 75%) were conducted at the participants’ residence or at the University of Pittsburgh’s Department of Rehabilitation Science and Technology Design Studio. Table 2 presents the demographic characteristics of the participants. Interview themes were classified as shown in Table 3.
Table 2. Demographic and wheelchair use characteristics of participants (n=12).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>74.2 (9.1)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Men</td>
<td>2 (17)</td>
</tr>
<tr>
<td><strong>Type of residence, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>House</td>
<td>5 (42)</td>
</tr>
<tr>
<td>Apartment</td>
<td>7 (58)</td>
</tr>
<tr>
<td><strong>Living arrangement, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Live alone without assistance</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Live alone with assistance</td>
<td>6 (50)</td>
</tr>
<tr>
<td>Live with a family member for assistance</td>
<td>4 (33)</td>
</tr>
<tr>
<td><strong>Mobility devices, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Manual wheelchair</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Electric wheelchair</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Manual and electric wheelchairs</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Manual wheelchair and walker</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Scooter</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Stated liking the mobility device, n (%)</td>
<td>12 (100)</td>
</tr>
<tr>
<td><strong>Disliked aspects about wheelchair or scooter, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Difficulty in avoiding bumping into objects when driving backward</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Going over thresholds</td>
<td>2 (17)</td>
</tr>
<tr>
<td>Decreased battery power</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Wheelchair is uncomfortable owing to spasticity in lower extremities</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Lack of portability</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Indoor mobility device activities, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>All indoor activities</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Some indoor activities (eg, mobility around apartment, house, or job; eating; transfers; mopping floors; and assisting with carrying things from kitchen to dining room)</td>
<td>8 (67)</td>
</tr>
<tr>
<td><strong>Outdoor mobility device activities, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>All outdoor activities, using vehicle for transportation</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Activities without vehicle, using own wheelchair</td>
<td>8 (67)</td>
</tr>
<tr>
<td>Activities just outside the house (eg, sit on the porch)</td>
<td>1 (8)</td>
</tr>
</tbody>
</table>
Table 3. Interview themes (n=12).

<table>
<thead>
<tr>
<th>Themes and categories</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WheelTrak app improvement</strong></td>
<td></td>
</tr>
<tr>
<td>Improve readability</td>
<td></td>
</tr>
<tr>
<td>Increase text size</td>
<td>6 (50)</td>
</tr>
<tr>
<td>Increase icon size</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Change pie graph colors</td>
<td>8 (67)</td>
</tr>
<tr>
<td>Prioritize simplicity</td>
<td></td>
</tr>
<tr>
<td>Simplify</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Make graphs and charts easy to understand</td>
<td>8 (67)</td>
</tr>
<tr>
<td>The awards screen of the app is unnecessary for older adults</td>
<td>11 (92)</td>
</tr>
<tr>
<td>Include personalization</td>
<td></td>
</tr>
<tr>
<td>Use a phone to contact the provider for maintenance, instead of connecting via WheelTrak</td>
<td>4 (33)</td>
</tr>
<tr>
<td>Send the maintenance notification to the provider through WheelTrak</td>
<td>8 (67)</td>
</tr>
<tr>
<td>Maintain privacy regarding maintenance events</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Provision of manual, guide, or video for maintenance</td>
<td>6 (50)</td>
</tr>
<tr>
<td>Barriers to maintenance</td>
<td></td>
</tr>
<tr>
<td>Lack of maintenance training</td>
<td>11 (92)</td>
</tr>
<tr>
<td>Low confidence in conducting maintenance</td>
<td>5 (42)</td>
</tr>
<tr>
<td>Lack of ability to conduct maintenance owing to health issues</td>
<td>8 (67)</td>
</tr>
<tr>
<td>Consequences related to mobility device failure</td>
<td></td>
</tr>
<tr>
<td>Reported failures and repairs</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Reported consequences after failures</td>
<td>6 (50)</td>
</tr>
<tr>
<td>Repairs by providers were not timely</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Smart technology use and acceptance</td>
<td></td>
</tr>
<tr>
<td>Use of smartphone or tablet</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Stated that they will use WheelTrak app</td>
<td></td>
</tr>
<tr>
<td>Only when they remember</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Once a day</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Once a week</td>
<td>1 (8)</td>
</tr>
<tr>
<td>If they had a bumpy ride or travel</td>
<td>1 (8)</td>
</tr>
<tr>
<td>When they received a maintenance notification</td>
<td>1 (8)</td>
</tr>
<tr>
<td>Everyday</td>
<td>4 (33)</td>
</tr>
</tbody>
</table>

Comments and Additional Findings on WheelTrak App Improvement

Comprehension of the information displayed on the app screens varied across participants and depended on their health conditions and technology use experiences:

*I’ve had a cataract operation, so it’s hard to read.*
[Participant 5]

*The text size is too condensed it would be more beneficial to have a font that everyone can read.*
[Participant 4]

In addition, 67% (8/12) of the participants reported difficulty in reading the chart, especially the pie graph owing to the similarity in colors. The size of icons at the base of the main WheelTrak screen created complications in navigating through various app screens. Participant 3 does not have a smartphone, so it was difficult for them to touch the icons and input information. Participant 10 and participant 6 experienced seizures and difficulty in touching the icons.

Simplification of the app was noted as a priority by 33% (4/12) of the participants:
The 33% (4/12) of the participants indicated improvements to the information and graphical elements on the main screen.

WheelTrak app screens for failure reporting, wheelchair information collection, and notifications were considered to be the most important by the participants. Convenience was a common feature among these pages owing to access to provider information. One-third of the participants (4/12, 33%) valued the connection with a wheelchair provider for maintenance and repair purposes. Half of the study participants (6/12, 50%) liked the functionality of reporting wheelchair failures using the app. By contrast, most participants (11/12, 92%) reported the awards page to be unnecessary and noted that they would not be motivated by it. Overall, 50% (6/12) of the participants indicated that it would be helpful to have a manual, guide, or training video for learning wheelchair parts, app functions, and maintenance tasks.

When asked how participants would like to be notified about maintenance, phone call, SMS text message, or email were the preferred choices. Overall, 75% (9/12) of the participants preferred privacy regarding maintenance notification. They feared that if their children or spouses were notified of maintenance, it will increase their care burden.

**Comments on Barriers to Maintenance**

In total, 92% (11/12) of the participants in the study were not trained in mobility device maintenance. A collective lack of maintenance knowledge and confidence was observed:

> You don’t want to mess anything up and just like me, I don’t know how to really do the parts, so I would have to learn about it first. [Participant 6]

Additional barriers noted were health issues, including vision problems and lack of strength and dexterity in the hands, making manipulating objects difficult. Of the 12 participants, 3 (25%) stated that they would ask someone to do maintenance. Approximately one-third of the study participants (4/12, 33%) conducted maintenance activities including fixing armrests, inflating tires, and replacing scooter spark plug. Participant 11 received training from a wheelchair maintenance program and conducted maintenance with assistance from their spouse.

**Comments on Consequences Related to Mobility Device Failure**

Overall, 83% (10/12) of the participants reported wheelchair repairs and failures before the study. Brakes, wheels, and tires incurred the most failures. These parts were replaced during repair. Other failures were found with armrests, battery connections, grip handles, back support, front struts, and cushions. In total, 50% (6/12) of the participants reported consequences associated with these failures. Participant 3 and participant 10 stated that they are aware that their brakes are not secure and may slip, causing them to be more cautious and fearful when standing up. The scooter of participant 2 broke down on a transportation bus and they could not perform their daily life activities afterward. Overall, 17% (2/12) of the participants reported receiving a loaner chair that they did not like, thus causing frustration. Most repairs were completed by wheelchair service providers or the Veteran Affairs. They were timely, except for 25% (3/12) of the participants. Participant 6 stated that their repair took a week, whereas participant 7 stated that it took 2 months. Participant 11 experienced a front caster fracture failure during regular use and stated that their wheelchair has not been repaired yet. Overall, >4 months have elapsed, and the participant still uses a loaner chair.

**Comments on Smart Technology Use and Acceptance**

More than half of the older wheelchair user participants (9/12, 75%) used smart technology, that is, smartphones and tablets. Participants expressed proficiency in using their flip phones, home phones, or smartphones. All participants (12/12, 100%) indicated that they keep their phones on them or close to them. These places include their bag, wheelchair pocket, pants pocket, rollator basket, and kitchen or bedroom. Overall, 17% (2/12) of the participants indicated that they kept their phones on them to report falls. Regarding usability, the most significant barrier noted was remembering to charge the devices. Almost all participants (11/12, 92%) noted that they charge their phone daily or when it alerts them about low battery. The participants disliked short charging cords or plugging the charging cable into the wall socket.

**Additional Feedback on WheelTrak Technology**

All study participants (12/12, 100%) recognized the importance of routine maintenance for wheelchairs during the interviews. In total, 75% (9/12) of the participants appreciated WheelTrak’s development and wanted to use the technology (app and sensor). Of the 12 participants, 3 (25%) participants (participant 6, participant 8, and participant 11) did not favor using the maintenance technology. Participant 6 reported that her wheelchair has electronics that track daily distance and speed:

> I know my chair pretty well when something is going to break, so I don’t need the app. [Participant 6]

Participant 8 and participant 11 reported that they would not use the app because they are inactive. Participant 6 and participant 8 experienced wheelchair failures and subsequent consequences before the study.

**Quantifying WheelTrak Usability**

The average SUS score was 60.25 (SD 16). Figure 4 shows the average score for each SUS item. As far as individual scores are concerned, participant 6 had the highest SUS score of 90 but stated that they would not use WheelTrak. They expressed that the app was designed well and easy to use. Participant 1 had the lowest SUS score of 35; they were unfamiliar with using a smartphone and experienced extreme difficulty in using the WheelTrak app.
Discussion

Principal Findings

Our study demonstrates that older adults using wheelchairs favor the use of the WheelTrak smartphone app for conducting preventative maintenance and suggests age-friendly modifications to improve the usability of the app.

Preventative maintenance reduces the frequency of wheelchair failures and breakdowns, thus preventing adverse health consequences to older users [7,16,17]. To encourage maintenance practices and compliance, we are leveraging the existing capabilities of smartphones and sensors and developing an mHealth app called WheelTrak. A systematic design process used to gather technology design requirements revealed multiple development areas for maintenance technology, including the app, from the perspective of diverse stakeholders. This study specifically aimed to evaluate the usability of and barriers to the WheelTrak smartphone app with the older subset of the target user population. All the older wheelchair and scooter user participants in the study (12/12, 100%) recognized the importance of conducting maintenance, and 75% (9/12) of them expressed interest in using WheelTrak for maintenance purposes. This demonstrates that although some older adults did not use smartphones or tablets, they still perceived WheelTrak positively. In addition, participants rated the WheelTrak app for close to average usability, according to the SUS score. They recommended design improvements to make WheelTrak more inclusive for use and adoption by the older population.

Facilitators for app-based preventative maintenance included smart technology use, relevant app features, and interest in conducting maintenance with support from family member. Overall, 75% (9/12) of the participants used smart technology, which is slightly more than the statistics reported by the Pew Research Center [33]. Most smart technology users (9/12, 75%) were interested in using the WheelTrak app for maintenance. Specifically, the features such as maintenance reminders, failure reporting using photos of failure, and ability to connect with the provider were much appreciated. This outcome indicates that the WheelTrak technology provides convenience and relief from pain points on failures and scheduling repairs, as listed in Table 1. These pain points have plagued the wheelchair user community for decades. Users were motivated to leverage the existing resources and family member or caregiver support for conducting maintenance. Aligning with the notable quotes listed in Table 1, users are interested in conducting small repairs, which can avert the occurrence of major failure-related consequences.

As measured by the SUS instrument, the usability of the WheelTrak app exhibits significant variability across the SUS items, as seen in Figure 4. Key reasons for such variability can be attributed to familiarity with and challenges in using digital technology, confidence in device maintenance, usability issues with the current WheelTrak app version, and existing bias toward mobility device capabilities, as found during the study. These reasons may be characteristic of the older population of wheelchair users. These app usability findings prompt the development of a new app version for older adults, based on feedback collected in this study.

The WheelTrak app has close to average usability rating (SUS score=60.5), as the app features are yet to be tailored to the capabilities of older adults. When app screens were...
demonstrated to older users, comprehension of wheelchair use information through animated line graphs was difficult. Challenges were encountered with navigation through different screens. Although all the participants (12/12, 100%) conceded the importance of maintenance, most of them (8/12, 67%) experienced health limitations to perform general maintenance. Such limitations are commonly noted as barriers in mHealth literature about older adults. Limited physical ability, cognition, perception, and motivation have been cited as barriers to using mHealth technologies [34]. In addition, as highlighted in a previous app review study [35], the variations in capabilities of our study cohort showed that one size does not fit all; hence, we need to tailor the app to different older adult personas. We plan to address the barriers found during usability testing and accommodate the recommendations suggested by the older adults.

Older adults recommended simplicity in visualization and information communication. Accordingly, we plan to develop new specifications for text and icon size. The amount of information on each page shall be reduced and displayed in a text format for easy comprehension. We may use paper prototyping and testing to develop such specifications and test new app workflows. Use graphs, color fading, and animations can be removed completely. Similarly, the awards and records screens are not appealing to the older adults and can be removed. The valuable feedback collected in this study on app design improvements—optimal typography, color contrast, icon size, and information personalization—can apply to other apps developed for the older population of wheelchair users.

Personalization of app features is highly valued by older adults. The users want to exercise control over app communication. Privacy of use and maintenance data and communication of maintenance events via phone or app will be based on user selection. These options could be selected during the installation of the app, with support from a technical person who may be a family member, caregiver, assistive technology professional, or wheelchair provider technician.

All except 1 participant (11/12, 92%) were unaware of maintenance training and lacked technical knowledge about wheelchairs. This finding was synonymous with the user feedback collected during needs assessment. Although barriers to training availability exist, half of the study participants (6/12, 50%) expressed willingness to educate themselves about maintenance through videos, manuals, and resources in the app. The WheelTrak app can incorporate such training materials to increase maintenance knowledge, familiarize users with WheelTrak app–enabled maintenance, and assist users in taking charge of maintenance. Furthermore, we anticipate the integration of WheelTrak technology in the in-person and remote wheelchair maintenance training programs to augment maintenance training capabilities and outcomes [36]. With increasing smartphone adoption among older adults and people with disabilities, a low-cost tool such as WheelTrak can assist in scaling maintenance training and enable great compliance.

Most participants (8/12, 67%) expected WheelTrak to alert providers about upcoming maintenance events. This motivates us to explore the development of WheelTrak for the wheelchair provider group. Providers incur wheelchair repair losses, a finding from provider feedback during needs assessment and a cross-sectional study that surveyed >125 providers [37]. According to the provider’s feedback obtained during the app design phase, these losses can be prevented if WheelTrak can streamline and schedule upcoming repairs. Accordingly, web-based platform development can be investigated in the future for providers.

Participants’ experiences with wheelchair failures and repairs are similar to those reported in the literature [3-5] and needs assessment phase. These events result in long time without the wheelchair and severely limited mobility, as repair wait times stretch beyond months, which means spending more time in bed [3]. It is not surprising that wheelchair failures are associated with pressure injuries and rehospitalization [6]. Consequences include time off work, numerous calls and trips to the provider, and expensive repair bills if insurance does not cover repair. In addition, older users’ frustrations with loaner wheelchairs were documented in this study.

A trained participant (participant 11) who conducts maintenance experienced a high-risk failure during regular wheelchair use. These findings support the development of WheelTrak technology for monitoring wheelchair condition and informing users about upcoming high-risk failures and maintenance events. WheelTrak plans to predict failures, especially those related to wheels and brakes. These parts need to be replaced to avoid risks related to tips and falling out of the wheelchair, which can cause injury to the user.

The barriers and consequences realized in this study and recent studies on repair experiences of wheelchair providers demonstrate that it is crucial to address the repair-related needs of mobility device users and providers. Furthermore, as WheelTrak develops, health care policies must enact provisions and support preventative maintenance practices. Reimbursements for service and maintenance activities can be allowed with justification provided by ground truth data collected by WheelTrak.

**Limitations**

First, the design phase cast a wide net for gathering WheelTrak app requirements, but the usability test was conducted only with older adults, a subset of the intended WheelTrak user population. This may have affected the SUS usability score. Second, the usability testing study concentrated on app navigation and interaction and was limited to specific app screens that older participants would use. For instance, user account setup, sensor connectivity via Bluetooth, cloud connectivity, and health information collection using the Functional Mobility Assessment tool are features available in the app. These features will be tested in future studies with other user cohorts. Finally, we aimed to limit potential bias in soliciting positive feedback for WheelTrak. For this purpose, AB was initially recruited to conduct interviews.

**Future Studies**

We plan to develop an age-friendly version of the app based on study results and conduct focus group testing. In addition to SUS, we plan to include other validated tools to understand
users’ cognitive load while performing tasks in the app. As part of ongoing studies, the app’s current version will be tested with other mobility-assistive device users (aged 60 years) to identify their perceptions and usability of WheelTrak. WheelTrak development will include other modes of informing participants about maintenance such as SMS text messages, email, and phone calls, in addition to app notifications. Along with app development, we are conducting field studies to monitor wheelchair use and develop a WWI-based preventative maintenance model.

Conclusions

The WheelTrak preventative maintenance app has been identified as a tool that older adults can use for maintenance notifications and reporting wheelchair failures to providers. Despite challenges in using smart technology, older adults expressed interest in educating themselves about maintenance and conducting WheelTrak-led maintenance with caregiver or provider support. The WheelTrak app has close to average usability for older adults with disabilities. Findings from the study informed the research team about improvements to the app, making it easy to be used and adopted by wheelchair users across their life span.

Acknowledgments

This study was supported by the National Institute on Aging (R03 #AG069836) through the National Academy of Medicine Healthy Longevity Challenge and the Year of Pitt Global funding from the University of Pittsburgh. The authors would like to sincerely thank the stakeholders and study participants for their time and contribution to this study. In addition, the authors thank Aaron Anderson for his contribution to initial needs, specification development, and app screen design in stage 1 and earlier. The authors wish to thank Jake Meadows and Adam Cunningham for taking handwritten notes, brainstorming stakeholder interview questions, and developing the app. The authors thank Brian Robinson, Vineeth Poojary, and Surya Sankaralingam for participating in the technology transfer programs and for their assistance with the project. The authors are grateful to the Pepper Center Registry administrators at the University of Pittsburgh for providing access to the registry and for recruitment.

Data Availability

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

Authors’ Contributions

AM led the study and manuscript writing while at the University of Pittsburgh’s Department of Rehabilitation Science and Technology and transitioned to The Ohio State University’s Division of Occupational Therapy during the review and editing process.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 21 KB - aging_v5j4e39301_app1.docx]

References


Abbreviations

- **mHealth**: mobile health
- **SUS**: System Usability Scale
- **WWI**: wheelchair wear index

© Alyssa Boccardi, Fangzheng Wu, Jon Pearlman, Anand Mhatre. Originally published in JMIR Aging (https://aging.jmir.org), 18.10.2022. 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.
The Factors Influencing Older Adults’ Decisions Surrounding Adoption of Technology: Quantitative Experimental Study

Jerad Moxley¹, PhD; Joseph Sharit², PhD; Sara J Czaja¹, PhD

¹Division of Geriatrics and Palliative Medicine, Center on Aging and Behavioral Research, Weill Cornell Medicine, New York, NY, United States
²Department of Industrial and Systems Engineering, University of Miami, Miami, FL, United States

Corresponding Author:
Sara J Czaja, PhD
Division of Geriatrics and Palliative Medicine
Center on Aging and Behavioral Research
Weill Cornell Medicine
420 70th St
New York, NY, 10021
United States
Phone: 1 6469627148
Email: sjc7004@med.cornell.edu

Abstract

Background: The rapid diffusion of technology apps may support older adults’ independence and improve the quality of their lives. Models for predicting technology acceptance in older adults are sparse, based on broad questions related to general technology acceptance, and largely not grounded in theories of aging.

Objective: This study aimed to use a mixed methods approach involving 5 technologies to comprehensively assess the causal relationships among factors that influence older adults’ willingness to adopt the technologies.

Methods: In total, 187 men and women aged 65 to 92 years participated in the study. Participants were given presentations on 5 different technologies spanning domains that included transportation, leisure, health, and new learning and provided ratings of each technology on various measures hypothesized to influence adoption. They were also administered other instruments to collect data on their actual and self-assessed cognitive abilities, rates of discounting of the technologies with respect to willingness to invest time to attain higher skills in the technologies, general technology experience, and attitudes toward technology. We used the machine learning technique of k-fold cross-validated regressions to select variables that predicted participants’ willingness to adopt the technologies.

Results: Willingness to adopt technologies was most impacted by 3 variables: perceived value of the technologies ($\beta=.54$), perceived improvement in quality of life attainable from the technologies ($\beta=.24$), and confidence in being able to use the technologies ($\beta=.15$). These variables, in turn, were mostly facilitated or inhibited by the perceived effort required to learn to use the technologies, a positive attitude toward technology as reflected in the optimism component of the technology readiness scale, the degree to which technologies were discounted, and the perceived help needed to learn to use the technologies.

Conclusions: Our findings demonstrate that participants’ willingness to adopt technologies is mainly determined by perceptions of 3 aspects of the technologies; these aspects possibly mediate many relationships with willingness to adopt. We discuss the implications of these findings for the design and marketing of technology products for older consumers.

(JMIR Aging 2022;5(4):e39890) doi:10.2196/39890

KEYWORDS

aging; technology; design; older adult; technology; application; independence; relationship; adopt; transportation; leisure; health; learning; adoption; cognition; cognitive; willingness; design; marketing; consumer; mobile phone
Introduction

Background

New technologies are diffusing into everyday life at an extraordinary pace. These technologies span domains that include health and wellness, communication and socialization, transportation, entertainment, lifelong learning, and home support and may be found in different forms such as websites, mobile apps, wearable devices, and consoles. Importantly, many existing and emerging technologies may afford older adults with opportunities for enhanced independence, quality of life [1], and more successful aging [2-5] by promoting the maintenance of mental and physical health and life-engagement activities and, more generally, the continuation of adaptation to age-related changes over the life span [6]. However, older adults consistently adopt technology at lower rates compared with younger age groups [7-9], which compromises their ability to derive benefits offered by technology.

General Models of Technology Acceptance

Given the broad and important implications of technology use, even beyond older adults, several models have been developed and refined with the purpose of delineating factors that predict technology adoption in the general population. A widely cited early model is the Technology Acceptance Model (TAM) [10,11]. It posits that the use of a technology system is predicated by an individual's motivation to use it, which depends on 3 variables: the perceived usefulness of the technology, the perceived ease of use of the technology, and overall attitude toward using the technology.

Various modifications of the TAM [12] led to TAM 2 [13], which largely focused on technology use in the workplace. It proposes additional variables that influence the perceived usefulness of a technology, such as job relevance and output quality. Another widely cited model for technology acceptance is the Unified Theory of Acceptance and Use of Technology (UTAUT) [14], which coalesces data from 8 prior technology adoption models that consider the roles of constructs such as social factors, job fit, subjective norms, perceived ease of use and usefulness, self-efficacy, and attitudes toward behavior. The UTAUT allows for age to interact with every relationship in the model.

The UTAUT model also posited 4 additional factors important in the behavioral intention to adopt technology: facilitating conditions, social influence, effort expectancy, and performance expectancy. However, the data that the UTAUT and its precursor models were based on were largely derived from students and the concern for people in work situations who were reticent to adopt workplace technologies that could potentially benefit themselves and their employers. As stated by these authors, “UTAUT thus provides a useful tool for managers needing to assess the likelihood of success for new technology introductions and helps them understand the drivers of acceptance to proactively design interventions (including training, marketing, etc) targeted at populations of users that may be less inclined to adopt and use new systems” [14]. In addition, these models were not explicitly tested with older age groups and thus did not specifically address factors associated with older populations. To this end, the UTAUT model has been extended with new constructs such as privacy and to new populations [15].

Models of Technology Acceptance Specific to Older Adults

The Senior TAM (STAM) [16,17] also attempts to address these issues. On the basis of the TAM and UTAUT modeling frameworks, the STAM was developed to predict the acceptance of general technology by Hong Kong Chinese older adults through the inclusion of several factors. Although both the TAM and UTAUT propose a causality flow whereby a set of constructs causes another set of constructs, which in turn causes the use of technology, the STAM differs from the TAM and UTAUT in that the causal specifications are much broader.

For example, in the TAM, the direct effect of perceived ease of use on the behavioral intention to use technology is predicted to be influenced by attitudes about the technology. In contrast, the STAM model found no support for a direct effect of perceived ease of use (or of perceived usefulness) on the actual use of technology. Although both perceived ease of use and perceived usefulness predicted attitudes toward the use of technology, attitudes did not significantly predict actual use. Instead, the STAM found a broad array of variables, including gerontechnology self-efficacy, gerontechnology anxiety, facilitating conditions (knowledge, guidance, and support from other people), health and ability characteristics, social relationships, and attitudes toward life and satisfaction, to have a direct effect on the actual use of technology. In addition, almost every one of these predictors had a direct effect on perceived ease of use, perceived usefulness, or both variables. However, the general nature of how causality is specified in this model implies that it is difficult to refute [18].

It should also be noted that the development of the STAM [16] was based on the outcome measure of how many technologies respondents had used in the previous 12 months. Therefore, in contrast to other models of technology adoption, the outcome was retrospective and not concurrent or prospective. This raises the possibility that the conclusions that formed the STAM model may have been due to differences in older adults' retrospective technology use rather than based on their concurrent or prospective attitudes toward adopting technology.

More recently, Harris and Rogers [19] developed a health care TAM based on older adults with hypertension. In total, 23 older adults were interviewed, and the interview transcripts were analyzed to identify factors that were frequently mentioned for the consideration of use of each of 3 health care technologies: a blood pressure monitor, an electronic pillbox, and a multifunction robot. Perceived ease of use, perceived usefulness, facilitating conditions, and social influences—4 predictors commonly used in the theoretical models of technology acceptance associated with the TAM—were assumed to be the primary drivers of behavioral intentions to adopt technology. The qualitative analysis revealed that a host of other factors could impact these predictors, including perceived need, privacy and trust in the technology, familiarity, and advice acceptance.
Study Objectives

This study focused on deriving a more comprehensive understanding of the factors underlying and interactively influencing older adults’ willingness to adopt technology within a concurrent context. Specifically, participants provided appraisals based on exposure to actual technologies, in contrast to eliciting responses from older adults regarding acceptance of “general technology” [17] through broadly based questions. We used a mixed methods experimental approach referred to as the Technology Assessment Procedure (TAP), which provided study participants with some requisite exposure to a set of specific technologies.

The focus, however, was not on these specific technologies; the technologies selected were exemplars of a potentially large number of existing and future technologies across several life domains. In selecting an exemplar set of representative technologies that could be used to experimentally investigate predictors of willingness to adopt technology, our objective was to ensure that the technologies spanned different domains (e.g., transportation, health and wellness, and lifelong learning), spanned different forms (e.g., mobile apps and websites), and were potentially relevant because of their ability to improve the quality of life of older adults [1]. Simultaneously, we wanted to ensure that it was feasible and comfortable for our older study participants to evaluate the selected technologies during a single experimental session, which we validated during pilot studies.

Using the TAP methodology, we obtained participants’ ratings for each of the technologies on measures such as perceived value, improvements to quality of life, confidence in the ability to use the technology, concerns for privacy, perceived effort needed to learn the technology, and perceived help needed from family and friends to use the technology. As part of our methodology, we also examined cognitive ability measures, subjective ability measures based on self-assessments, participants’ discounting behavior, and general technology experience.

The primary goal of this study was to determine, using a concurrent framework whereby participants’ appraisals are provided within the context of actual technologies that are presented to them, those variables that are most critical in directly impacting the willingness to adopt these technologies. We also sought to establish possible meditative roles by identifying facilitating and inhibitory influences on these variables. As discussed in the ensuing section, a number of these variables were derived based on our prior findings regarding older adults’ use of technology and from cognitive aging theory. Understanding the interplay of these influences is critical, both to our theoretical knowledge concerning older adults and technology adoption and for providing a blueprint for more effective design of technology products for older populations and strategies for marketing these products to older consumers.

A Modeling Framework: Predicting Older Adults’ Willingness to Adopt Technologies

In total, 5 exemplar technologies were targeted in this study. For each of these technologies the variables examined, using the TAP methodology, included participants’ self-appraisals of the willingness to adopt the specified technology, the value or importance of the technology, the (mental) effort needed to learn and master the technology, the confidence in one’s ability to learn and master the technology, the degree to which help would be needed from family and friends to learn and master the technology, the degree to which the technology is perceived as improving one’s quality of life, concerns with issues of privacy and trust associated with the technology, and the willingness to pay for the technology.

In addition, other factors related to willingness to adopt technology that could be impacted by age were considered. These included perceived health status, openness to and readiness to take on technology, the degree to which the technology is discounted because of the investment of time needed to obtain skills on it, self-assessment of one’s cognitive abilities, cognitive abilities (based on cognitive tests), experience using computer technologies, knowledge and skills related to the use of technology, and the degree and nature of support available from family or friends for learning or using the technology.

The machine learning technique of k-fold cross-validated regressions was used to select variables that directly predicted participants’ willingness to adopt the technologies. We then used multiple regression analyses to determine the best predictors of the variables selected by k-fold cross-validated regressions. This technique is well suited to this study, as data were collected on many variables, including those that are highly correlated with each other. In addition, the models of technology adoption reviewed serve mostly as frameworks for describing the types of factors that likely influence technology adoption and are not intended to be rigorous enough to be the basis for a fully confirmatory model. We hypothesized that the perceived value of the technology, based on empirical studies involving older adults [1], would be a strong predictor of its adoption. We also predicted that confidence in one’s ability to learn the technology, which relates to the construct of self-efficacy in technology acceptance-based models (e.g., STAM), and the degree to which the technology is perceived as improving one’s quality of life would be strong predictors of willingness to adopt technology based on our past findings [1].

In addition, the cognitive effort perceived to be needed to learn and master a technology was expected to have an inhibitory influence on the intention to adopt technologies, given the general tendency for people to minimize expenditure of cognitive effort [20] and the reductions with age in “metacognitive beliefs” [21] concerning cognitive capabilities. In addition, from the perspective of learning and skill acquisition [22,23], given that older adults learn new material more slowly than younger adults, the possible requirement for a greater investment of mental effort for older adults to learn the technology may inhibit their intention to adopt it. Willingness to learn new things, which is related to the trait known as “openness to experience” and to the construct of technology readiness, was also expected to indirectly influence the willingness to adopt technologies as it typically diminishes with aging [24].
Another age-related factor that we believe would influence the willingness to adopt technologies is the extent to which rewards received later in time are discounted. For decisions based on more realistic (i.e., not hypothetical monetary) types of future rewards, Melenhorst [25] found increased discounting with age, which is consistent with economic perspectives on aging and discounting [26]. However, Sharit et al [27] found that older adults discounted less with increasing age when rewards consisted of attaining greater skills on technologies. In this context, the willingness to invest more time than someone else to achieve the same reward (i.e., level of skill in a technology) would reflect lesser discounting, similar to the willingness to wait a longer period than someone else to accrue the same amount of monetary reward. We hypothesized that lower discounting would imply greater willingness to adopt the technologies, especially if the technologies are perceived to provide improvements to the quality of life.

There may be concerns with privacy that older adults harbor, which may depend on the technology, for instance, apps that are designed to support health or financial management [1]. In addition, willingness to pay for the technology (for those technologies or apps for which such costs apply) may also influence the intention of older adults to adopt technologies [28]. The hypothesized effects of these and some additional variables are presented in Table 1.

Through the identification of variables that directly influence the willingness to adopt the technologies presented to participants and determining their possible facilitating and inhibitory indirect influences and thus possible mediating roles, as indicated in some of the examples considered above, we hoped to develop a better understanding of the interplay of influences on the adoption of technologies by older adults. Overall, the goal of this study was to measure these variables, test their hypothesized influences, and ultimately derive an efficient model that reliably captures, across a range of technologies relevant to older adults, the dynamic interplay of factors governing the willingness of older adults to adopt technologies.

### Table 1. Hypothesized effects of increases in selected study variables on willingness to adopt the technologies.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Expected effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived value</td>
<td>Positive</td>
</tr>
<tr>
<td>Confidence in ability to use the technology</td>
<td>Positive</td>
</tr>
<tr>
<td>Perceived ability for the technology to improve quality of life</td>
<td>Positive</td>
</tr>
<tr>
<td>Perceived help needed to learn the technology</td>
<td>Negative</td>
</tr>
<tr>
<td>Perceived cognitive effort needed to learn the technology</td>
<td>Negative</td>
</tr>
<tr>
<td>Technology readiness</td>
<td>Positive</td>
</tr>
<tr>
<td>Discounting of time willing to invest to learn the technology</td>
<td>Negative</td>
</tr>
<tr>
<td>Concerns with privacy</td>
<td>Negative</td>
</tr>
<tr>
<td>Willingness to pay for the technology</td>
<td>Positive</td>
</tr>
<tr>
<td>Self-assessment of abilities</td>
<td>Positive</td>
</tr>
<tr>
<td>General technology experience</td>
<td>Positive</td>
</tr>
<tr>
<td>Availability of technology assistance</td>
<td>Positive</td>
</tr>
</tbody>
</table>

### Methods

#### Participants

Participants were recruited from 2 large US cities through advertisement in local media and newsletters, interactions with agencies serving older adults, and participant registries. Interested participants completed an initial telephone interview that assessed basic eligibility, which included being ≥65 years of age; able to read and understand English at the sixth grade level; having no problems related to hearing (with correction), vision (at least 20/70 with correction), or arthritis that would impair their ability to write or use a laptop computer (only 2 people were excluded based on this criterion); being noncognitively impaired as measured by the Telephone Interview for Cognitive Status instrument [29], with cutoff scores adjusted for age and education (e.g., for people between 70 and 79 years of age, a minimal score of 29 was required for those with less than a high school education, and a minimal score of 31 was required for those with at least a high school education); and having no experience with any of the 5 technologies presented in the study.

#### Ethics Approval

The participants provided written informed consent and were compensated US $40 (and any parking expenses) for their participation. The Institutional Review Boards of the University of Miami and the Weill Cornell Medicine approved the study (approval number 1808019538).

#### Procedure

The experimental procedure used a modified version of a mixed methods data collection procedure referred to as the TAP. This method [1] involves the following: (1) presenting study participants with in-depth overviews of various technologies; (2) after each technology presentation, completing a questionnaire to rate the technology on various criteria related to its adoption; (3) completing additional questionnaires and
other assessments intended for complementing the data on technology ratings; and (4) participation in postpresentation focus groups. In this study, the focus group feature of TAP was not implemented.

Each study session involved groups of 2 to 4 people. Participants were introduced to the study, provided written informed consent, and then individually administered a demographics questionnaire, the Wide Range Achievement Test [30] to assess literacy, and a vision test. Participants who did not meet the inclusion criteria were compensated US $10 for their time. Participants who met the inclusion criteria proceeded, in a sequential order, through the ensuing session steps, with rest breaks given to them between several of these steps.

The assessment typically took approximately 4 hours. Participants were provided with snacks and drinks during the sessions, a formal break after the technology ratings were completed, and restroom breaks as needed. Although the order in which the technologies were presented and rated was randomized, the order of the instruments did not vary. As the most important measure for this study was the technology ratings, these ratings were completed first. We believed, based on prior experience, that cognitive testing would constitute the most taxing aspect of the study for the participants; thus, these measures were collected last to lessen the effect of fatigue on the other components of the study. Furthermore, as the study was not concerned with the level of the cognitive measures with respect to classifying cognitive status or abilities but rather with examining the potential impact of individual differences, this approach seemed better than adding an additional design factor whereby some participants completed the cognitive measures at the beginning of the study and were more fatigued before doing different parts of the study than other participants.

The assessment typically took approximately 4 hours. Participants were provided with snacks and drinks during the sessions, a formal break after the technology ratings were completed, and restroom breaks as needed. Although the order in which the technologies were presented and rated was randomized, the order of the instruments did not vary. As the most important measure for this study was the technology ratings, these ratings were completed first. We believed, based on prior experience, that cognitive testing would constitute the most taxing aspect of the study for the participants; thus, these measures were collected last to lessen the effect of fatigue on the other components of the study. Furthermore, as the study was not concerned with the level of the cognitive measures with respect to classifying cognitive status or abilities but rather with examining the potential impact of individual differences, this approach seemed better than adding an additional design factor whereby some participants completed the cognitive measures at the beginning of the study and were more fatigued before doing different parts of the study than other participants.

Technology Presentations and the Technology Ratings Questionnaire

Overview

Participants as a group were shown PowerPoint presentations on 5 technologies in a predetermined random order to minimize order effects. The five technologies were (1) Lyft, a ride-sharing app; (2) eCareCompanion, an app that allows sharing of health information with your care team, tracking of health tasks, and optional devices to measure vital statistics; (3) Curious, a website dedicated to providing lessons for lifelong learners on a variety of topics; (4) InteliChart, a patient portal that allows an individual to view medical charts, schedule appointments, and manage other aspects of health care; and (5) Fittle, an app that uses an internet-based coach to help people meet health and fitness goals. Each presentation lasted for approximately 10 minutes. Participants were allowed to ask clarifying questions about each technology; however, discussion among the participants was not permitted. Figure 1 shows examples of the slides used in the presentation of the technologies.

Following the presentation of each technology, participants completed a technology rating questionnaire in which they rated the technology on various criteria using a Likert-type 9-point scale (except for the willingness-to-pay criterion), with verbal descriptors provided for the 2 endpoints and the 3 intermediary points on the scale. After the presentations on all 5 technologies were completed, a summary of the 5 technologies was presented, and participants were able to review their ratings and make changes if desired. The participants rated each technology based on the following criteria.

Figure 1. Example of slides used in the presentations of the technologies to participants going clockwise from the top left: Lyft, eCareCompanion, Curious, Fittle (bottom right), and InteliChart.

Willingness to Adopt the Technology

How willing are you to adopt the technology that was just presented to you? In other words, how willing are you to “take it up” and start using it (1 being completely unwilling, 5 being somewhat willing, and 9 being completely willing)? This was the primary dependent variable in the analyses. We also measured willingness to adopt using paired comparisons within the Analytic Hierarchy Process [31,32]. Although the paired comparisons measure validated our willingness to adopt measures, we chose not to use this measure in our analyses because it was highly correlated with the willingness to adopt measure and had lesser correlation with other measures in our analyses. Thus, we only used the willingness to adopt this measure.
**Perceived Value**
How would you rate the importance or value of the technology that was just presented to you (1 being not at all important, 5 being somewhat important, and 9 being extremely important)?

**Perceived Effort**
How much effort do you think you would have to put into learning and mastering the technology that was just presented to you (1 being none, 5 being some, and 9 being a lot)?

**Confidence in Using the Technology**
How confident are you that you have the ability to learn and master the technology that was just presented to you (1 being not at all confident, 5 being somewhat confident, and 9 being extremely confident)?

**Help With Technology**
How much help would you need from family and friends to learn and master the technology that was just presented to you (1 being none, 5 being some, and 9 being a lot)?

**Quality of Life**
Think about the things that are most important to you that contribute to your quality of life. How much can the technology just presented to you help improve your quality of life (1 being not at all, 5 being some, and 9 being a lot)?

**Concern With Privacy**
How worried or concerned are you about privacy and trust issues associated with the technology just presented to you (1 being not at all worried, 5 being somewhat worried, and 9 being extremely worried)?

**Willingness to Pay**
Select how much you would be willing to pay to own the technology just presented to you. When making your selection, do not include any ongoing or recurring costs for services or subscriptions associated with the technology. (Participants could choose from nothing to more than US $100, with 10 categories in between these extremes).

**Discounting the Investment of Time to Learn the Technologies**
The experimenter, following a script, guided the administration of 2 complementary instruments designed for assessing discounting behavior that were presented to the individual participants on the laptops provided to them. First, the Time Allocation to Attain Skill instrument was used for participants to indicate the amount of time (in hours and minutes) that they would be willing to spend to achieve a certain level of skill on the specified technology. Five levels of skill—basic, moderate, intermediate, advanced, and mastery—were defined, and training was given to clarify differences among these skill levels. Participants responded either yes or no regarding their desire to attain the next level of skill in the technology. If their choice was “yes,” they were also asked to indicate the additional amount of time they would be willing to invest to achieve that skill level; however, they could choose to stop at any level if they did not desire to acquire any further skill in that technology.

After completing this instrument for each of the 5 technologies, the participants completed the Assigning Importance to a Skill Level instrument. Using their laptops, they were instructed to rate the importance of attaining the desired skill levels that they had previously indicated for each of the 5 technologies on a scale that ranged from 1 to 10 (1 indicated no importance, 5 indicated average importance, and 10 indicated extremely important). Participants were cued (by the computer interface) to assign importance values only for those skill levels for which they indicated that they were willing to invest time to attain.

These 2 instruments enabled the collection of data for determining the degree to which participants discounted the time they were willing to invest to acquire skills for each of the 5 technologies [27]. In addition, the level of skill desired, defined as the highest level of skill participants wished to attain for each technology (ranging from 1 for basic skills to 5 for a skill level of mastery) was also used as a measure as desire to acquire greater skills was believed to be indicative of willingness to adopt the technology.

**Additional Instruments**

**Cognitive Test Battery**
Participants were administered the Trail Making Tests A and B [33], which measure overall cognitive functioning; Digit Span, forward and backward [34], which measures working memory; the Shipley Vocabulary test [35], which measures crystallized and fluid cognitive ability; and the Multidimensional Aptitude Battery [36] test, which measures life knowledge.

**Self-assessment of Abilities**
Participants completed an 8-item rating scale, adapted from Ackerman and Wolman [37], which was used to assess their self-appraisal of the following abilities on a 9-point scale (1=very low ability; 9=very high ability): vocabulary, comprehension, numeric ability, memory, learning ability, problem-solving and reasoning, detection, and grasping and manipulative skill.

**Openness to New Experiences**
Participants answered 2 questions from the Ten-Item Personality Inventory [38] related to traits associated with being open to new experiences.

**Perceptions of Aging**
Participants answered the 10-item Attitudes Toward Age-Related Change [39], which is divided into two 5-item sections measuring the perceptions of positive and negative aspects of aging. Each item ranges from 1 to 5 or not at all to very much.

**Technology Readiness**
Participants completed the Technology Readiness Index, a 16-item questionnaire that uses a 5-point Likert scale designed to determine an individual’s predisposition to adopting new technologies [40] and thus capture potentially important attitudes related to technology acceptance. It comprises 2 positive dimensions, optimism (belief that technology increases control, flexibility, and efficiency) and innovativeness (one’s view of being a “technology pioneer”), and 2 negative dimensions,
discomfort (a tendency to being uncomfortable with or overwhelmed by technology) and insecurity (a general feeling of skepticism or fear toward technology). Participants were asked to what extent they agree or disagree with 16 statements across the 4 dimensions.

**General Technology Use Survey**

This instrument was divided into three sections: (1) a section which asked participants about their access to and use of computer technologies, such as desktops, laptops, and tablets, smartphones, internet, and email; (2) a section comprised of 4 questions to which participants indicated, on a 9-point scale (1=very low; 9=very high), their degree of basic computer technology skill (eg, the ability to easily use the equipment associated with basic computer technologies such as a keyboard or a mouse); internet and email skill or knowledge; computer programs knowledge; and computer applications knowledge (eg, about different applications or “apps” on a computer or smartphone and how to use them); and (3) a section related to participants’ needs for assistance and support in use of technology. This last section consisted of 2 questions to which participants responded, using a 9-point scale (1=none of the time, 5=some of the time, and 9=all the time), how often they needed assistance to help them learn and master a new technology and how often someone was available to them to learn and master a new technology. In addition, participants were asked to check off items indicating who they relied on for help learning to use a new technology and who they listen to for advice and recommendations when considering whether to use a new technology.

Table 2 summarizes the measures used in this study as well as the instruments from which they were obtained.
### Table 2. Variables collected in the study.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Type of variable(^a)</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Individual (continuous)</td>
<td>Demographics questionnaire</td>
</tr>
<tr>
<td>Perceived overall health</td>
<td>Individual (scale value: 1-5)</td>
<td>Demographics questionnaire</td>
</tr>
<tr>
<td>Willingness to adopt the technology</td>
<td>Technology (scale value: 1-9)</td>
<td>Technology Rating Questionnaire</td>
</tr>
<tr>
<td>Perceived value of the technology</td>
<td>Technology (scale value: 1-9)</td>
<td>Technology Rating Questionnaire</td>
</tr>
<tr>
<td>Perceived effort to learn the technology</td>
<td>Technology (scale value: 1-9)</td>
<td>Technology Rating Questionnaire</td>
</tr>
<tr>
<td>Self-confidence in ability to learn and use the technology</td>
<td>Technology (scale value: 1-9)</td>
<td>Technology Rating Questionnaire</td>
</tr>
<tr>
<td>Perceived help needed to learn the technology</td>
<td>Technology (scale value: 1-9)</td>
<td>Technology Rating Questionnaire</td>
</tr>
<tr>
<td>Quality of life improvement from the technology</td>
<td>Technology (scale value: 1-9)</td>
<td>Technology Rating Questionnaire</td>
</tr>
<tr>
<td>Privacy or trust issues with the technology</td>
<td>Technology (scale value: 1-9)</td>
<td>Technology Rating Questionnaire</td>
</tr>
<tr>
<td>Willingness to pay to own the technology</td>
<td>Technology (categorical)</td>
<td>Technology Rating Questionnaire</td>
</tr>
<tr>
<td>Relative comparisons in adopting the technologies</td>
<td>Technology (a set of relative weights of each technology)</td>
<td>Paired Comparison Ratings Instrument</td>
</tr>
<tr>
<td>Discounting rate</td>
<td>Technology (continuous)</td>
<td>Time Allocation to Attain Skill Instrument</td>
</tr>
<tr>
<td>Level of skill desired</td>
<td>Technology (scale value: 1-5)</td>
<td>Time Allocation to Attain Skill Instrument</td>
</tr>
<tr>
<td>Overall cognitive functioning</td>
<td>Individual (test score)</td>
<td>Cognitive Test Battery (Trail Making Tests A and B)</td>
</tr>
<tr>
<td>Working memory</td>
<td>Individual (test score)</td>
<td>Cognitive Test Battery (Digit Span)</td>
</tr>
<tr>
<td>Crystallized and fluid cognitive ability</td>
<td>Individual (test score)</td>
<td>Cognitive Test Battery (Shipley Vocabulary)</td>
</tr>
<tr>
<td>Life knowledge</td>
<td>Individual (test score)</td>
<td>Cognitive Test Battery (Multidimensional Aptitude Battery)</td>
</tr>
<tr>
<td>Self-assessment of cognitive abilities</td>
<td>Individual (average score; item scale value: 1-9)</td>
<td>Self-Assessment of Abilities Questionnaire (6 of 8 items)</td>
</tr>
<tr>
<td>Openness to new experiences</td>
<td>Individual (average score; item scale value: 1-7)</td>
<td>Ten-Item Personality Inventory (2 of the items)</td>
</tr>
<tr>
<td>Perceptions of aging: gain</td>
<td>Individual (sum score; item scale value: 1-5)</td>
<td>Perceptions of Aging (5 of 10 items)</td>
</tr>
<tr>
<td>Perceptions of aging: loss</td>
<td>Individual (sum score; item scale value: 1-5)</td>
<td>Perceptions of Aging (5 of 10 items)</td>
</tr>
<tr>
<td>Technology readiness: optimism, innovativeness, discomfort, and insecurity</td>
<td>Individual (average total and subscale scores; item scale value: 1-5)</td>
<td>Technology Readiness Index Questionnaire</td>
</tr>
<tr>
<td>Self-assessment of technology skill</td>
<td>Individual (average of 5 items; item scale value: 1-9)</td>
<td>General Technology Use Survey</td>
</tr>
<tr>
<td>Needs or availability for technology assistance support</td>
<td>Individual (average of items; item scale value: 1-9)</td>
<td>General Technology Use Survey</td>
</tr>
<tr>
<td>General tech experience</td>
<td>Individual (sum score of yes or no to having used 5 technologies)</td>
<td>General Technology Use Survey</td>
</tr>
</tbody>
</table>

\(^a\)Technology variables were collected for each of the 5 technologies; individual variables were collected once.

### Analytic Approach

As noted earlier, because of the large number of variables and the difficulty of specifying a predictive model a priori with such a large number of parameters and potential collinearity, we adopted a systemic exploration analytic method based on machine learning techniques, the k-fold cross-validation regression technique, to derive our model of willingness to adopt the technologies. We chose this technique as it helps reduce model overfitting and provides a better estimate of how our derived model would perform in general, beyond the data generated by our sample [41].

Initially, because of missing data (<1% of total observations), for some of the variables, we used multiple imputation to create 20 different complete data records for each of the 187 participants. For each participant, each of these data records contained complete data for all variables, with the prior missing data replaced by imputed values. Thus, we generated 20 data records for each participant, each with 20 different imputed values for each variable with missing data. Although typically 5 imputations for missing data are considered sufficient, we opted to be conservative and instead created 20 different data records for each participant for reasons involving our variable selection method explained in the following paragraphs.
Following imputation and the generation of 20 data records, the next step was to identify the set of variables that would best predict our main dependent variable, willingness to adopt the technologies. We tested each data record with a k-fold cross-validation regression program as implemented in \textit{glmnet} for the R statistical environment [42]. This program uses a penalized regression technique to handle collinearity and is consistent with the techniques of Ridge regression and Lasso regression. The k-fold cross-validation estimates model parameters on part of the data and then validates those parameters on a separate subsample not used to estimate the parameters. The program attempts to find the set of parameters that best fits the separate subsample while varying the lambda penalty (a value that shrinks regression parameters toward 0) from 0 (equivalent to a Ridge regression) to 1 (equivalent to a Lasso regression).

Once the 20 k-fold cross-validated regressions were computed, we recorded the number of times across the 20 data records each variable was predictive and the average parameter value each time it was predictive. The criterion we adopted was that variables would be selected for further exploration if they were found to be significant in at least half (10/20, 50%) of the regression models, as this would result in a model that is more generalizable and less biased. We viewed variables that were predictive in all 20 data records as more likely to be producing a replicable effect than those that were not predictive in all 20 data records. We also viewed those variables that were predictive more often than not as more likely to replicate in future research than those that were not predictive more often than not, and we encourage the reader to use the same heuristic.

Although we report model data for those variables that were selected by the model in <10 of the data records based on the previous reasoning, we do not provide an interpretation of the parameters for these variables.

We then conducted a series of regression analyses for the purpose of enhancing our understanding of the variables predicting willingness to adopt (eg, understand the effect sizes associated with the variables), using the set of variables selected by the k-fold cross-validated regressions.

However, consistent with our study objectives in understanding the facilitating and inhibitory roles of various variables in influencing technology adoption willingness, we were also interested in understanding if the strongest predictors of willingness to adopt technologies were potentially mediating the relationships of other variables (eg, crystallized intelligence and technology readiness) related with willingness to adopt technologies. For example, the findings from our analysis of willingness to adopt indicated that it was most strongly predicted by the participants’ ratings of improvement in quality of life from the technology, perceived value of the technology, and confidence in using the technology. To examine the potential mediating role of these 3 variables, we again performed k-fold cross-validated regressions in \textit{glmnet}, this time with the ratings of quality of life, perceived value, and confidence serving as the dependent variables.

Having identified a set of predictors for each of the 3 variables using cross-validation, we again conducted separate regression analyses on each of the 3 primary predictors of willingness to adopt technologies, once again with the goal of getting the more intuitive output with effect sizes and statistical significance. For example, as an illustration of this analytic process, in the k-fold regressions, crystallized ability had no direct relationship with adopting technologies, once again with the goal of getting the more intuitive output with effect sizes and statistical significance.

Results

Overview

The study sample included 187 adults aged 65 to 92 (mean 74.1, SD 6.3) years, who were primarily women (145/187, 77.5%); diverse in age, with 41.1% (77/187) of the participants aged ≥75 years; and diverse in ethnicity and race—21.3% (40/187) of the participants identified as Hispanic and 35.8% (67/187) identified as Black or African American. Most participants (157/187, 83.9%) reported having at least some college education, 82.9% (155/187) reported being retired, and 89.8% (168/187) self-reported their health as at least good. Table 3 includes descriptive statistics for the sample demographics. Table 4 shows the results of the k-fold cross-validated regressions with the parameter estimate and the number of multiply imputed data sets for the parameter that was selected as predictive; again, we urge caution in interpreting variables that were not selected in most of the models. For each of the 20 models, the following 4 variables were predictive of higher ratings of willingness to adopt the technologies: higher ratings of perceived value of the technologies, higher ratings of perceived improvement in quality of life by the technologies, higher rating of confidence in using the technology. To examine the potential mediating role of these 3 variables, we again performed k-fold cross-validated regression analyses on each of the 3 primary predictors of willingness to adopt technologies, once again with the goal of getting the more intuitive output with effect sizes and statistical significance.
Table 3. Demographics of the sample (n=187).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tech experience, mean (SD)</td>
<td>3.87 (1.50)</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>74.11 (6.33)</td>
</tr>
<tr>
<td>General health, mean (SD)</td>
<td>3.43 (0.85)</td>
</tr>
<tr>
<td>Gender (women), n (%)</td>
<td>145 (77.5)</td>
</tr>
</tbody>
</table>

**Education, n (%)**

- High school or less: 22 (11.8)
- Some college or associates: 51 (27.4)
- Bachelor’s degree: 48 (25.8)
- Postgraduate: 58 (31.2)
- Vocational: 7 (3.8)
Table 4. Results of 20 multiple imputed cross-validated regression.

<table>
<thead>
<tr>
<th>Willingness to adopt technology</th>
<th>Quality of life from technology</th>
<th>Perceived value of technology</th>
<th>Confidence using technology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M&lt;sup&gt;a&lt;/sup&gt;</td>
<td>β</td>
<td>M</td>
</tr>
<tr>
<td>Intercept</td>
<td>20</td>
<td>−.11</td>
<td>20</td>
</tr>
<tr>
<td>Tech readiness optimism</td>
<td>—&lt;sup&gt;b&lt;/sup&gt;</td>
<td>—</td>
<td>20</td>
</tr>
<tr>
<td>Tech readiness innovativeness</td>
<td>15</td>
<td>.01</td>
<td>—</td>
</tr>
<tr>
<td>Tech readiness insecurity</td>
<td>14</td>
<td>.00</td>
<td>2</td>
</tr>
<tr>
<td>Positive tech readiness</td>
<td>18</td>
<td>.01</td>
<td>19</td>
</tr>
<tr>
<td>General tech experience</td>
<td>20</td>
<td>.06</td>
<td>—</td>
</tr>
<tr>
<td>Gender (women)</td>
<td>8</td>
<td>−.06</td>
<td>—</td>
</tr>
<tr>
<td>Education</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>General health</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Self-assessed comprehension</td>
<td>—</td>
<td>—</td>
<td>11</td>
</tr>
<tr>
<td>Self-assessed learning ability</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Self-assessed applying new knowl-</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>edge</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Self-assessed problem-solving or</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>reasoning</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Cognitive abilities: fluid</td>
<td>—</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td>Cognitive abilities: crystalized</td>
<td>—</td>
<td>—</td>
<td>18</td>
</tr>
<tr>
<td>Perceptions of aging: gains</td>
<td>—</td>
<td>—</td>
<td>12</td>
</tr>
<tr>
<td>Openness to experience</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Help with technology</td>
<td>20</td>
<td>−.03</td>
<td>—</td>
</tr>
<tr>
<td>Confidence using technology</td>
<td>20</td>
<td>.15</td>
<td>N/A&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Concern with privacy of technology</td>
<td>—</td>
<td>—</td>
<td>1</td>
</tr>
<tr>
<td>Perceived effort of technology</td>
<td>—</td>
<td>—</td>
<td>20</td>
</tr>
<tr>
<td>Perceived value of technology</td>
<td>20</td>
<td>.54</td>
<td>N/A</td>
</tr>
<tr>
<td>Quality of life from technology</td>
<td>20</td>
<td>.24</td>
<td>N/A</td>
</tr>
<tr>
<td>Discounting parameter&lt;sup&gt;d&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
<td>15</td>
</tr>
</tbody>
</table>

<sup>a</sup>M: number of models in which the parameter was included.

<sup>b</sup>—: variable was not selected as predictive in any data record for this dependent variable.

<sup>c</sup>N/A: not applicable.

<sup>d</sup>Tech readiness discomfort, negative tech readiness, age, self-assessed vocabulary, self-assessed numeric, self-assessed memory, self-assessed grasping, and perceptions of aging losses were included in the model; however, they are not presented because they were not predictive in any model computed.

Factors Effecting Willingness to Adopt Technology

The k-fold analyses on willingness to adopt technologies found that the following variables were predictive in ≥10 of the 20 imputed data sets: perceived value of the technologies, perceived improvement of quality of life by the technologies, confidence using the technologies, perceived help needed with the technologies, the innovativeness component of technology readiness, the insecurity component of technology readiness, positive technology readiness, and technology experience. In the multiple regression analysis, these 8 variables selected by cross-validation explained a large amount of variance ($F_{8,178}=59.7$, $P<.001$, $R^2=0.73$). Five of these variables were significantly predictive in the multiple regression. For each statistically significant predictor, we report first the regression parameter and then the zero-order correlation. The zero-order correlation is reported to allow the reader to see how the predictor relates to the target variable, in this case willingness to adopt technology, without the rest of the variables and thus how the other variables included in the multiple regression are altering the effect size. Of these, the perceived value of the
technologies was the strongest predictor ($\beta=0.60, P<.001; r=0.77, P<.001$), followed by perceived improvement in quality of life by the technologies ($\beta=0.32, P<.001; r=0.71, P<.001$). Confidence using the technologies was the next strongest predictor ($\beta=0.15, P=.009; r=0.45, P<.001$), followed by perceived help needed with the technologies ($\beta=-0.10, P=0.02; r=0.17, P<.001$) and then technology experience ($\beta=0.12, P=.03; r=0.30, P<.001$). Therefore, individually, perceived value and improvement in quality of life accounted for >50% of the variance in willingness to adopt the technologies, confidence accounted for 20%, and technology experience accounted for 9%. It is also notable that in this model of only 8 predictor variables, the technology readiness variables were not significant, although they were selected in over half of the models as improving the models' predictive accuracy.

**Factors Effecting Perceived Value of Technologies**

As noted, given the strong relationship of willingness to adopt technologies with perceived value, improvement in quality of life, and confidence in using the technologies, we hypothesized that these variables were potentially mediating the relationships of other important variables. The findings from the k-fold cross-validation analyses indicated that across all 20 imputed data records, higher scores in the optimism component of the technology readiness scale and higher scores in perceived effort needed to learn the technologies were positively related to perceptions regarding the impact of the technologies on perceived value.

The results of the multiple regression analyses indicated that these 2 variables together strongly positively predicted perceived value ($F_{2,184}=29.88, P<.001, R^2=0.25$). The individual parameters were $\beta=0.19 (P<.001)$ and $r=0.43 (P<.001)$ for the optimism component of technology readiness and $\beta=0.21 (P<.001)$ and $r=0.25 (P=0.01)$ for perceived effort required by the technologies.

**Factors Effecting Quality of Life From Technologies**

With respect to perceptions of improvements in quality of life from the technologies, the models across all 20 data records selected the optimism component of technology readiness and perceived effort required by the technology, with higher values in both being associated with higher quality of life from adopting the technologies. Selected by most models but not all models as predicting higher levels of quality of life from the technology were higher levels of positive technology readiness, lower crystallized intelligence, lower values in the discounting of the technology within the context of concrete technologies, we considered the administration of various tools intended for assessing other variables predicting higher confidence using the technologies. With respect to confidence in using the technologies, the variables predicting higher confidence using the technology found by a k-fold cross-validation regression were higher crystallized intelligence, technology experience, self-assessed applying of new knowledge, self-assessed problem-solving or reasoning, general education, scores on the optimism component of technology readiness, and needing less help with technologies. Selected by most models but not all models as predicting higher confidence using the technologies were the following variables: higher self-assessed comprehension, higher self-assessed detection, higher openness to experience, and higher self-assessed learning ability. Interestingly, however, although the multiple regression on this set of variables predicting higher confidence explained a large amount of variance ($F_{11,175}=13.14, P<.001, R^2=0.45$), likely due to collinearity, only 2 variables were predictive in the multiple regression. Lower scores in perceived help needed with the technologies ($\beta=-0.25, P<.001; r=0.31, P<.001$) and higher scores in the optimism component of technology readiness ($\beta=0.07, P=.02; r=-0.54, P<.001$) were predictive of higher confidence in using the technologies.

**Factors Effecting Confidence in Using Technologies**

With respect to confidence in using the technologies, the variables predicting higher confidence using the technology found by a k-fold cross-validation regression were higher crystallized intelligence, technology experience, self-assessed applying of new knowledge, self-assessed problem-solving or reasoning, general education, scores on the optimism component of technology readiness, and needing less help with technologies. Selected by most models but not all models as predicting higher confidence using the technologies were the following variables: higher self-assessed comprehension, higher self-assessed detection, higher openness to experience, and higher self-assessed learning ability. Interestingly, however, although the multiple regression on this set of variables predicting higher confidence explained a large amount of variance ($F_{11,175}=13.14, P<.001, R^2=0.45$), likely due to collinearity, only 2 variables were predictive in the multiple regression. Lower scores in perceived help needed with the technologies ($\beta=-0.25, P<.001; r=0.31, P<.001$) and higher scores in the optimism component of technology readiness ($\beta=0.07, P=.02; r=-0.54, P<.001$) were predictive of higher confidence in using the technologies.

**Discussion**

**Summary**

The importance of identifying and understanding the factors that contribute to people adopting technology stems from the considerable benefits that numerous powerful and easily accessible technologies can potentially provide. For many older adults, these benefits may be especially beneficial, as these individuals may be experiencing cognitive and physical limitations or declines in health states related to social isolation or declines in functional abilities, which could possibly be circumvented or even overcome with the aid of technologies. However, research related to technology adoption for this population of adults, unlike prior studies on technology adoption for younger populations, has been limited to identifying factors based on adults’ responses to questionnaire items directed at general technology adoption. In this study, we implemented an innovative approach referred to as the TAP, which was intended to capture study participants’ willingness to adopt each of the 5 technologies included in this study. This methodology enabled participants to assess their perceptions related to specific technologies by providing them, through a series of formal presentations, basic information associated with a set of technologies that were believed to provide unique benefits to the health and well-being of older adults [1]. In addition to being able to assess the willingness to adopt technology within the context of concrete technologies, we incorporated an empirical experimental perspective that included the administration of various tools intended for assessing other variables that were hypothesized to either facilitate or inhibit the behavioral intentions underlying the primary dependent variable, the willingness to adopt technology, and a machine learning approach in the selection of predictor variables coupled with regression analysis for assessing the possible mediating roles of those predictors. Furthermore, our sample included a...
Implications for Models of Technology Adoption

A robust study result was the finding that higher ratings on the following 4 variables predicted higher ratings of willingness to adopt the technologies, and higher ratings on a fifth variable predicted lower ratings of willingness to adopt the technologies. These variables consisted of perceived value of the technologies, perceived improvement of quality of life by the technologies, perceived confidence in using the technologies, higher scores on technology experience, and higher ratings of perceived help needed with the technologies (which was predictive of lower ratings of willingness to adopt the technologies). The results for perceived value, improvement in quality of life, and confidence are consistent with past studies of the TAM [10,11] and STAM [16,17], as well as with a study that implemented the TAP method [1]. Davis argued that intention to use technology was determined by perceived ease of use and perceived usefulness. The largest predictors that we found, perceived value and perceived improvements in quality of life, could be interpreted as corresponding to the construct of perceived usefulness, and our next strongest predictors, confidence in using the technologies and perceived help needed with the technologies, could be construed as relating to perceived ease of use. In the STAM [16,17], which focused on older adults (aged >55 years), the factors found to predict general technology acceptance were relatively broad, such as self-efficacy, gerontechnology anxiety, and health and ability characteristics, and could be viewed as related to the predictor of confidence in using the technologies.

Unlike the STAM and TAM, this study was based on a concurrent perspective which used presentations of concrete contexts of specific technologies as a basis for providing appraisals related to willingness to adopt technology. Therefore, we suggest that the variables perceived value and perceived improvements in quality of life, although clearly related conceptually, are distinctive predictors. Because participants were given the opportunity to consider the present adoption of the technology, they may have been able to dissociate general usefulness or value from more specific ways in which the quality and independence of their lives could be influenced by the technologies. Thus, we also suggest that the perceived value associated with a technology captures an appraisal more closely linked to the general appraisal of a technology as useful, whereas perceived improvements in quality of life represent appraisals that enable technologies to be more differentiable based on the extent to which they might positively modify one’s life. Similarly, although self-efficacy resembles the degree of confidence in one’s ability to learn and master the technology (the variable used in our study), the self-appraisals of confidence collected following presentations on each technology may be more dependent on the nature of the specific technology than on a more general self-assessed state of self-efficacy. Greater emphasis on current confidence in using a specific technology, as opposed to the more general retrospective assessment of one’s self-efficacy, is more consistent with the perspective of Lee and Coughlin [43] on the importance of confidence in older adults’ adoption of technology and, as will be discussed later, provides a more direct bridge to strategies intended to market technologies for older users.

The use of concrete contexts as a basis for assessing technology adoption also likely influenced the finding that a greater perceived need for help was found to reduce willingness to adopt the technology, as this information would provide the participant with a greater understanding of the predicaments in which they might find themselves when attempting to adopt or use a specific technology without available support. Finally, the positive relationship between general technology experience and willingness to adopt the technologies, although not related to a self-appraisal linked to the presentation of a specific technology, is also informative, as it provides for an assessment of self-efficacy as it pertains to technology use.

However, to establish both a more comprehensive understanding of factors predicting technology adoption for older adults and to develop strategies for marketing technologies to increase the likelihood of their adoption by these users, other variables need to be considered that could have facilitated or inhibited the possible mediating roles the primary predictors discussed above had on willingness to adopt the technologies. The most robust finding (ie, across all 20 imputed data records) was that for both perceived value and improvements in quality of life, the optimism component of the technology readiness scale and perceived effort needed to learn the technologies each had a positive influence on perceived value and improvements to quality of life from the technology. Optimism in technology readiness (the belief that technology increases control, flexibility, and efficiency) appears to be the more critical of the 2 positive dimensions on this scale, and within the context of considering adopting specific technologies (such as those considered in this study), likely to represent a powerful attitudinal perspective to the behavioral intention to adopt a technology and thus underly ratings of perceived value and improvements to quality of life.

Perhaps less intuitive to the possible mediating roles of these 2 predictors of willingness to adopt technologies is the positive relationship the perceived effort needed to learn the technologies had with them. Earlier, we had hypothesized that increased perceived cognitive effort needed to learn technologies would negatively impact willingness to adopt, as people tend to minimize cognitive effort. However, the overall findings may suggest that if older users demonstrate optimism in their technology readiness attitudes, this may override the tendency to avoid investing effort in learning, especially if the technologies are perceived as capable of providing improvements to quality of life.

Other variables found in most (but not all) of the 20 imputed data sets that significantly influenced improvements in quality of life included crystallized intelligence, discounting of the technologies, and perceived health, with lower values for each of these variables positively associated with improvements in quality of life and willingness to adopt the technologies. For this study sample, increased age was found to be associated with decreased discounting [27], suggesting that the greater time older adults were willing to invest to achieve higher skill levels in these technologies may be linked to their perceptions.
that attaining these higher-level skills could translate to improvements in the quality of their lives. The weaker relationships that were found between lower levels of crystallized intelligence and lower perceived health on perceptions of increased improvements to quality of life and willingness to adopt the technologies are less understood; they may suggest self-awareness by these participants of the need to compensate for these lower cognitive and health levels through technologies that could potentially benefit their health and well-being.

Expectedly, confidence in having the ability to learn the technologies as a positive predictor of willingness to adopt them was predicted by higher self-assessments of abilities such as comprehension and learning abilities, higher scores in the optimism component of the technology readiness scale, and lower ratings of perceived need for help to learn to use the technologies. Taken together with the findings for the other 2 main predictors of willingness to adopt the technologies, perceived value, and improvements in quality of life, some strategies for inducing adoption of technologies by older adults are suggested. For example, in marketing these technologies and developing methods for instruction on their use, emphasis should be given to very specific ways the technology can benefit independence and quality of life and how efficient these technologies can be in meeting these goals. Although designing technological products that are easy for older adults to use is critical [44], if these designs are usable, older adults are likely to not be deterred if cognitive effort, within reason, is needed to learn to use the technologies. In addition, they may be willing to invest additional time to attain higher levels of mastery, provided the benefits of the technology are evident. Messages that promote optimism in technology are also recommended as they provide the basis for positive underlying attitudes.

Limitations

The main limitation of this study is that it is a concurrent, cross-sectional study, and participants did not have the opportunity to actually engage with the technologies. Thus, although a great deal of attention was paid to familiarize the participants with each technology so that participants could provide responses that were as accurate as possible about their willingness to adopt each of the technologies that were presented, collecting prospective real-world data on actual use patterns would be preferred, but this was beyond the scope of this study.

However, this study overcame many limitations associated with retrospective, questionnaire-based data by providing more realistic contexts for assessing technology adoption and a larger array of variables informed by an understanding of the cognitive capabilities and limitations of older adults. In addition, 2 major strengths of the study are the comprehensive explanation and walkthrough of the target technology, which we feel is necessary for participants to have an accurate understanding of the technologies, and our rigorous measurement of their perceptions and many related constructs. However, these comprehensive explanations and measurements limited us to only the 5 technologies selected. Although we feel the findings of our study likely generalize beyond these specific technologies, that must be confirmed by future research.

In addition, because the study used only 5 technologies, we were not able to study the differences in technology adoption between technologies within the scope of this project. Future research should expand upon the technologies used here and potentially look at heterogeneity in what predicts technology adoption between different types of technologies. In addition, the participants self-selected into the study, and the study advertisement stated that the study was about technology and might require participants to travel to the University of Miami-Miller School of Medicine or Weill Cornell Medicine. This may have impacted the sample recruitment. In this regard, the sample was likely healthier, more interested in technology, and more educated and thus not representative of the diverse population of older adults living independently. In addition, we could not include people with cognitive deficits because of the nature of the study requirements.

Conclusions

This study provided a conceptual basis for identifying variables that could influence older adults’ willingness to adopt technology and used a concurrent framework whereby participants’ appraisals regarding their willingness to adopt technology were made within the context of exposure to 5 exemplary technologies with potential benefits to older populations. The analytic approach taken enabled direct predictors of willingness to adopt technology, as well as variables that had inhibitory and facilitating influences on these predictors to be determined. Future research examining the issues of technology adoption could benefit from the methods used in this study and examining the complex patterns of relationships found. On the basis of the variables identified as important in this study, in future studies, it should be easier to select the number of variables for investigation and further expand our causal model. However, the ultimate criterion for a model of technology acceptance among older adults is the longitudinal measurement of the use of technology in the naturalistic environment, which for many reasons remains a challenging problem.

Acknowledgments

This study was funded by the National Institute on Aging and National Institutes of Health (NIA 3 PO1 AG17211, Project CREATE IV—Center for Research and Education on Aging and Technology Enhancement). The authors are thankful to Ronald Berkowsky and Mario Hernandez for their contribution to this study. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.
Data Availability

Data are available upon request. The authors are currently merging data sets on technology adoption across the Center on Aging at the University of Miami and the Center for Aging and Behavioral Research at Weill Cornell Medicine.

Conflicts of Interest

None declared.

References


Abbreviations

STAM: Senior Technology Acceptance Model
TAM: Technology Acceptance Model
TAP: Technology Assessment Procedure
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by T Leung; submitted 26.05.22; peer-reviewed by M Rampioni, E Boucher, O Ogundaini, N Sahari; comments to author 10.07.22; revised version received 31.07.22; accepted 30.08.22; published 23.11.22.

Please cite as:
Moxley J, Sharit J, Czaja SJ
The Factors Influencing Older Adults' Decisions Surrounding Adoption of Technology: Quantitative Experimental Study
JMIR Aging 2022;5(4):e39890
URL: https://aging.jmir.org/2022/4/e39890
doi: 10.2196/39890
PMID: 36416885
An Online Dual-Task Cognitive and Motor Exercise Program for Individuals With Parkinson Disease (PD3 Move Program): Acceptability Study

Josefa Domingos¹,²,³, MSc; John Dean², MSc; Júlio Belo Fernandes³, PhD; Catarina Godinho³, PhD

¹Department of Neurology, Radboud University Medical Center, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, Netherlands
²Triad Health AI, Aurora, CO, United States
³Grupo de Patologia Médica, Nutrição e Exercício Clínico (PaMNEC), Centro de Investigação Interdisciplinar Egas Moniz, Almada, Portugal

Corresponding Author:
Josefa Domingos, MSc
Department of Neurology
Radboud University Medical Center
Donders Institute for Brain, Cognition and Behaviour
6500 HB Nijmegen
Nijmegen, 1495-688 Cruz Quebrada
Netherlands
Phone: 31 +351967142254
Email: domingojs Josefa@gmail.com

Abstract

Background: Dual-task training is an emerging field used for people with Parkinson disease (PD) to improve their physical and cognitive well-being, but the patients’ acceptability, safety, and adherence to such training in online settings are unknown.

Objective: This study aims to evaluate the acceptability of a dual-task cognitive and motor online training program for people with PD as a group online community program.

Methods: People with PD were invited to participate in an online program (PD3 Move) consisting of physical and vocal exercises in response to different cognitive challenges displayed as dynamic backgrounds on Zoom. The program ran twice per week for 16 weeks. Patient acceptability was assessed at 4 months by monitoring attendance rates and feedback from an exit questionnaire emailed to all participants assessing satisfaction, perceived benefit, safety, and willingness to continue and recommend to others.

Results: The online program was delivered to 15 participants (n=9, 60%, females) with a diagnosis of PD, a mean age of 69.4 (SD 9.3) years, and Hoehn and Yahr (H&Y) stages I-IV. The attendance rate was high, with participants coming to more than 13 (81%) of the sessions. Participants were very satisfied (n=8, 53%) or satisfied (n=7, 47%) with the program. Participants reported that what they most liked were the new cognitive physical challenges. The 3 main facilitators to participating were perceiving the benefits, instructor’s flexibility and engagement, and the social interaction moments with others. The 3 main difficulties were dealing with motor fluctuations (n=3, 20%), difficulties in using technology (n=2, 13%), and difficulty hearing instructions due to hearing loss (n=2, 13%). Patients had favorable perceived benefits of the program, with 14 (93%) considering it very useful for the current management of health and 1 (7%) moderately useful. No adverse events were reported, and all participants said that they were willing to continue the program and recommend it to others.

Conclusions: Our findings suggest that the online cognitive and motor program was well received, safe, and perceived to be of benefit to this group of medically stable people with PD in H&Y stages I-IV. Access to specialized care and enhancement of long-term adherence to regular exercise can be achieved with online community group programs.

(JMIR Aging 2022;5(4):e40325) doi:10.2196/40325

KEYWORDS
Parkinson disease; dual-task training; exercise; digital intervention; online intervention; physical activity; physical therapy; elder; older adult; geriatric; neurodegenerative; adherence; acceptability; community based; group program; online program; physiotherapy; cognitive training online exercise; Parkinson's; neuromuscular; task training; physiotherapist; motor; movement; cognitive; cognition; vocal; voice; speech
Introduction

Parkinson disease (PD) is considered 1 of the fastest-growing neurological disorders in the world [1]. It causes significant functional disabilities, affecting posture, gait, daily living activities, and cognition [2,3]. Impairments in frontal executive function and attention in PD are common and have been associated with loss of balance and an increased risk of falls [2,4-8].

There is growing evidence suggesting that nonpharmacological interventions, such as exercise/physiotherapy [2,9,10] and cognitive training [11-15], benefit people with PD in both physical and cognitive outcomes. Combining interventions may be a new potential treatment and comes in line with the growing evidence of the feasibility and potential benefits of dual tasking or multitasking in older adults [16,17] and in people with PD [14,18-20].

Initial research in this area of dual-task training concluded that 30 minutes once a week for 3 weeks of multiple-task gait training is feasible in 5 people with PD (Hoehn and Yahr [H&Y] stages I-III, mean age of 61 years with a mean of 8 years with PD) [21], with sustained benefits in multiple-task walking velocity and levels of fatigue and anxiety. More recently, a randomized clinical trial in this area with 121 individuals with early to mid-stage PD showed that dual-task gait improved when compared to a control period without training [19]. The study implemented 2 dual-task training programs, one with consecutive training and the other with concurrent (ie, integrated) dual-task training delivered in the home setting. Importantly, effects transferred to activities in daily life that were not trained and benefits were retained after a 12-week follow-up. This novel training program had excellent compliance from people with PD and did not increase the risk of falls.

Given the recent need for online solutions associated with the coronavirus pandemic and with the advances in technology facilitating access to specialized care [22-25], online exercise programs have emerged as a means for people with PD to stay physically and mentally active [26,27]. Online community-based dual-task exercise programs specifically adapted for people with PD have not been studied yet but may be an excellent tool to facilitate access to the reported benefits of dual-task training.

Here, we aim to assess the acceptability and safety of delivering such dual-task programs in an online group format with people with PD in the early to late stages of PD.

Methods

Design

Program acceptability was evaluated with mixed methods quantitative and qualitative assessments [28]. To ensure quality in the research report, we followed the Good Reporting of A Mixed Methods Study (GRAMMS) checklist [29].

Sampling and Recruitment

The sampling method selection was nonprobabilistic by convenience. Recruitment took place for 3 months from October 2021. All people with PD registered at the Portuguese Parkinson Disease Patient Association were invited to join the program online. People were included if they had (1) a diagnosis of PD (self-reported by patients), (2) H&Y stages I-IV, (3) the ability to connect online via Zoom (Zoom Video Communications, Inc) safely or have a care partner to assist if needed, (4) the ability to communicate with the investigator to understand and comply with the study procedures, and (5) the willingness and ability to provide written informed consent to participate and understand the right to withdraw their consent at any time without prejudice toward future medical care.

Participants were excluded if they had self-reported severe cognitive difficulties and significant active psychiatric problems that would incapacitate them from participating.

A therapist from the patient association clinic carried out the selection process and included participants that met the study criteria via email. The PD3 Move program was provided as an online community exercise program offered by the patient association.

Program

The program was led by a physiotherapist with expertise in PD and cognitive training, as well as with experience in building and implementing community dual- and multitask programs in PD. The program was delivered in 2 (1-hour) group sessions per week for 16 weeks. The program consisted of combining cognitive exercises projected on a dynamic Zoom background, with participants responding with physical and vocal exercises. The use of different types of Zoom backgrounds to support the cognitive challenge, as well as the use of voice to respond, was expected to be associated with higher levels of engagement, learning benefits, and exercise adherence.

Dual-task training was defined as the capacity to simultaneously perform 2 or more tasks during transfers, ambulation, and other movement-related activities [30]. The performance of these simultaneous 2 attention-demanding tasks had different goals, requiring patients to shift attention between tasks or placing equal amounts of attention on both tasks [31].

The cognitive exercises selected targeted the 4 main cognitive domains (examples in Table 1) particularly affected in PD [32]: attention (ability to apply different cognitive senses), working memory (temporarily storing and managing information), executive function (ability to manage time and attention, switch focus, plan and organize, remember details in sequence), and visual spatial skills (orientation in space, taking in and organizing visual information from the screen).
Table 1. Example of 1 exercise per cognitive domain.

<table>
<thead>
<tr>
<th>Cognitive domain targeted</th>
<th>Exercise instructions</th>
<th>Primary motor goal</th>
<th>Primary cognitive goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attention</td>
<td>“Start by stepping in place. Pay attention to random days of the week on the screen and say them out loud. Every time you see Monday, Thursday, and Sunday, lift your arms up.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increasing physical capacity, stepping, global amplitude, vocal volume, rhythm, and speed of motor responses</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Training sustained and divided attention</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive function</td>
<td>“Reorganize the activities you see in chronological order as you step in place. 1: I wash my hair; 2: I turn on the shower; 3: I brush my hair.” A video graphic of a 10-second timer is inserted with each new prompt.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increasing physical capacity and tolerance to dual-task interference of walking/stepping and thinking</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Training managing time, planning, and organizing a sequence of activities</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual spatial</td>
<td>“If you see a dog inside the house, sit down. If the dog is on the left, take a step to the left. If the dog is on the right, take a step to the right. If the dog is in the front, take a step forward.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increasing physical capacity, stepping, and tolerance to dual-task interference on quick stepping</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Training quick decision-making and working memory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working memory</td>
<td>“Pay attention to a shopping list of 5 items. Then, with a new background, perform side steps for 30 seconds and then say the names of all [items] on the list.”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Increasing physical capacity and stepping</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Retaining information temporarily</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Exercises were modified between sessions to maintain motivation and reduce memorization. Physical exercises consisted of an array of frequently recommended movements in PD that directly enhance functional activities of daily living, and relevant to PD, such as sitting and standing, reaching, and stepping or walking in place [2]. All physical exercises focused on high-amplitude, multidirectional movements, increasing in complexity and speed gradually and enough to foster motor learning and motivation but not so quickly as to cause frustration.

Based on the clinical judgment of the instructor and the patient’s performance, exercises were progressively increased through several levels of difficulty via (1) increasing the physical or vocal challenge or (2) manipulation of features on the cognitive exercises (number and type of prompts per time, their intrinsic complexity, or the interval between prompts). Gamification principles were also adapted to some exercises to act as motivational drivers for the participants and enhance social interaction [33]. An example of the organization of an exercise class is given in Table 2.

Table 2. An example of the outline of an exercise class.

<table>
<thead>
<tr>
<th>Phase</th>
<th>Example of a group exercise session</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase 1: arrival (10 minutes)</td>
<td>• Social interaction (eg, greeting friends)</td>
</tr>
<tr>
<td></td>
<td>• Brief discussion on safety issues and how to participate (first timers)</td>
</tr>
<tr>
<td></td>
<td>• Assessment of new health and logistical issues since the last session (second timers)</td>
</tr>
<tr>
<td>Phase 2: warm-up (10 minutes)</td>
<td>• Group warm-up using whole-body amplitude movements incorporating movements of the neck, shoulders, hands, trunk, hips, and knees, combined with walking in place or in a chair, in a rhythmic routine using appropriate music</td>
</tr>
<tr>
<td></td>
<td>• Voice warm-up exercises with loud ahh, glides, and humming exercises</td>
</tr>
<tr>
<td>Phase 3: exercise 1 (learning 5 minutes and training 10 minutes)</td>
<td>• Learning movements to respond to a cognitive exercise that will be projected in the background behind the instructor (ie, stand up and sit again every time you see a 5 on the screen; 5 minutes)</td>
</tr>
<tr>
<td></td>
<td>• Training exercise (10 minutes)</td>
</tr>
<tr>
<td></td>
<td>• Progress to more speed-based movement while maintaining amplitude and then to more cognitive demands (ie, lift arms if you see a 3 and count how many fives appear; 5 minutes)</td>
</tr>
<tr>
<td>Phase 4: exercise 2 (learning 5 minutes and training 10 minutes)</td>
<td>• Learning movements to respond to a cognitive exercise that will be projected (ie, identify the direction of the arrow and take a step in that direction; 5 minutes)</td>
</tr>
<tr>
<td></td>
<td>• Training (10 minutes)</td>
</tr>
<tr>
<td></td>
<td>• Progress to more speed-based movement while maintaining amplitude and then to more cognitive demands (ie, if the arrow is red, do the opposite move; 5 minutes)</td>
</tr>
<tr>
<td>Phase 5: fire down (5 minutes)</td>
<td>• Active, slow-amplitude movements with music, stretching, and breathing exercises</td>
</tr>
<tr>
<td>Phase 6: session assessment (5 minutes)</td>
<td>• Brief group discussion to gather participant feedback on each exercise to guide future sessions</td>
</tr>
</tbody>
</table>
Data Collection
A pretest was first applied to a group of 5 individuals with a diagnosis of PD. Researchers questioned these individuals to understand their perceptions about the online questionnaires. The instrument was considered sufficiently clear, objective, and comprehensive and did not present questions that could be ambiguous or equivocal. This procedure allowed researchers to determine that the survey was suitable for this study.

Before attending the program, all participants completed a questionnaire collecting general information about demographics, clinical problems, past medical conditions, current exercise habits, and perceived facilitators/barriers to exercise. Importantly, when patients reported falls in the assessment questionnaire, they were asked to take safety precautions, namely have a care partner present, always exercise in the sitting position, and signal to the therapist to monitor ongoing risk.

A postassessment anonymous online questionnaire was sent to participants immediately after the program. This questionnaire assessed patients' overall satisfaction, preferences, and barriers and facilitators to participating and any type of problem or adverse events that occurred during the program.

The overall satisfaction with the online program was assessed using a 5-point Likert scale (1=very satisfied, 2=satisfied, 3=neither satisfied nor unsatisfied, 4=unsatisfied, and 5=very unsatisfied). Patients were also asked whether they would recommend the program to a friend (1=yes, 2=no, and 3=maybe) and how likely they are to return to a similar program (1=very likely, 2=likely, 3=neither likely nor unlikely, 4=unlikely, and 5=very unlikely).

To assess the program design, patients were asked which part of the session they preferred: 1=warm-up with simple amplitude-based movements; 2=physical exercises on their own (eg, stepping, raising arms, sit stand); 3=cognitive exercises with movements (answering the cognitive challenges with movement and voice); 4=social interaction moments before, during, and after the session; 5=explanation of the teacher to Parkinson-related questions or reasons for the specific exercise; 6=final relaxation with soft music and breathing exercises; 7=preferred all the parts; and 8=did not prefer any parts.

The instructor assessed patient presence and feedback at the end of each session and prepared a monthly report regarding difficulties or learning points from the application of the program. Importantly, when patients reported falls in the assessment questionnaire, they were asked to take safety precautions, namely have a care partner present, always exercise in the sitting position, and signal to the therapist to monitor ongoing risk.

A postassessment anonymous online questionnaire was sent to participants immediately after the program. This questionnaire assessed patients' overall satisfaction, preferences, and barriers and facilitators to participating and any type of problem or adverse events that occurred during the program.

The overall satisfaction with the online program was assessed using a 5-point Likert scale (1=very satisfied, 2=satisfied, 3=neither satisfied nor unsatisfied, 4=unsatisfied, and 5=very unsatisfied). Patients were also asked whether they would recommend the program to a friend (1=yes, 2=no, and 3=maybe) and how likely they are to return to a similar program (1=very likely, 2=likely, 3=neither likely nor unlikely, 4=unlikely, and 5=very unlikely).

To assess the program design, patients were asked which part of the session they preferred: 1=warm-up with simple amplitude-based movements; 2=physical exercises on their own (eg, stepping, raising arms, sit stand); 3=cognitive exercises with movements (answering the cognitive challenges with movement and voice); 4=social interaction moments before, during, and after the session; 5=explanation of the teacher to Parkinson-related questions or reasons for the specific exercise; 6=final relaxation with soft music and breathing exercises; 7=preferred all the parts; and 8=did not prefer any parts.

The instructor assessed patient presence and feedback at the end of each session and prepared a monthly report regarding difficulties or learning points from the application of the program. In addition, the therapist made follow-up phone calls, when needed, to assess people’s feedback, concerns, or any questions. To ensure that the data collected by the instructor remained pertinent to the aims of the study, an interview guide was developed, comprising a set of open-ended questions focused on the difficulties in and facilitators to participating in PD3 Move. Examples of questions used in the guide are:

- Are there any factors regarding technology or other issues limiting your participation in the program?
- Do you have someone to help you during the classes?
- Do you feel the exercises are being delivered at a good speed for you?
- Do you take resting periods, when needed?
- What do you think would assist or facilitate you participation in the exercises?

Data Analysis
Descriptive statistics was performed using the IBM Statistic Package for the Social Sciences software (SPSS Statistics for Windows, version 27.0). Textual data from open-ended questions and the instructor’s notes were analyzed using the QDA Miner Lite database. The Braun, Clarke, Hayfield, and Terry (2019) process of content analysis was applied. This method of analysis focuses on identifying recurring common themes, ideas, and patterns of meaning in data. This method comprises 4 stages: preanalysis, encoding, categorization, and interpretation of data.

Ethical Considerations
Researchers followed the principles of the Declaration of Helsinki. The study protocol was approved by the Egas Moniz Research Ethics Board (ID: 948, date: March 25, 2021). Prior to starting the program, all participants received information regarding the study procedures and provided their written informed consent.

The online survey was set up so that participants were free to decide not to answer any question, change or review their responses, or voluntarily quit at any time. To comply with the ethical principles of anonymity and confidentiality, all data collected were free of any personally identifying information, including any form of electronic identifiers.

The archive of essential documents was carried out in a locked file cabinet at Egas Moniz University for 5 years. After this retention period, all data will be destroyed.

No individual data will be available.

Results
Participants
In total, 15 individuals with a diagnosis of PD participated in the program in 16 group sessions of 1 hour each performed twice a week. Participants were mainly female (n=9, 60%) with a mean age of 69.4 (SD 9.3) years and H&Y stages I-IV. The participants’ demographics and clinical characteristics can be found in Table 3.
Table 3. Participants’ general and clinical characteristics (N=15).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>69.4 (9.3)</td>
</tr>
<tr>
<td>Median</td>
<td>71.0</td>
</tr>
<tr>
<td>Minimum</td>
<td>48</td>
</tr>
<tr>
<td>Maximum</td>
<td>80</td>
</tr>
<tr>
<td><strong>Time since diagnosis (years)</strong></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>10.7 (7.1)</td>
</tr>
<tr>
<td>Median</td>
<td>10.0</td>
</tr>
<tr>
<td>Minimum</td>
<td>3</td>
</tr>
<tr>
<td>Maximum</td>
<td>31</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>9 (60)</td>
</tr>
<tr>
<td>Male</td>
<td>6 (40)</td>
</tr>
<tr>
<td><strong>Main problems, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>General fatigue</td>
<td>12 (80)</td>
</tr>
<tr>
<td>Urinary problems and constipation</td>
<td>10 (67)</td>
</tr>
<tr>
<td>Difficulties in daily activities</td>
<td>9 (60)</td>
</tr>
<tr>
<td>Difficulties in walking</td>
<td>9 (60)</td>
</tr>
<tr>
<td>Difficulties in balance and falls</td>
<td>7 (47)</td>
</tr>
<tr>
<td>Problems with sleep</td>
<td>7 (47)</td>
</tr>
<tr>
<td>Pain</td>
<td>7 (47)</td>
</tr>
<tr>
<td>Anxiety or apathy or depression</td>
<td>6 (40)</td>
</tr>
<tr>
<td>Difficulties in memory, thinking, and attention</td>
<td>6 (40)</td>
</tr>
<tr>
<td>Tremor</td>
<td>5 (33)</td>
</tr>
<tr>
<td>Difficulties in talking</td>
<td>4 (27)</td>
</tr>
<tr>
<td><strong>Fall history (past 3 months), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>No falls</td>
<td>6 (40)</td>
</tr>
<tr>
<td>1 fall</td>
<td>4 (27)</td>
</tr>
<tr>
<td>2 falls</td>
<td>2 (13)</td>
</tr>
<tr>
<td>Frequent falls</td>
<td>3 (20)</td>
</tr>
</tbody>
</table>

Regarding exercise habits, of the 15 participants, 8 (53%) were doing some form of exercise besides the online program. All participants considered that it is important to exercise regularly, having identified several factors that influence them to exercise, namely the benefits of exercise for their health (n=15, 100%), feeling exacerbation of their symptoms when they do not exercise regularly (n=9, 60%), having fun when exercising (n=6, 40%), and feeling guilty when not exercising (n=2, 13%).

**Program Attendance**

The attendance rate was high, with participants coming to more than 13 (81%) of the sessions. People did not attend some sessions, but none totally dropped out of the program. Reasons for not attending reported by participants varied but included medical appointments, health-related issues, and family events. Adherence with the exercises throughout the sessions was also perceived as high by the instructor, with participants continuously engaging in the exercises (assessed via general vocal responses to the exercises).

**Safety**

In the exit questionnaire, all participants reported having no major problems during the sessions. Several participants (n=8, 53%) had supervision from family or care partners during the sessions. Based on the instructor’s notes, all participants completed the program, with no report of severe problems during the sessions and some reporting only some fatigue.
Participants’ Satisfaction and Perceived Benefits

After the program, participants provided favorable feedback, with 8 (53%) feeling very satisfied and 7 (47%) feeling satisfied with the program. In addition, all participants reported their willingness to attend future online classes and said they would recommend it to another person with PD. Patients had favorable perceived benefits of the program (n=15, 100%) for the current management of exercise/health habits, with 14 (93%) considering it very useful and 1 (7%) moderately useful.

Feedback Regarding the Program Format and Delivery

The frequency of 2 sessions per week was considered ideal by 8 (53%) participants. The remaining participants expressed a preference to participate 3 (n=3, 20%) to 4 (n=4, 27%) times a week in this type of sessions. Regarding the type of physical and cognitive exercises, only 2 (13%) of the participants were familiar with this type of exercises performed. For the remaining 13 (87%) participants, the combination was something new.

Considering all activities carried out, 9 (60%) participants reported enjoying all these activities in general. All participants expressed a preference for 2 components: (1) performing cognitive exercises, responding verbally, and movements and (2) moments of explanation by the health professional regarding Parkinson-related issues or about the reason for the specific exercise. The less preferred components were (1) performing physical movements (eg, walking in place, sitting/rising, taking steps, raising arms; n=1, 7%) and (2) relaxing at the end with soft music (n=1, 7%).

The social moments that allowed the exchange of experiences and ideas was considered useful by 14 (93%) participants, with 1 (7%) participant expressing neutrality in relation to this component of the program.

Instructor’s Notes Regarding Program Delivery

Instructor notes highlighted that some time had to be allocated to the logistics of the program (ie, helping with technology difficulties, phone calls for follow-up, registering attendance), and this was critical for program success. However, the time requirements necessary decreased as the participants became more agile with the technology and program. Additionally, the instructor reported that the most stressful aspect was the need for constant modification of the exercises during the sessions. This included adjustments to the type of physical activities, the length of the exercise, the use of verbal feedback, the time for learning in a group setting, and resting periods suitable for all.

The instructor noted that alerts for the risk of falling and abnormalities in posture and performance, enabling health care providers to provide timely feedback so that online training will be a valid option for disease management of people with PD showing a 100% attendance rate [26]. However, the patient’s acceptability, preferences, and long-term adherence in online group settings is still unknown. Our findings suggest that the clinician-delivered dual-task motor cognitive program is acceptable, safe, and perceived to be of benefit to this group of medically stable people with PD in H&Y stages I-IV. Online programs will likely remain a key element of future delivery of care for people with PD, and our findings inform that the previously reported feasibility of dual-task training may be replicated in this format and fuel further development of such online community-based programs, resources, and research.

Attendance and satisfaction was high, with people attending 81% of the sessions. This was concordant with previous feasibility studies on online dance therapy programs for patients with PD showing a 100% attendance rate [26].

When compared to in-person programs, online programs can be a way to provide personalized and timely care to more people with PD. This may ultimately impact the change in care models as online therapies can complement in-person care as a means of ongoing and easy access to specialized care.

We believe that online training will be a valid option for disease management in the future as the development of key enabling technologies will allow health professionals to provide training guidance and monitor movement, behavior, and cognitive/motor learning. This state of play will allow real-time assessment of patients’ posture and performance, enabling health care professionals to provide timely feedback so that online training contributes more effectively to treating PD [34,35].
Several aspects of this program contributed to the high attendance and satisfaction. First, adherence may be related to highly motivated patients who have continuous encouragement to participate [36]. The phone calls from the therapist when they were absent was appreciated by participants and was a potential motivating factor to come to the next session and may have enhanced adherence [37]. In addition, the familiar relationship that the physiotherapist/instructor had with each of the participants also allowed the therapist to preferentially select the types of exercises that participants would enjoy, while maintaining a strong focus on tasks that would be a match for the physical capabilities of the various participants with PD. This proximity might also explain why all participants indicated that they were willing to continue in the program. Nevertheless, in line with previous studies [19] using dual-task activities, all participants were prepared to continue the same training if it were offered again. Second, the small group size is believed to have enhanced adherence via group social interaction with fellow participants. It allowed for good group dynamics, allowed visual assessment of major safety concerns related to balance limitations, and facilitated follow-up calls. Importantly, even though it is difficult to replicate the social aspects of group in-person classes in online formats, the constant engagement through voice in this program is believed to contribute to greater social interactions. Third, intrinsic motivation, via experiencing and recognizing the actual benefit of the exercises and enjoying participation are important factors for long-term training adherence [38,39]. Fourth, the therapist’s PD-specific and cognitive knowledge was recognized as an important factor to facilitate participation. Additionally, it allowed us to adequately anticipate and act quickly on problems that arose that may have influenced patients’ satisfaction [36,40]. Expertise allowed the physiotherapist to recognize safety issues and to anticipate abnormal postural behaviors that may arise when participants performed additional tasks. Fifth, the program used the participants’ continuous feedback to constantly adapt and develop new exercises. Importantly, participants were able to bypass common safety concerns and technological difficulties with support from care partners and the instructor. Even though 47% of the participants had reported initially having difficulties with balance and falls and 60% had walking difficulties, no one experienced major problems. This was also in concordance with previous feasibility studies on online programs for PD showing no adverse events [26].

Important insights regarding the effects of combining dual modes of exercise in PD were identified. The physiotherapist/instructor reported that the most stressful problem/factor was the need for constant modification of the exercises during the sessions. Importantly, this constant modification may represent a barrier to replicating programs easily, and further detailed description and research on the exercises delivered should be carried out and shared with others in educational courses [41].

Limitations

Our study was not without limitations mainly intrinsic to its acceptability nature, small sample size, and single-center design, which imposed restrictions on the generalizability of the findings. Second, although primary data were collected through an anonymous online questionnaire, we should not exclude that the participants’ actual reports may diverge from what they revealed due to biases, such as a lack of confidence in guaranteeing anonymity or protecting identity values or beliefs. In addition, we also must consider that if participants perceive it to be socially desirable, they might overstate the frequency of positive items. Third, we included a fairly heterogeneous group of people with PD from different backgrounds and did not specifically include people with PD with cognitive impairment. Even though it remains to be determined how such people can undergo dual-task training, given the potential benefits of this type of program on cognition due to its cognitively demanding components (memorizing the instructions, quick decision-making to reply, cognitive or dual tasking when dividing attention between physical and cognitive exercises, and dealing with environment with constant changes), these programs may be of particular interest in treating cognitive dysfunction in individuals who already display mild cognitive impairment (to potentially delay or slow down further decline). Future programs should attempt to include individuals with cognitive limitations that reflect the type of frontal lobe deficits that are more impaired in fallers [42]. Importantly, a growing body of evidence suggests that there is an increased risk of falls in the presence of cognitive impairment [5,43,44] as well as in dementia [45,46], and this trend is present in both community-dwelling and institutionalized older populations. Hence, interventions that can potentially improve executive function and cognitive processes, in particular attention, have been recognized as a significant element in the process of treating balance and gait deficits in people with PD [18].

Even so, the use of dual-task training may already be of particular interest in preventing or delaying cognitive dysfunction in people with PD who are not (yet) affected. Additionally, people with PD in early disease stages are more amenable to these types of innovative therapeutic interventions. Any contribution to reduce the obvious pitfalls in such dual motor cognitive interventions will be a valuable adjunct to reduce the evolving needs of people with PD and their care partners. An additional limitation in group settings is the impossibility of effectively monitoring patients during the sessions and registering potential situations of wearing off, sudden offs, patients in off (in other words, patients without the effect of medication), etc. Reaction to the exercises and safety was the primary assessment being conducted during the sessions by the therapist. The type of medication taken by participants is another important issue as the dose and type of pharmacological treatments will highly impact patients’ overall functionality.

Conclusion

Dual-task training is an emerging field for PD, but access to such specialized care remains limited. Online community-based dual-task exercise programs, such as the PD3 Move program, specifically adapted for people with PD may be an excellent tool to facilitate access to previously reported benefits of dual-task training. It can provide a safe and enjoyable way to reduce physical and cognitive inactivity commonly seen in PD.
Yet, the design, type of visual display, type of sessions, and participants included will need further reflection. Sharing the concept of such a program implementation may fuel the development of future research and similar community exercise care services for PD that incorporate the complexity of the cognitive challenges in PD. This ultimately may lead to—at least partly—initial treatment suggestions for those who decide to start such an online program for PD and be useful to guarantee safety and better care to the target population.

Acknowledgments
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. This publication was financed by national funds through the Foundation for Science and Technology (FCT), IP, under the project UIDB/04585/2020. The researchers would like to thank the Centro de Investigação Interdisciplinar Egas Moniz (CiiEM) for the support provided for the publication of this paper and all the participants.

Data Availability
The data presented in this study are available on request from the corresponding author.

Authors' Contributions
JD (Josefa Domingos) performed conceptualization, formal analysis, investigation, methodology, writing, and project administration. JD (John Dean) performed conceptualization, methodology, and writing and reviewing. JBF conducted formal analysis, methodology, and writing and reviewing. CG performed data curation, formal analysis, methodology, writing and reviewing, editing, and supervision. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References


Abbreviations

H&Y: Hoehn and Yahr
PD: Parkinson disease

Edited by T Leung; submitted 15.06.22; peer-reviewed by S Pandey, L Krejcová, C Baixinho; comments to author 04.10.22; revised version received 16.10.22; accepted 25.10.22; published 22.12.22.

Please cite as:

Domingos J, Dean J, Fernandes JB, Godinho C. An Online Dual-Task Cognitive and Motor Exercise Program for Individuals With Parkinson Disease (PD3 Move Program): Acceptability Study. JMIR Aging 2022;5(4):e40325
URL: https://aging.jmir.org/2022/4/e40325
PMID: 34501947
Evidence and User Considerations of Home Health Monitoring for Older Adults: Scoping Review

Andrew Chan1,2, BSc, MD, PhD; Rachel Cohen3,4, BASc, MASc; Katherine-Marie Robinson5,6, BSc, BASc, MSc; Devvrat Bhardwaj7, BSc, MASc; Geoffrey Gregson1,2, BKin, MSc, MBA, LLM, PhD; Jeffrey W Jutai8,9, BSc, MSc, PhD; Jason Millar5,6, BScE, BA, MA, PhD; Adriana Ríos Rincón1,2, BSc, MSc, PhD; Atena Roshan Fekr3,4, BSc, MSc, PhD

1Faculty of Rehabilitation Medicine, Department of Occupational Therapy, University of Alberta, Edmonton, AB, Canada
2Innovation and Technology Hub, Glenrose Rehabilitation Research, Edmonton, AB, Canada
3KITE Research Institute, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
4Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada
5School of Engineering Design and Teaching Innovation, Faculty of Engineering, University of Ottawa, Ottawa, ON, Canada
6Department of Philosophy, Faculty of Arts, University of Ottawa, Ottawa, ON, Canada
7Department of Electrical Engineering and Computer Science, Faculty of Engineering, University of Ottawa, Ottawa, ON, Canada
8Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada
9LIFE Research Institute, University of Ottawa, Ottawa, ON, Canada

Corresponding Author:
Rachel Cohen, BASc, MASc
KITE Research Institute
Toronto Rehabilitation Institute
University Health Network
550 University Avenue
Toronto, ON, M5G 2A2
Canada
Phone: 1 416 597 3422 ext 7622
Email: rachelj.cohen@mail.utoronto.ca

Abstract

**Background:** Home health monitoring shows promise in improving health outcomes; however, navigating the literature remains challenging given the breadth of evidence. There is a need to summarize the effectiveness of monitoring across health domains and identify gaps in the literature. In addition, ethical and user-centered frameworks are important to maximize the acceptability of health monitoring technologies.

**Objective:** This review aimed to summarize the clinical evidence on home-based health monitoring through a scoping review and outline ethical and user concerns and discuss the challenges of the current user-oriented conceptual frameworks.

**Methods:** A total of 2 literature reviews were conducted. We conducted a scoping review of systematic reviews in Scopus, MEDLINE, Embase, and CINAHL in July 2021. We included reviews examining the effectiveness of home-based health monitoring in older adults. The exclusion criteria included reviews with no clinical outcomes and lack of monitoring interventions (mobile health, telephone, video interventions, virtual reality, and robots). We conducted a quality assessment using the Assessment of Multiple Systematic Reviews (AMSTAR-2). We organized the outcomes by disease and summarized the type of outcomes as positive, inconclusive, or negative. Second, we conducted a literature review including both systematic reviews and original articles to identify ethical concerns and user-centered frameworks for smart home technology. The search was halted after saturation of the basic themes presented.

**Results:** The scoping review found 822 systematic reviews, of which 94 (11%) were included and of those, 23 (24%) were of medium or high quality. Of these 23 studies, monitoring for heart failure or chronic obstructive pulmonary disease reduced exacerbations (4/7, 57%) and hospitalizations (5/6, 83%); improved hemoglobin A1c (1/2, 50%); improved safety for older adults at home and detected changing cognitive status (2/3, 66%) reviews; and improved physical activity, motor control in stroke, and pain in arthritis in (3/3, 100%) rehabilitation studies. The second literature review on ethics and user-centered frameworks found...
19 papers focused on ethical concerns, with privacy (12/19, 63%), autonomy (12/19, 63%), and control (10/19, 53%) being the most common. An additional 7 user-centered frameworks were studied.

Conclusions: Home health monitoring can improve health outcomes in heart failure, chronic obstructive pulmonary disease, and diabetes and increase physical activity, although review quality and consistency were limited. Long-term generalized monitoring has the least amount of evidence and requires further study. The concept of trade-offs between technology usefulness and acceptability is critical to consider, as older adults have a hierarchy of concerns. Implementing user-oriented frameworks can allow long-term and larger studies to be conducted to improve the evidence base for monitoring and increase the receptiveness of clinicians, policy makers, and end users.

JMIR Aging 2022;5(4):e40079 doi:10.2196/40079

KEYWORDS
smart homes; concerns; user-centered frameworks; clinical evidence; home health monitoring; gerontechnology; telemonitoring; older adults

Introduction

Background

Current health care systems are being pushed to use the capabilities of modern technology outside the hospital to increase efficiency and effectiveness of health delivery [1]. Transforming care processes by using digital platforms and remote monitoring may help address our increasingly older population and higher life expectancies [2]. Smart home and health monitoring technologies have been touted as the future for managing chronic diseases and allowing people to age in place [3-6].

With the technological advancements in the Internet of Things and the widespread use of machine learning and artificial intelligence, the application of smart home technology for aging in place has become more realistic and feasible. Numerous studies on technology development [3,7-10], clinical applicability [11-14], and user considerations [15-17] have been conducted to demonstrate that the technology is ready for mainstream use. There is a plethora of clinically evaluated activity and health recording devices readily available in the market, including wearables (eg, wrist bands, chest bands, and textiles) and ambient sensors (eg, motion sensors, cameras, and pressure sensors) [18-21]. However, the widespread adoption and development of health monitoring platforms remain limited for 2 reasons.

First, the evidence remains siloed within disease-specific reviews or secondary prevention, whether in heart failure [22,23], chronic obstructive pulmonary disease (COPD) [24], diabetes [25], or cardiometabolic health [26], making it difficult to compare effective monitoring models, delineate the overall evidence for home monitoring, and identify where gaps remain [27-29]. There is also a need to differentiate automated monitoring from user-based monitoring, which involves patients texting or phoning in their results.

Second, user and ethical concerns related to monitoring technologies remain major barriers to user adoption, particularly in research [30-33]. Much of the smart home research is focused on the technical development of devices rather than the reliability and usability of smart home systems [34]. Although the feasibility of smart home technology is high in most studies, the acceptability is a critical factor [35]. The benefits of the technology are touted without considering if, how, and at what cost a user may be willing to integrate the technology in their lives [17,36]. In addition, concerns related to data privacy and control, autonomy, and social connectivity are sometimes neglected when designing such systems [37-39].

Objective

This review aimed to map out the literature on two major research questions: (1) what is the evidence for the effectiveness of home-based patient monitoring technologies for improving the health and well-being of older adults and (2) what are the ethical concerns that older adults have with home-based patient monitoring technologies, and what frameworks have been proposed to address these concerns? By addressing both questions, we aimed to provide a tool for researchers in this field to understand what needs to be studied and how to study them while keeping in mind ethical and user-centered practices.

In this study, we defined home health monitoring as the use of technology, omitting telecommunications, to monitor the health of users over time (ie, remotely) [40]. This could include using a variety of technologies, including wearables and ambient sensors, to track physiological parameters, activity levels, and routines or to facilitate rehabilitation and treatment [40].

The first part of this review outlined the clinical evidence for using home health monitoring through a scoping review of systematic reviews on home monitoring interventions. We highlighted the domains being researched; performed quality assessments; and determined whether the evidence is positive, neutral, or negative for home monitoring.

The second part of this review considered ethical concerns when researching and developing smart home monitoring technology and user-centered frameworks that address issues of acceptance and adoption. We outlined the ethical and user-centered frameworks that are available to improve these trials and described the current level of adoption of these frameworks in smart home monitoring.

Methods

To answer the first research question, we conducted a scoping review by searching for systematic reviews following the PRISMA (Preferred Reporting Items for Systematic Reviews
and Meta-Analyses) methodology [41,42]. The PRISMA checklist can be found in Multimedia Appendix 1.

**Data Sources and Search Strategy**

Two researchers (AC and RC) conducted a probing search on MEDLINE for studies using smart homes related to older adults. We decided upon a summary of search terms with consultation with an interdisciplinary team of clinicians and engineers. We completed a systematic scoping search on Scopus, MEDLINE, Embase, and CINAHL in July 2021. The search focused on systematic reviews and meta-analyses using smart homes and remote monitoring of older adults (Table S1 in Multimedia Appendix 2).

**Study Screening and Inclusion and Exclusion Criteria**

Three researchers (AC, RC, and KR) completed abstract screening and full-text screening. Inclusion criteria included systematic reviews written in English with >80% articles focused on older participants (aged >65 years) and published from 2010 to 2021. We excluded studies that focused on assistive technology, mobile health interventions, telephone- or videoconferencing-based interventions, mobile phones and apps, virtual reality, and robots because we focused on automated sensing technologies with clinician intervention. We also excluded studies in which >80% of the articles involved users texting or phoning in the results, rather than using automated monitoring systems. Narrative reviews and technical articles outlining the implementation of these technologies were excluded. We excluded studies that were not journal articles because our focus was on identifying gaps rather than estimating effect sizes.

**Data Extraction and Quality Assessment**

We created an extraction table and completed extractions independently. The extracted items included population characteristics (number of articles, included diseases or disorders, and percentage of articles including older adults), study type (meta-analysis and quantitative or qualitative study inclusion), and monitoring methods (automated monitoring, mixed automated monitoring, and user-reported monitoring).

We extracted outcome measures, including physiological outcomes (vital signs and blood tests), symptoms or health events (falls, exacerbations, or mortality), health care use, cognitive decline, functional status, adherence to rehabilitation, and activity levels. Studies on human factors included user perspectives on the ethics of home monitoring, the acceptability and usability of devices, and changes in quality of life (QoL) or social connections through technology. We coded outcomes as positive, negative, mixed, or “not enough evidence” according to the author’s assessment and summarized in tables. Finally, we extracted the challenges that the author listed on health monitoring research, ranging from privacy and security concerns and technology acceptability to technical challenges and lack of clinical evidence. The challenges were coded based on what the authors reported. We developed a spreadsheet with a separate column for each challenge that each author would raise. If a new challenge was listed, it was added to a separate column. After we added all the articles and challenges, the 3 authors reduced the number of challenges by grouping together similar challenges by consensus.

Four authors (AC, RC, KR, and DB) conducted a quality appraisal using the AMSTAR-2 checklist and appraised independently after deciding upon critical, important, and unimportant categories. Categories rated as “Critical” included a comprehensive literature search strategy, describing included studies in adequate detail, and accounting for the risk of bias on evidence synthesis. Categories rated as “unimportant” in quality assessment included explaining the selection of study designs, providing a list of excluded studies, and reporting sources of funding for studies included in the review. All the other items were considered important. Studies with ≥2 “No” ratings on critical categories were considered critically low quality, whereas studies with one “No” rating were considered low quality. Moderate-quality studies had ≥2 “important” categories scored as “No,” whereas high-quality studies had <2 “important” categories scored as “No.” The entire list can be found in Multimedia Appendix 3.

**Ethics and Stakeholder Participation Review**

To answer the second research question, we conducted 2 searches. The first search focused on the ethical challenges and concerns related to smart home technology. We conducted a systematic search from April to May 2021 using Scopus, Web of Science, and Dimensions AI. The search terms are listed in Table S2 in Multimedia Appendix 2. Once we collected all the papers, we screened the titles and abstracts to select the papers to be used for this review. Studies published before 2015 were excluded.

The second search focused on studies involving stakeholder participation while developing health care services and apps. We used Google Scholar to gather related chapters, journals, and articles with keywords (Multimedia Appendix 2). We excluded studies with health care frameworks that did not focus on user centeredness (ie, stakeholder involvement).

**Results**

**Study Characteristics for Question 1: Evidence for Smart Home Technologies**

The search yielded 1022 articles, which was reduced to 822 after deduplication (Figure 1). Screening abstracts yielded 480 articles, whereas full-text and additional screening during extractions yielded 94 systematic reviews. Most articles were excluded because they did not focus on telemonitoring technologies (325/728, 44.6%), followed by articles that did not focus on older adult users (120/728, 16.5%) and that were not systematic reviews (104/728, 14.3%). The extraction procedures are presented in Multimedia Appendix 4.
Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram outlining the extraction process. mHealth: mobile health.

Figure 2 shows an increasing trend for systematic reviews on home-based patient monitoring technologies, particularly starting in 2018. Live and automonitoring has only become more commonly reviewed since 2018. The number of systematic reviews showed an upward trend starting in 2018.

Figure 3 displays the study designs used in each systematic review, organized according to the outcomes discussed. Many articles have presented >1 outcome category. Explanations of the outcomes are described in the methodology.

Meta-analyses were most common when reporting physiological evidence, health events, and health use (9/18, 50.0% for physiological; 18/31, 58.1% for health events; and 20/30, 66.7% for health use). Studies focusing on cognition, safety, and activities of daily living (ADLs) were mostly mixed methods or nonpatient studies (13/18, 72.2%), whereas studies focusing on exercise were mostly meta-analyses or quantitative clinical studies (10/15, 66.7%). Most systematic reviews on ethics, acceptability, and usability used mixed methods or nonpatient studies (4/5, 80% for ethics; 20/32, 62.5% for acceptability; and 19/24, 79.2% for usability), and no single study design dominated QoL or social-focused studies.

Figure 4 displays the disease processes studied according to the category of disease monitoring. Acute prevention studies focused on reducing heart failure and COPD exacerbations and fall prevention (24/61, 39.3% reviews). Chronic management studies focused on blood pressure reduction, blood glucose control, and metabolic disease management (13/61, 21.3% reviews). Home monitoring studies focused on monitoring aging and status of patients with dementia (13/61, 21.3% reviews). Studies on physical activity monitoring focused on rehabilitation or increasing physical activity in the older population (11/61, 18.0% reviews).

The challenges related to current evidence and implementation of home monitoring were also analyzed. Many studies listed multiple challenges, whereas 9.3% (9/94) studies listed no challenge. The most common challenges in the literature included a lack of strong clinical evidence for monitoring (510/94, 53.1%), poor descriptions of methodologies of how patients were monitored (279/94, 28.7%), and applicability to broader patient populations (10/94, 10.6%). On the implementation side related to human factors, the acceptance of technology (33/94, 35.1%), usability of devices (18/94, 19.1%), privacy concerns (17/94, 18%), cost-effectiveness (23/94, 24.2%), and safety concerns with devices (7/94, 7.6%)
were listed. Technical challenges were not as commonly reported, although concerns with accuracy (10/94, 10.7%) and connectivity of devices (9/94, 9.7%) were more common than the others.

**Figure 5** shows the AMSTAR-2 quality assessments organized according to the monitoring category. Most studies were either of critically low or low quality across all systematic reviews (36/94, 38.3%, and 35/94, 37.2%, respectively). Most studies did not discuss the impact of risk of bias on results (57/94, 60.6%), did not discuss heterogeneity (55/94, 58.5%), or did not include an explicit statement on following a protocol (50/94, 53.2%).

**Figure 2.** Number of systematic review articles published in each year which included only live or automated monitoring, or had mixed modes of monitoring.

**Figure 3.** Number of articles according to outcome measures versus study design. ADL: activities of daily living; QoL: quality of life; RCT: randomized controlled trial.

- Meta-analysis, RCT only
- Meta-analysis
- RCT only
- Quantitative clinical
- Mixed methods
- Nonpatient studies
Summary of Systematic Reviews on Evidence for Remote Monitoring

Table 1 reports the outcomes from articles that focused on clinical outcomes, as shown in Figure 4. The disease-specific details are presented in Multimedia Appendix 5. In total, 51 of 64 (80%) or more systematic reviews reported positive results across all categories. Chronic disease management and physical activity were among the categories with the largest number of systematic reviews with positive evidence. Evidence was most limited in managing degenerative diseases, health events, and use in rehabilitation interventions.
Table 1. Summary of evidence for home health monitoring according to category and outcome.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Reviews</th>
<th>Positive evidence</th>
<th>Negative or inconclusive evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acute prevention</strong></td>
<td>22/24, (91%)</td>
<td>17/22 (77%) studies reported reduced exacerbations or mortality [43-57]</td>
<td>5/22 (23%) studies showed no difference in mortality [58-62]</td>
</tr>
<tr>
<td>Health event (exacerbations, mortality, and falls)</td>
<td>5/22 (23%) studies showed no difference in mortality [58-62]</td>
<td>17/22 (77%) studies reported reduced exacerbations or mortality [43-57]</td>
<td>2/22 (9%) studies showed inconclusive results on mortality and exacerbations [63,64]</td>
</tr>
<tr>
<td></td>
<td>2/22 (9%) studies showed inconclusive results on mortality and exacerbations [63,64]</td>
<td>High- or medium- quality: 4/22 (18%) studies reported reduced exacerbations [43,46,47,56]</td>
<td>High- or medium- quality: 3/22 (14%) studies showed no difference in mortality [58,59,61]</td>
</tr>
<tr>
<td>Health use (hospitalizations and ER visits)</td>
<td>21/24, (88%)</td>
<td>16/21 (76%) studies reported reduced hospitalizations and ER visits [47-56,58-61,64,66]</td>
<td>2/21 (10%) studies showed no change in hospitalization [43,57]</td>
</tr>
<tr>
<td></td>
<td>2/21 (10%) studies showed no change in hospitalization [43,57]</td>
<td>High- or medium- quality: 5/21 (24%) studies reported reduced hospitalizations [45,47,56,61]</td>
<td>2/21 (10%) studies showed no change in hospitalization [43,57]</td>
</tr>
<tr>
<td>Chronic disease management</td>
<td>10/14, (71%)</td>
<td>8/10 (80%) studies showed improved HbA1c [69-76]</td>
<td>1/10 (10%) study showed mixed evidence for change in HbA1c [77]</td>
</tr>
<tr>
<td>Physiological outcomes (blood pressure, HbA1c, and blood lipids)</td>
<td>5/10 (50%) showed improved blood pressure [69-71,73,77]</td>
<td>5/10 (50%) showed improved blood pressure [69-71,73,77]</td>
<td>3/10 (30%) studies showed mixed evidence for blood pressure reduction [72,76,78]</td>
</tr>
<tr>
<td></td>
<td>2/10 (20%) showed blood lipid reduction [69,76]</td>
<td>2/10 (20%) showed blood lipid reduction [69,76]</td>
<td>1/10 (10%) study found no change in blood lipids [72]</td>
</tr>
<tr>
<td></td>
<td>2/10 (20%) showed blood lipid reduction [69,76]</td>
<td>High- or medium- quality: 1/10 (10%) study showed improved HbA1c [75]</td>
<td>High- or medium- quality: 1/10 (10%) study showed mixed evidence for change in HbA1c [77]</td>
</tr>
<tr>
<td>Health events (mortality, medical events, and pain)</td>
<td>4/14 (28%)</td>
<td>2/4 (50%) studies showed reduced mortality or adverse health events [69,70]</td>
<td>1/4 (25%) study neutral for mortality for patients with chronic kidney disease [78]</td>
</tr>
<tr>
<td></td>
<td>No high- or medium-quality studies</td>
<td>No high- or medium-quality studies</td>
<td>1/4 (25%) study found medication adherence was mixed for varying medical conditions [79]</td>
</tr>
<tr>
<td>Health use for chronic disease (hospitalizations and ER visits)</td>
<td>5/14 (36%)</td>
<td>4/5 (80%) studies showed reduced admissions and ER visits [69,70,80] and treatment adherence [81]</td>
<td>1/5 (20%) study neutral for hospitalizations [80]</td>
</tr>
<tr>
<td></td>
<td>No high- or medium-quality studies</td>
<td>No high- or medium-quality studies</td>
<td>No high- or medium-quality studies</td>
</tr>
<tr>
<td>Degenerative disease monitoring</td>
<td>8/13 (61%)</td>
<td>3/8 (37%) studies showing that ADLs can be detected [82-84]</td>
<td>2/8 (25%) studies presented that technology is not mature enough to detect functional independence/ADLs [88,89]</td>
</tr>
<tr>
<td>Function and independence</td>
<td>2/8 (25%) studies showing that QoL related to independence improved [85,86]</td>
<td>2/8 (25%) studies showing that QoL related to independence improved [85,86]</td>
<td>High- or medium- quality: 1/8 (13%) study showing improved safety with assistive technology [87]</td>
</tr>
<tr>
<td></td>
<td>High- or medium- quality: 1/8 (13%) study showing improved safety with assistive technology [87]</td>
<td>2/8 (25%) studies presented that technology is not mature enough to detect functional independence/ADLs [88,89]</td>
<td>High- or medium- quality: 1/5 (20%) study presented high evidence for monitoring cognitive status and mental health [14]</td>
</tr>
<tr>
<td>Cognitive status</td>
<td>5/13 (38%)</td>
<td>2/5 (40%) studies presented weak evidence for detecting cognitive impairment or agitation [90,91]</td>
<td>2/5 (40%) studies presented that technology is not mature enough to detect cognitive status [92,93]</td>
</tr>
<tr>
<td></td>
<td>High- or medium- quality studies: 1/5 (20%) study found high evidence for monitoring cognitive status and mental health [14]</td>
<td>High- or medium- quality studies: 1/5 (20%) study presented that overall technology readiness is low [14]</td>
<td></td>
</tr>
</tbody>
</table>
The first theme involved detecting acute events including exacerbations of heart failure, COPD, or falls (Table 1 and Table S1 in Multimedia Appendix 5). Most reviews reported on exacerbations of heart failure and COPD. Monitoring reduced hospitalizations for patients with COPD (5/7, 71.4% studies) and for patients with heart failure (10/14, 71.4% studies). Mortality was unchanged for COPD in a few reviews (3/7, 42.8%) but for most reviews for heart failure (11/14, 78.6%). Only 1 study had reviewed fall interventions in an older population [46]. No clinical studies have focused on atrial fibrillation, although devices showed high specificity and sensitivity in detecting atrial fibrillation.

Management of chronic diseases included managing diabetes, blood pressure, kidney function, or multiple diseases simultaneously (Table 1 and Table S2 in Multimedia Appendix 5). In most studies, home health monitoring helped to reduce hemoglobin A1c (8/10, 80%) and blood pressure (5/9, 55.6%), although the results were more mixed when only considering high- or medium-quality studies. In total, 2/4 (50%) of studies reported reduced mortality, reduced health events, and fewer hospitalizations when monitoring chronic cardiovascular diseases.

Smart home monitoring for ensuring safety of older adults with dementia or cognitive impairment had scarce evidence (Table 1 and Table S3 in Multimedia Appendix 5). Only 1/2 (50%) of studies found no reduction in admission to care homes for monitored patients [87]. In addition, 2/5 (40%) of studies noted progress in the ability to detect cognitive decline [90] and aggression [91]. Most studies found that technologies were not mature enough to detect activity changes or improve independence [88,89,92,105]. For monitoring older adults in general, 3/N studies noted weak evidence for detecting changes in ADLs [82-84], and 2 studies showed improved QoL [85,86].

The fourth theme, rehabilitation adherence and encouraging an active lifestyle at home, included 11 studies (Table 1 and Table S4 in Multimedia Appendix 5). All studies showed positive outcomes for both disease-specific rehabilitation programs and for monitoring older adults in general. Monitoring improved adherence to cardiac rehabilitation, increased activity levels in patients with COPD, and improved motor control in patients after stroke [94-96]. Monitoring increased physical activity in older adults in general in 4 studies, and telerehabilitation with monitoring was found to be as effective as traditional rehabilitation [104]. Wearables helped to increase activity in patients with cancer, improve functioning in patients with arthritis, and improve QoL in postoperative patients [97-99].

### Summary of Systematic Reviews on Human Factors

As part of the scoping review, we included studies that focused on human factors such as acceptability of technology, ethical considerations, and costs. Although this was not the focus of question 1, we performed a basic analysis of the outcomes from these studies, recognizing the importance of human factors related to remote monitoring. Table 2 summarizes these 33 studies.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Reviews</th>
<th>Positive evidence</th>
<th>Negative or inconclusive evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health use</td>
<td>2/13 (15%)</td>
<td>• 1/2 (50%) study found that monitoring cognitive status and mental health lowered hospital visits [14]</td>
<td>• No reduced admission to care homes in 1/2 (50%) study [87]</td>
</tr>
<tr>
<td>Physical activity</td>
<td></td>
<td>• 10/11 (90%) studies showed improved physical activity or adherence to rehabilitation [94-103]</td>
<td></td>
</tr>
<tr>
<td>Rehabilitation adherence or physical activity increase</td>
<td>11/11 (100%)</td>
<td>• 1/11 (9%) study showed no change in rehabilitation outcomes compared with traditional rehabilitation [104]</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• High- or medium-quality studies: improved physical activity in older adults, improved motor control in stroke, and improved pain in patients with arthritis [96,98,101]</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**

aER: emergency room.
bCOPD: chronic obstructive pulmonary disease.
cHbA1c: glycated hemoglobin.
dADLs: activities of daily living.
eQoL: quality of life.
fNot available.
Table 2. Summary of studies focused on human factors (N=33).

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Reviews, n (%)</th>
<th>Positive evidence</th>
<th>Negative or inconclusive evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability for managing chronic diseases</td>
<td>9 (27)</td>
<td>5 studies showing high acceptability of monitoring technology [106-110]</td>
<td>• 2 studies were inconclusive on monitoring acceptability [111,112]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 2 studies were descriptive studies [17,113]</td>
</tr>
<tr>
<td>Acceptability of telerehabilitation</td>
<td>1 (3)</td>
<td>_a</td>
<td>• One study was inconclusive on acceptability of monitoring in telerehabilitation [114]</td>
</tr>
<tr>
<td>Acceptability for home health monitoring</td>
<td>14 (42)</td>
<td>3 studies showing good acceptability of monitoring technologies [115-117]</td>
<td>• 2 studies showed inconclusive results on acceptability [105,118]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• 9 studies describing acceptability [36,119-126]</td>
</tr>
<tr>
<td>Costs</td>
<td>3 (9)</td>
<td>2 studies showed weak evidence that remote patient monitoring is cost-effective [127,128]</td>
<td>• One study was inconclusive on cost-effectiveness of monitoring [129]</td>
</tr>
<tr>
<td>Ethics</td>
<td>3 (9)</td>
<td>_b</td>
<td>• 3 descriptive studies on ethical frameworks for remote monitoring [37,39,130]</td>
</tr>
<tr>
<td>QoL</td>
<td>3 (9)</td>
<td>1 study showed improved QoL with monitoring, though not by validated measures [32]</td>
<td>• 2 descriptive studies on QoL and social interaction with monitoring [131,132]</td>
</tr>
</tbody>
</table>

aNot available.
bQoL: quality of life.

Of the 24 studies on acceptability, 8/24 (33%) showed good acceptability of monitoring technologies in general. For chronic diseases, 5/9 (56%) showed good acceptability in monitoring for heart failure, COPD, and management of amyotrophic lateral sclerosis, whereas 2/9 (22%) studies on heart failure and COPD were inconclusive. Monitoring acceptability was less clear for generalized home health monitoring and rehabilitation for older adult patients with fatigue. Costs, ethical considerations, and QoL were inconclusive or were descriptive studies.

Although there appear to be positive results from specific diseases, there is a need to continue studying whether remote health monitoring is acceptable for the older adult population. Human factor considerations have been prominent in several studies and deserve closer inspection. Applying appropriate frameworks to the design and development of home health monitoring technology that address human factors such as users’ ethical concerns, like privacy and usability, need to be explored to ensure that users feel at ease while using and installing devices that monitor their lives. Therefore, the second part of our review focused on the ethical concerns that older adults have regarding home health monitoring technology and frameworks that have been suggested to address such concerns.

Study Characteristics for the Ethics and User-Centered Frameworks

The second research question in this study was “what frameworks have been proposed to address the ethical concerns that stakeholders, specifically older adults, have toward home health monitoring?” To learn which frameworks are available to address the ethical issues that arise with home health monitoring, the ethical issues themselves must be uncovered and discussed in more detail.

To explore the ethical issues, we conducted a literature review of ethical challenges. An initial search of 132 papers published since 2015 was conducted, focusing on the ethical challenges and concerns regarding smart home technology. After screening titles and abstracts, 19 papers discussing the ethical challenges and concerns regarding smart home technology were identified. The most discussed concerns are listed in Table 3. The papers were a mix of systematic reviews, literature reviews, qualitative research (including focus groups and interviews), and mixed methods research (qualitative and quantitative research, many using surveys to obtain results).
Ethical Challenges and Concerns

Overview

The first half of this review made it apparent that home monitoring of older adults is a useful intervention. However, many studies have highlighted the ethical challenges and impacts on user acceptance and adoption that arise when using monitoring technology.

When home monitoring technology for aging in place is used appropriately, it can improve QoL, maintain health and wellness of older adults, and support other stakeholders [37]. Such technology may allow increased autonomy and independence in older adults while providing additional support for family members or health care professionals [133]. However, stakeholders and researchers have raised many ethical concerns regarding the design, development, and deployment of home monitoring technologies. This section expands on the most discussed ethical questions and concerns regarding home health monitoring technology, as outlined in Table 3. By understanding these concerns, solutions may be discovered to better design and implement home monitoring technology for older adults and other critical stakeholders [37].

Privacy

Privacy is one of the most critical factors affecting older adults’ willingness to participate in and use smart home technology [31,33,37,133-138,143]. Privacy can be classified into two types: physical and informational. Physical privacy relates to the degree to which a person or their personal space is physically accessible [136]. The home is a refuge for privacy and intimacy [139], so it is understandable that some users of home monitoring technology may feel discomfort or apprehension toward any device that can watch them, like an “invisible person” in the room [134,135,137,143]. Any technology that impinges on this refuge will have lower acceptance rates [31,33]. Informational privacy refers to the desire of a person to control the sharing of personal information with others [136]. Informational privacy can be violated when information is used against the wishes of the stakeholder [135]. As home monitoring devices may store and transmit intimate personal data, many older adults, their family members, and health care professionals have reservations about data collection policies including types of data collected, use of the data, and access levels [31,134,139]. Interestingly, the results of a study focused on designing home monitoring technology found that although information privacy was a priority for engineers and designers, physical privacy implications were not considered [143]. This highlighted the need to understand both the user’s and designer’s perspectives and to think broadly about privacy.

Control

Control has multiple meanings including controlling device data, settings, and who makes decisions about the device and the data it collects [31,37,134]. Older adults desire to maintain control of their lives and surroundings for as long as possible, with many seeing value in home monitoring, but more as a last resort or to be used later in life [31,134,138-140]. Older adults want to control whether and how to use the technology, when it is turned on or off, where it is placed, and with whom the collected data are shared [37,134]. From the designer’s perspective, Birchley et al [143] found that a common trend was to give end users the responsibility to decide what the technology did and how much control it had in their lives. However, it was pointed out that many of the ethical challenges regarding the design and implementation of technology, such as an opt-out feature for data sharing, were addressed in the early phases of design, which placed the responsibility of ethical design with the engineers and designers, not the end users [143].

Social Concerns

As shown in Table 3, social interaction is a growing ethical concern. Many older adults strongly indicate that technology should not replace human contact but should foster and promote human communication and support [140,141]. An increase in assistive technology could mean a decrease in human care and

Table 3. Summary of ethical concerns with smart home technology for the older population (N=19).

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Reviews, n (%)</th>
<th>Methods used</th>
<th>Reason for concern</th>
</tr>
</thead>
</table>
| Privacy [31,33,37,133-142]       | 15 (79)        | Literature review, mixed methods research, SR, and qualitative research | Personal privacy concerns (watched and monitored)  
|                                  |                |                                                   | Informational privacy concerns (data sharing)                                       |
| Control [33,37,133-137,139,140,143] | 10 (53)       | Literature review, mixed methods research, SR, and qualitative research | Fear of losing control and desire to make decisions about the technology (use, on or off, placement, and data collection) |
| Social concerns [33,37,133,135,137,140-142,144] | 9 (47)        | Literature review, mixed methods research, SR, and qualitative research | Desire for face-to-face communication and fear of losing in-person interaction |
| Autonomy [31,33,37,133-135,137,140,143] | 9 (47)        | Literature review, mixed methods research, SR, and qualitative research | Very important for older adults  
|                                  |                |                                                   | Do not want to be a burden on others  
|                                  |                |                                                   | Fear of being dependent                                                             |
| Stereotypes and stigmatization [31,33,134,138] | 5 (26)        | Literature review, mixed methods research, and qualitative research | Fear of judgment and its consequences and stereotypes |

SR: systematic review.
human contact for older adults [37,140,141,145]. Older adults, family members, and health care professionals insist that face-to-face interactions should not be systematized or replaced by technology [37,145]. For many older adults, visits from health care workers are often the only human contact they receive on a day-to-day basis, making this interaction critical to their health and well-being [37,140].

**Stereotypes and Stigmatization**

Stereotypes of “oldness” are often depicted as a time of ill health, dependency, and loneliness [31,33,37]. However, older adults want to be perceived as strong, capable, and independent individuals [31,134]. Any device that projects negative aging stereotypes is likely to be rejected by older adults, even if the device is helpful [31]. Of the older adults interviewed by Astell et al [31], many expressed fears of being judged or discriminated against if they used devices that would stigmatize them as being different, incompetent, or lonely. Participants were more likely to decline social events than to use devices that could incur judgment [31]. Home monitoring devices for older adults must not reflect stereotypes or stigmatizations but should augment how older adults view themselves as independent, competent, and self-reliant users.

**Autonomy**

Autonomy for smart home technology means “the assistive technology developed for elderly care must not interfere with the will of the person it is designed to care for” [133]. Older adults strive to maintain independence and personal autonomy to avoid burdening their loved ones or society [31,37,140,141]. Therefore, devices that enable or prolong independent performance in meaningful activities are met with great enthusiasm, although some still hold reservations regarding how home health monitoring could affect their autonomy and independence [31,37]. Some older adults expressed concerns about becoming overreliant on the devices [146], whereas others did not want technology to complete a task without them [140]. Nevertheless, older adults overwhelmingly agreed that if home health monitoring technologies could preserve their autonomy and accommodate their preferences, using the technology was preferable to moving into a nursing home [140]. Table 3 summarizes the most discussed ethical concerns along with how often they were mentioned in the collected papers.

Although many older adults admit that they see value in home health technology, they also have wide-ranging reservations about it. The ability to anticipate, address, and respond to ethical challenges and concerns is critical for future development and adoption for stakeholders. Communication between all involved stakeholders must occur to better understand the attitudes, concerns, and demands of those who are most impacted by the technology [31,134,147,148].

**Impacts on User Acceptance and Adoption**

**Overview**

In addition to the ethical concerns and challenges for home health monitoring technology discussed, many studies have also examined how ethical challenges and concerns impact user acceptance and adoption of smart home monitoring devices [31,33,37,134,138,142,149-155]. Table 4 highlights the most discussed user aspects found in the literature that influence user acceptance and adoption of home health monitoring technology.

<table>
<thead>
<tr>
<th>Concern</th>
<th>Papers, n (%)</th>
<th>Methods used</th>
<th>Feedback from users</th>
</tr>
</thead>
<tbody>
<tr>
<td>User thoughts and feelings (eg, attitudes, preferences, and knowledge)</td>
<td>10 (53)</td>
<td>Literature review, mixed methods research, and qualitative research</td>
<td>• Positive and negative views on aging change acceptance level</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Limited knowledge on technology but older adults are willing to learn and desire</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>customizability</td>
</tr>
<tr>
<td>User acceptance [31,33,37,134-137,142,143]</td>
<td>9 (47)</td>
<td>Literature review, mixed methods approach, SR, and qualitative research</td>
<td>• Many factors influence acceptance and excitement for smart home technology but</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>do not think they need it</td>
</tr>
<tr>
<td>Usability [37,134,142,144,145]</td>
<td>6 (32)</td>
<td>Literature review, mixed methods research, and qualitative research</td>
<td>• Technology must respect certain values</td>
</tr>
<tr>
<td>Usefulness [37,133,134,142,145,149]</td>
<td>6 (32)</td>
<td>Literature review, mixed methods research, and qualitative research</td>
<td>• Technology must have a purpose</td>
</tr>
<tr>
<td>User adoption or abandonment [31,33,37,142]</td>
<td>4 (21)</td>
<td>Literature review, mixed methods research, and SR</td>
<td>• To promote adoption, stakeholders need to be included in discussions around</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>developing technology</td>
</tr>
</tbody>
</table>

*SR: systematic review.*

Often, health care tools are designed without considering the values and characteristics of the intended users and other key stakeholders, their literacy levels, or their information goals and preferences. This can result in the technology suffering “social
failure modes” [156], ultimately leading to the abandonment of the technology and the need to redesign it [150]. Web-based medical platforms are one example, having low adoptability or acceptance rates [1], potentially because of inadequate user involvement, especially in the early development phases [151-153].

User Acceptance, Usefulness, and Usability
User acceptance is not only based on excitement with technology [31,37] but also on appearance, values and principles, situation, usability, opinions of other direct stakeholders [134], and especially attitude [31,37,134,142]. Older adults who have negative views on aging and associate it with traits such as “illness,” “loneliness,” or “dependency” tend to neglect assistive technology they believe reflects those ideals, whereas those who see aging as positive are more inclined to accept assistive technology and integrate it into their daily lives [31,134]. If end users see no benefit in using an assistive device, there is little chance of acceptance [37,149]. A device that is difficult to operate or understand can lead to frustration and lack of confidence [138,142,149]. Usefulness and usability also tie into developing technology that respects and accommodates end user’s values. If older adults believe that their values are being threatened without explanation, they will likely refuse to accept the technology, whereas they will be more likely to adopt it if those values are upheld [33,134]. All the factors discussed need to be weighed to determine the value of the technology. Tools should be tailored to the users to increase the usability and utility of the technology [154]. The transition from “doing for” users to “doing with” users requires considerable adjustments to be made in both attitude and practice [155].

User Adoption or Abandonment
Beyond the acceptability of technology is the adoption of these devices by end users. Chung et al [37] noted that for aging in place technology to be truly adopted by older adults, devices should address older adults’ values, self-perceptions, and ethical issues at the intersection of aging, technology, and the home environment. Device abandonment is a common reality, whether it is because the device impedes a user’s independence or because of the fear of judgment from their peers [31]. Careful decisions must be made throughout the development process to ensure that the final deliverable is something that is beneficial to the end user and aligns with how the end user sees themselves.

Frameworks for User-Centered Design
Health care tools are often designed without considering the intended users, their literacy levels, information goals, and preferences, which results in user dissatisfaction, leading to the abandonment of the technology and eventually the need to redesign it [150]. Despite constant advancements, web-based medical platforms have low adoptability or acceptance rates [1]. A reason for this is inadequate user involvement, especially in the early development phases [151-153]. We should aim to tailor the technology to the stakeholders’ needs and requirements to increase their usability and utility [154]. User-centered design (UCD) is a framework in which the requirements of stakeholders (including end users) are considered extensively at every stage of the product’s development and design [157]. Table 5 summarizes the various user-centered frameworks developed and adopted in various health care settings based on systematic reviews and individual articles. Moreover, the use of UCD-based evaluation instruments, such as the 11-item measure “UCD-11,” developed by Witteman et al [154] to quantitatively determine the user centeredness of the design and development of health care tools will ensure production of reliable and valid constructs.
Table 5. User-centered frameworks in medical design.

<table>
<thead>
<tr>
<th>Framework</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-phase framework: narrative, metaphorical, and structured [2]</td>
<td>Narrative phase gathers patients’ every-day stories to develop design goals in the metaphorical phase, followed by building technological prototypes in the structured phase</td>
</tr>
<tr>
<td>3-phase iterative framework in repeated cycles [153]</td>
<td>Identifying users and understanding their environments, developing a prototype, and refining the prototype through observations and feedback</td>
</tr>
<tr>
<td>Design thinking framework (5 modes executed iteratively, either sequentially or parallelly) [158]</td>
<td>Understanding user perspective in the empathize mode, synthesizing user feedback into identifiable needs in the define mode, identifying diverse solutions to the problem statement in the ideate mode, constructing basic but actual representations of the ideas in the prototype mode, and simulating the developed prototypes in context of the problem and refining or improving them in the test mode</td>
</tr>
<tr>
<td>Information systems research framework (3-phase framework) in combination with the 5 modes of the design thinking framework [158,159]</td>
<td>The relevance cycle (understanding the end user environment) combined with the empathize and define modes, the design cycle (creating and assessing objects related to the problem) combined with the ideate and prototype modes, and the rigor cycle (appending findings from evaluation in existing knowledge base) combined with the test mode</td>
</tr>
<tr>
<td>The website development model for consumers [150]</td>
<td>Analyzing user requirements; evaluating environments; defining website functions, constraints, visuals and structure; and testing with small-scale user groups</td>
</tr>
<tr>
<td>Wearable device design framework (3-level, top-down) [160]</td>
<td></td>
</tr>
<tr>
<td>The eHix framework [161,162]</td>
<td>Consists of a 20-cell matrix constructed from 4 business model domains (Service, Technology, Organization, and Financial), each having 5 phases of innovation (Inventory—new ideas explored and user needs and requirements identified, Design and Development—construction of prototype application or technology, Experimental—testing the developed prototype in a laboratory setting, Pilot—testing the prototype in daily life scenarios and providing feedback, and Implementation—the finished prototype is deployed). Each cell lays out the required steps and instructions for a particular phase in each domain</td>
</tr>
</tbody>
</table>

*aHix: eHealth innovation matrix.

Discussion

This paper had 2 aims: to present the evidence for home health monitoring through a scoping review and to provide the ethical, user-centered considerations and potential frameworks that can be adopted for developing user-centered health care platforms and personalized support systems. The implementation of remote monitoring requires both a strong backing of the evidence for monitoring to demonstrate its benefits, while maintaining a focus on the ethical and user implications of remote monitoring.

Principal Findings

Evidence for Home Health Monitoring

The first half of this paper presented the evidence related to home health monitoring. We found numerous systematic reviews, with the majority showing positive evidence in monitoring for acute exacerbations of COPD and heart failure, improving blood pressure or diabetes markers, and increasing physical activity levels. When only high- or medium-quality articles were considered, monitoring for COPD and heart failure had the strongest evidence, diabetes had mixed results, and rehabilitation and physical activity had positive results.

Compared to a previous review of systematic reviews on telemedicine in 2010 [163], advancement can be seen in managing heart failure, COPD and for improving HbA1c levels. This result corresponds with the current literature, with a Cochrane review showing positive results in COPD and heart failure [164], another 2 showing improved diabetes management [165,166], and another showing equivalent outcomes comparing telerehabilitation with conventional rehabilitation, although monitoring was not the focus of this review [167].

Ethical Concerns and User-Centered Frameworks

Usability and ethical acceptability are critical for fostering adoption among older adults. The concept of a “trade-off” was key—where multiple factors cannot be attained at the same time, stemming from value tensions between older adults and technology [168]. The value tension between autonomy and privacy was common, where older adults were against being monitored by technology but were willing to relinquish some privacy if they could stay at home longer [37,134,136,139-141,143,149]. Other examples of trade-offs included safety and privacy [134,144], utility and privacy [33,139], and social interaction and privacy [149].

Moreover, end users find disparity between their requirements and the eHealth solutions provided to solve their problems [152].

https://aging.jmir.org/2022/4/e40079
JMIR Aging 2022 | vol. 5 | iss. 4 | e40079 | p.87
(page number not for citation purposes)
Although developers may not approach the problem from the users’ perspective [150,152], UCD in eHealth comes with additional practical challenges, including ambiguity regarding the number of iterative cycles to be conducted, time and cost constraints associated with the process, difficulty in overcoming designer bias, and difficulties in establishing multidisciplinary collaboration [169]. To mitigate such challenges, researchers and developers are recommended to refer to established, flexible, reliable, and valid UCD frameworks with respect to the problem statement and the target population.

**Limitations and Gaps in the Evidence**

The literature on home monitoring is limited. First, the study quality was limited to secondary prevention in heart failure and COPD from 7 high- or medium-quality studies and encouraging rehabilitation and physical activity in 3 studies. Most systematic reviews were of low or critically low quality. We have presented the results of a subgroup analysis of only high- and medium-quality studies to mitigate our own risk of bias. Higher-quality systematic reviews are critical for chronic disease management and generalized monitoring.

Second, the older adult population is not always the focus of reviews. We found only 1 fall detection study in older adults using monitoring devices. A technology survey of fall detection systems found that only 4 out of 57 studies even included older adults in their study [170]. In the screening process, >120 reviews were excluded, as they did not focus on adults aged >65 years. It is important to differentiate these populations, as the acceptability of technology and challenges of multiple morbidities may change the effectiveness of monitoring.

Third, the evidence was sparse for monitoring older adults in general or patients, with most studies suggesting that the technology is not yet mature enough to detect ADLs or cognitive decline [89,92,93]. Most acceptability studies have focused on the domain of general home health monitoring, with the majority being descriptive or showing mixed acceptability. The question of whether it is worthwhile to monitor older adults in general or those with dementia remains unanswered. In addition, telemonitoring and telemedicine are often not differentiated, making it difficult to differentiate the benefits of clinician support from devices and technological support.

From an evidence-focused perspective, monitoring is effective for specific diseases, but challenges remain in researching generalized monitoring for older adults. Clinical outcomes are difficult to measure because objective physiological markers of aging and health utilization need to be measured across years to determine effectiveness. Second, technical obstacles remain in acquiring data from multiple sensors, synchronizing outputs in real time, storing the data, and performing data analytics to detect anomalies or track wellness, all of which can be time consuming and expensive. Finally, the drivers of long-term monitoring are older adults or families focused on the personal benefits of aging in place. Outcomes focused on improving function and independence and older adults’ sense of security and identity may not draw as much broad attention or funding compared with long-term clinical effectiveness and cost-effectiveness studies. Considering the promising results of remote monitoring in specific apps, it is worth researching broader generalized monitoring to improve clinical outcomes.

**Limitations of This Review**

The main limitation of our review was that we did not have a 2-person validation for screenings and extractions. This may have introduced a bias in the included studies and the information extracted. To mitigate the validation of screenings and extractions, we used common inclusion and exclusion criteria and communicated with team members regularly on questionable papers. In addition, our search was limited to clinical databases to focus on clinical effectiveness rather than technological developments; this may have biased our results related to human factors and monitoring. Moreover, it is possible that multiple studies were included in multiple reviews, which may have biased the results. However, this cannot be mitigated. Finally, we identified gaps in the data but were unable to quantify the effect sizes.

Regarding the ethics- and user-centered framework portion of the review, we did not follow a scoping review methodology, as we found that the qualitative nature of the findings could be summarized with a narrative review. However, we may have missed some trade-offs regarding the acceptability of monitoring technologies.

**Future Work**

An interesting area for future research is the extraction of what acceptability tools were used in each study. For the second half of this review, we did not provide an exhaustive list of the ethical concerns that end users have regarding home health monitoring technology. We also limited our study to older adults and did not consider family members; health care professionals; and other stakeholders such as engineers, computer scientists, and designers. More value tensions will arise and need to be addressed with the involvement of more stakeholders.

**Conclusions**

This scoping review provided a summary of the clinical evidence for monitoring older adults in their homes, ethical implications, and user-oriented frameworks found in the literature. Overall, there is promising evidence for monitoring specific diseases and for rehabilitation support, but generalized monitoring for older adults, including cognitive and physical decline, has not been well researched. More clinical research is required for the long-term monitoring of aging in place to provide evidence for its use. To conduct these future studies, we performed a review of important ethical and user considerations and existing user-centered frameworks that must be considered when conducting these monitoring studies. This study demonstrated the need to develop technology with stakeholders rather than for stakeholders to build the evidence for home health monitoring. User-centered frameworks allow stakeholders’ ethical concerns to be addressed and open iterative design opportunities to improve adoption. In developing a system that achieves ethical and UCD, researchers can collect long-term, meaningful data to demonstrate the efficacy of home health monitoring systems for aging in place.
Acknowledgments
This research was funded by the AGE-WELL Core Research Program (AW PP2019 PP3).

Data Availability
We have included the search strings in Multimedia Appendix 2, quality assessments in Multimedia Appendix 3, the extracted articles in Multimedia Appendix 4, and more detailed summaries of the articles extracted in Multimedia Appendix 5 to provide transparency on extractions.

Conflicts of Interest
None declared.

Multimedia Appendix 1
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) checklist for evidence for remote monitoring.
[PDF File (Adobe PDF File), 499 KB - aging_v5i4e40079_app1.pdf]

Multimedia Appendix 2
Search strategy on evidence for remote monitoring.
[DOCX File, 26 KB - aging_v5i4e40079_app2.docx]

Multimedia Appendix 3
Quality assessments for evidence on remote monitoring.
[XLSX File (Microsoft Excel File), 21 KB - aging_v5i4e40079_app3.xlsx]

Multimedia Appendix 4
Extraction tables for evidence on remote monitoring.
[XLSX File (Microsoft Excel File), 45 KB - aging_v5i4e40079_app4.xlsx]

Multimedia Appendix 5
Summary of evidence for remote monitoring.
[DOCX File, 126 KB - aging_v5i4e40079_app5.docx]

References


Abbreviations

ADL: activity of daily living
COPD: chronic obstructive pulmonary disease
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QoL: quality of life
UCD: user-centered design

©Andrew Chan, Rachel Cohen, Katherine-Marie Robinson, Devvrat Bhardwaj, Geoffrey Gregson, Jeffrey W Jutai, Jason Millar, Adriana Ríos Rincón, Atena Roshan Fekr. Originally published in JMIR Aging (https://aging.jmir.org), 28.11.2022. 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.
Impact and Perceived Value of iGeriCare e-Learning Among Dementia Care Partners and Others: Pilot Evaluation Using the IAM4all Questionnaire

Andrew F Scott¹; Stephanie Ayers²*, BA, BEd; Pierre Pluye³, MD, PhD; Roland Grad³, MSc, MD; Richard Sztramko⁴, BSc, MD; Sharon Marr⁴, BSc, Med, MD; Alexandra Papaioannou⁴, BSc, MSc, MD; Sandra Clark², BSc, MSc; Patricia Gerantonis², BSc; Anthony J Levinson¹²*, BA, MA, MSc, MD

¹Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
²Division of e-Learning Innovation, McMaster University, Hamilton, ON, Canada
³Department of Family Medicine, McGill University, Montreal, QC, Canada
⁴Department of Medicine, McMaster University, Hamilton, ON, Canada
*these authors contributed equally

Corresponding Author:
Anthony J Levinson, BA, MA, MSc, MD
Faculty of Health Sciences
McMaster University
1280 Main Street West
Hamilton, ON, L8S 4L8
Canada
Phone: 1 905 525 9140 ext 22210
Email: levinson@mcmaster.ca

Abstract

Background: Care partners of people living with dementia may benefit from web-based education. We developed iGeriCare, an award-winning internet-based platform with 12 multimedia e-learning lessons about dementia.

Objective: Our objective was to evaluate users’ perceptions of impact.

Methods: From March 17, 2021 to May 16, 2022, data were collected upon lesson completion. We used the content-validated Information Assessment Method for all (IAM4all) for patients and the public adapted for dementia care partners. The IAM4all questionnaire assesses outcomes of web-based consumer health information. Responses were collected using SurveyMonkey, and data were analyzed using IBM SPSS Statistics (version 28).

Results: A total of 409 responses were collected, with 389 (95.1%) survey respondents completing the survey. Of 409 respondents, 179 (43.8%) identified as a family or friend care partner, 84 (20.5%) identified as an individual concerned they may have mild cognitive impairment or dementia, 380 (92.9%) identified the lesson as relevant or very relevant, and 403 (98.5%) understood the lesson well or very well. Over half of respondents felt they were motivated to learn more, they were taught something new, or they felt validated in what they do, while some felt reassured or felt that the lesson refreshed their memory. Of 409 respondents, 401 (98%) said they would use the information, in particular, to better understand something, discuss the information with someone else, do things differently, or do something.

Conclusions: Users identified iGeriCare as relevant and beneficial and said that they would use the information. To our knowledge, this is the first time the IAM4all questionnaire has been used to assess patient and caregiver feedback on internet-based dementia education resources. A randomized controlled trial to study feasibility and impact on caregiver knowledge, self-efficacy, and burden is in progress.

(JMIR Aging 2022;5(4):e40357) doi:10.2196/40357

KEYWORDS
dementia; caregiver; web-based education; internet; consumer health information; feedback; perception; survey; questionnaire; patient education; health education; care partner; caregiving; spousal care; informal care; Alzheimer; cognitive impairment; cognitively impaired; Lewy body; gerontology; geriatric

https://aging.jmir.org/2022/4/e40357

XSL-FO

RenderX
Introduction

Informal family care for people living with dementia is important, as 61% of Canadian seniors with dementia live at home [1]. Those living with dementia are often supported by informal care partners who each provide an average of 26 hours of care per week [1,2]. Care partners of people living with dementia may benefit from web-based education to help them develop knowledge and skills to support themselves and the person to whom they are providing care [2,3]. Caregiver education is most commonly delivered through face-to-face interactions during a clinical visit [4]. This can cause issues for clinicians because of time and resource constraints. Access to in-person dementia education resources for caregivers may also be very limited in some rural communities. Internet-based caregiver education has the potential to address some of these challenges and allow dementia caregivers and others to access educational materials at any time and at their own pace.

Although a systematic review performed by Moore and colleagues [5] reported inconclusive evidence supporting the effectiveness of educational interventions focused specifically on dementia progression on dementia caregivers’ knowledge and mental health, individual studies have shown favorable outcomes. The heterogeneity in format, length, instructional design methods, and content of interventions may play a role in some of the inconsistent findings of systematic reviews. This suggests that it may be important to compare educational interventions with more clearly defined instructional designs [6], especially those that follow best practices for web-based learning environments [7]. With the benefits of internet-based interventions on mental health for care partners of adults living with a chronic condition being shown by Sherifali and colleagues [8] and the reported need for high-quality, trusted internet resources for dementia caregivers, it is worthwhile to continue to study high-quality educational interventions that touch on a variety of topics and take into consideration the needs of caregivers [2,3].

For these reasons, we developed iGeriCare, an internet-based platform with 12 multimedia e-learning lessons about dementia. iGeriCare lessons cover a wide range of dementia-related topics, including stages of dementia, treatment options, and caregiver wellness, among others. The lessons were created using best practices in evidence-based instructional design for e-learning [9,10] by experts in dementia care with input from care partners, winning the first-place e-Learning Excellence Award at the 2021 European Conference on e-Learning [11-13]. Since launching in July 2018, iGeriCare has had over 182,000 unique users and over 258,000 user sessions. Lessons are multimedia content with images, audio narration, authentic case studies and examples, as well as some interactivity such as review quizzes, approximately 20 minutes in duration each (Multimedia Appendix 1). All iGeriCare lessons are publicly available and free to access. Our objective was to evaluate users’ impressions of the multimedia lessons, in particular, their perceived relevance, understandability, usefulness, and expected benefit.

Methods

Overview

From March 17, 2021, to May 16, 2022, anonymous data were collected from any person who completed an iGeriCare lesson and opted to fill out a postlesson survey. A total of 12 lessons were available, and the survey was presented after each lesson. Respondents could access iGeriCare in a variety of ways, including a recommendation from a health care professional or teacher, clicking on a link from a social media advertisement or from another website, doing an internet search, receiving the link from family or friends, or being an existing email subscriber to the iGeriCare newsletter campaign.

After lesson completion, participants were asked to rate the lesson on the Net Promoter Score (NPS) scale. NPS is a management tool that can be used to gauge customer satisfaction [14-16]. Using an 11-point scale, the NPS asks respondents their “likelihood to recommend” a product or service based on their experience. Participants were then invited to take the postlesson survey through a link. The 11-question postlesson survey was developed using the Information Assessment Method for all (IAM4all) questionnaire adapted for dementia care partners. The IAM4all is a content-validated questionnaire designed to collect feedback from health information consumers (eg, the public or patients) about resources, and it is structured by four levels of outcomes, including situational relevance, cognitive impact, information use, and health benefits [17]. The survey incorporated conditional logic that allowed respondents to skip some questions based on their responses (eg, if a respondent answered “no” to “Will you use the information from this lesson?” they would automatically skip the question “Please tell us how you will use this information”; Multimedia Appendix 2).

SurveyMonkey was used to collect survey responses, and all data were transferred into IBM SPSS Statistics (version 28; IBM Corp) for statistical analyses. All data from respondents—including those who only partially completed the survey—were included in the data analysis, and the survey data for all lessons were pooled. For analyses of differences in responses by respondent type or role, the Independent-Samples Kruskal-Wallis test with post hoc pairwise comparisons adjusted by the Bonferroni correction for multiple tests was used.

Ethics Approval

The Hamilton Integrated Research Ethics Board reviewed the study protocol and granted exemption from full review per their review process, as this was considered a quality improvement initiative.

Results

Data Trends and Respondent Characteristics

During the 14-month data collection period, there were a total of 2915 iGeriCare lesson completions. Of those, 409 respondents (14% of lesson completers) started an IAM4all survey. Of 409 respondents, 389 (95.1%) who started the postlesson survey completed it, with partial completion for the remaining 20
respondents. Average time spent for survey completion was 3 minutes and 11 seconds. The most commonly completed lessons for survey respondents were the following: “What is dementia?” with 153 of 409 (37.4%) lesson completions; “What is mild cognitive impairment?” with 71 of 409 (17.4%) lesson completions; and “How to promote brain health” completed by 46 of 409 (11.2%) respondents.

A total of 179 of 409 (43.8%) respondents self-identified as a family or friend care partner of a person living with dementia. Those concerned they may have cognitive impairment made up 84 of 409 (20.5%) respondents, with a further 37 (9%) having a diagnosis of Mild Cognitive Impairment (MCI) or dementia. The smallest groups were health students (35/409, 8.6%) and health care professionals (41/409, 10%); 33 of 409 (8.1%) self-identified as “other.” Respondents arrived at the iGeriCare site in various ways, with the most common being the recommendation of a healthcare professional.

IAM4all Domains

Responses were extremely positive with respect to the relevance, understandability, information use, and benefits of the lessons. Of the 409 respondents, 263 (64.3%) perceived the lesson as very relevant, and 117 (28.6%) perceived the lesson as relevant, compared to 27 (6.6%) who said it was somewhat relevant and only 2 (0.5%) who found it not very relevant. 306 of 409 (74.8%) respondents understood the lesson very well, and 97 (23.7%) understood the lesson well; only 6 (1.5%) said they understood it poorly, and no one said they understood it very poorly. Similarly, 401 of 409 (98%) said they would use the information from the lesson, and 378 of 405 (93.3%) expected to benefit from it.

Participants identified a range of thoughts about the lesson they reviewed, with over half stating that it taught them something new, motivated them to learn more, or validated what they were already doing; 194 of 409 (47.4%) respondents found the lessons reassuring, and 154 of 409 (37.7%) said it refreshed their memory. For those who said they would use the information from the lesson (n=396, as 13 respondents skipped this item), 281 of 396 (71%) respondents said they would use it to better understand something; 206 of 396 (52%) said they would use it to discuss with someone else; 134 of 396 (33.8%) said they would use the information to do things differently; and 121 of 396 (30.6%) said they would use it to do something.

Nearly half of 372 respondents identified that the lesson would benefit them by improving their health or well-being or that of the person they care for (Table 1): 177 of 372 (47.6%) said it would help them handle a problem or the worsening of a problem; 120 of 372 (32.3%) said it would prevent a problem; 160 of 372 (43%) felt they would be less worried; and 105 of 372 (28.2%) said it would benefit them by helping them to decide something for someone else.

Table 1. Selected summary of Information Assessment Method for all (IAM4all) questions and participant responses.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Responses, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What did you think about this lesson? (N=409)</strong></td>
<td></td>
</tr>
<tr>
<td>Motivated me to learn</td>
<td>236 (57.7)</td>
</tr>
<tr>
<td>Learned something new</td>
<td>219 (53.5)</td>
</tr>
<tr>
<td>Validated what I do</td>
<td>218 (53.3)</td>
</tr>
<tr>
<td>Reassured me</td>
<td>194 (47.4)</td>
</tr>
<tr>
<td>Refreshed my memory</td>
<td>154 (37.7)</td>
</tr>
<tr>
<td>Did not like this lesson</td>
<td>5 (1.2)</td>
</tr>
<tr>
<td><strong>I will use this information to (N=396)</strong></td>
<td></td>
</tr>
<tr>
<td>Better understand something</td>
<td>281 (71)</td>
</tr>
<tr>
<td>Discuss with someone</td>
<td>206 (52)</td>
</tr>
<tr>
<td>Do things differently</td>
<td>134 (33.8)</td>
</tr>
<tr>
<td>Do something</td>
<td>121 (30.6)</td>
</tr>
<tr>
<td>Other</td>
<td>42 (10.6)</td>
</tr>
<tr>
<td><strong>This information will help me to (N=372)</strong></td>
<td></td>
</tr>
<tr>
<td>Improve my health</td>
<td>181 (48.7)</td>
</tr>
<tr>
<td>Improve the health of someone I care for</td>
<td>177 (47.6)</td>
</tr>
<tr>
<td>Handle a problem</td>
<td>177 (47.6)</td>
</tr>
<tr>
<td>Be less worried</td>
<td>160 (43)</td>
</tr>
<tr>
<td>Prevent a problem</td>
<td>120 (32.3)</td>
</tr>
<tr>
<td>Decide something for someone else</td>
<td>105 (28.2)</td>
</tr>
<tr>
<td>Other</td>
<td>42 (11.3)</td>
</tr>
</tbody>
</table>
Response Differences Between Participant Types

Data collected from IAM4all questions were then analyzed by participant type, that is, family or friend care partner, a respondent concerned they might have MCI or dementia, someone diagnosed with MCI or dementia, a student or trainee, a health care provider, or other participant types. Perceived relevance was significantly higher among family or friend care partners compared to people concerned with possibly having MCI or dementia \([H(5) = 14.533; P = .01]\). Those with a diagnosis or concerned they might have MCI or dementia were significantly less likely to understand a lesson “very well” compared to the other groups \([H(5) = 41.762; P < .001]\), although 117 of those 121 (96.7%) respondents still reported understanding the information “well” or “very well,” with only 4 reporting they understood it “poorly.” Analyses of other survey data did not show any statistically significant differences by respondent role.

Discussion

Principal Findings

The purpose of this work was to evaluate iGeriCare users’ impressions of the multimedia lessons, including their perceived relevance, understandability, usefulness, and expected benefit, using the IAM4all survey. Respondents to the survey identified the iGeriCare lessons as relevant and beneficial, and they said they would use the information. Responses were relatively consistent among different types of participants.

These findings are of interest for two main reasons: first, although the iGeriCare site was primarily designed for family care partners, less than half of respondents identified as a family or friend care partner, implying that the iGeriCare site has appeal to many different audiences, including those concerned they may have a cognitive impairment, those who already have MCI or dementia, and even health care professionals and trainees. Second, the similarity in responses despite the varied audience implies that the iGeriCare lessons are not only appealing to a broader demographic but also relevant and valuable to a wider audience. Although 380 of 409 (92.9%) respondents perceived the lessons as either relevant or very relevant, the finding that the perceived relevance was greater for family and friend care partners compared to those concerned with having MCI or dementia is consistent with the initial design and intent of the site, as the scenarios presented within the e-learning lessons use family care partners.

Furthermore, even though 403 of 409 (98.5%) of respondents understood the lesson well or very well, those with a diagnosis or concerned they might have MCI or dementia were significantly less likely to understand a lesson “very well” compared to other respondent types. With the iGeriCare site having been co-designed with family care partners and not specifically designed for those with cognitive impairment, this finding seems reasonable and highlights a need for similar resources with lessons and scenarios tailored to individuals with these concerns or diagnoses.

Comparison to Prior Work

IAM4all questions measure participants’ perceived relevance, intention to use information, and expected benefit. The high percentage of respondents stating that they would use the information and expect benefits is consistent with the NPS ratings of iGeriCare lessons and comparable to IAM4all feedback on a high-quality parenting website [18]. The percentage of ‘symbolic use’—referring to those that would use the information to discuss with someone else—was higher on iGeriCare, consistent with promotion primarily to family or friend care partners and the fact that care partners made up the largest respondent type for this survey. This likely represents a mix of different people that respondents will discuss the information with, including the person they are caring for, other family or friend care partners, and their health care providers.

Respondents noted that the information in the lessons motivated them to learn more, taught them something new or refreshed their memory, validated their actions, and provided reassurance. This suggests that iGeriCare lessons are helpful in transferring knowledge and improving confidence of dementia caregivers. Considering four IAM4all constructs, ‘conceptual use’ (eg, better understanding) was most frequently noted, followed by ‘symbolic use,’ ‘instrumental use’ (ie, doing things differently), and ‘legitimating use’ (ie, doing something) [19,20]. These constructs within the IAM4all shed a different light on how dementia care partners and others interact with information resources. Previous systematic reviews have shown the impact of internet-based interventions on care partners’ mental health but have rarely described specific uses of information or perceived benefits using a content-validated tool [8,21].

Pluyé and colleagues [17] highlight three ways in which individuals seek health-based information on the internet: professionally mediated (ie, provided by a clinician or librarian), direct access (ie, individual search), and peer mediated (ie, recommendations from family or friends and social media). Respondents in our study arrived at the iGeriCare lessons through all three of these ways and, irrespective of how they arrived at the site, the educational content was highly valued.

Limitations

There are several limitations to this work. First, there is a potential selection bias, as data are from only those users who both completed a lesson and opted to fill out the additional IAM4all questionnaire. Those users who completed a lesson may be more likely to be positively predisposed to the lesson’s relevance, understandability, usefulness, and perceived benefit. However, NPS ratings between respondents and other users of iGeriCare are similar, suggesting that the participants may be comparable to iGeriCare users more broadly. Additionally, despite this potential selection bias, users who completed lessons were the desired target population for the postlesson surveys, since they had engaged with the complete intervention as intended. Nonetheless, future work should consider including a prompt asking respondents who abandon a lesson why they have opted to leave the lesson before completion. Since users can leave and come back to iGeriCare lessons at any time, an option of ‘I intend to finish this lesson later’ should be included. Second, the same individual may have filled out the IAM4all
multiple times for different lessons. Due to the anonymity of the site and survey, we had a limited idea of how many respondents completed the IAM4all for different lessons as well as limited information about participant demographics. However, this was part of an intentional design to facilitate easy access to iGeriCare lessons, rather than requiring users to create a free account. Furthermore, it is reasonable to assume that the same respondent may react differently to each lesson, eliminating the need to identify unique respondents between lessons. For these reasons, future work should continue to remove as many barriers as possible to educational interventions, even if this may impact the extent of research data about participants.

Implications and Future Directions
To our knowledge, this is the first time the IAM4all has been used to assess dementia education resources specifically. It is also the first time that the questionnaire has been used to evaluate instructionally designed, multimedia e-learning, rather than predominantly text-based formats of health information (eg, web pages and email-based content). The questionnaire has helped to better delineate care partners’ and others’ perspectives on the value and perceived impact of the iGeriCare lessons.

More work is needed to determine the effectiveness of web-based dementia education for care partners on other validated outcome measures. A randomized controlled trial to study impact on care partner knowledge, self-efficacy, and sense of burden is in progress (Clinicaltrials.gov identifier NCT05114187).

Acknowledgments
The authors disclose receipt of the following financial support for the research, authorship, and publication of this paper: the iGeriCare initiative was funded by the Canadian Centre for Aging and Brain Health Innovation powered by Baycrest. Additional support was provided by the GERAS Centre for Aging Research at Hamilton Health Sciences; Hamilton Health Sciences Foundation; the Alzheimer Society Foundation of Brant, Haldimand Norfolk, Hamilton, and Halton; McMaster University; the McMaster University Department of Medicine Internal Career Award (RS); and the Division of e-Learning Innovation at McMaster University. AJL receives support through the John R Evans Chair in Health Sciences Educational Research and Instructional Development, Faculty of Health Sciences, McMaster University.

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

Conflicts of Interest
AJL and RS are co-owners of the iGeriCare intervention with McMaster University. RG and PP were involved in the development process of the IAM4all questionnaire [22].

Multimedia Appendix 1
Screenshot from an iGeriCare lesson.
[ PNG File , 429 KB - aging_v5i4e40357_app1.png ]

Multimedia Appendix 2
A sample of the IAM4all questions in the postlesson survey.
[ PNG File , 126 KB - aging_v5i4e40357_app2.png ]

References


Abbreviations
IAM4all: Information Assessment Method for all
MCI: mild cognitive impairment
NPS: net promoter score

Edited by T Leung; submitted 12.07.22; peer-reviewed by A Ramaprasad, T Gladman; comments to author 17.08.22; revised version received 13.09.22; accepted 22.09.22; published 22.12.22.

Please cite as:
Impact and Perceived Value of iGeriCare e-Learning Among Dementia Care Partners and Others: Pilot Evaluation Using the IAM4all Questionnaire
JMIR Aging 2022;5(4):e40357
URL: https://aging.jmir.org/2022/4/e40357
doi:10.2196/40357
PMID:
Electronic Tracking Devices for People With Dementia: Content Analysis of Company Websites

Jared Howes, MA, MSc; Yvonne Denier, PhD; Chris Gastmans, PhD
Centre for Biomedical Ethics and Law, KU Leuven, Leuven, Belgium

Corresponding Author: Jared Howes, MA, MSc
Centre for Biomedical Ethics and Law
KU Leuven
Kapucijnenvoer 35
Box 7001
Leuven, 3000
Belgium
Phone: 32 16372182
Email: jaredmichael.howes@kuleuven.be

Abstract

Background: Electronic tracking devices, also known as locators, monitors, or surveillance devices, are increasingly being used to manage dementia-related wandering and, subsequently, raising various ethical questions. Despite the known importance technology design has on the ethics of technologies, little research has focused on the companies responsible for the design and development of electronic tracking devices. This paper is the first to perform a qualitative analysis of the ethically related content of the websites of companies that design and develop electronic tracking devices.

Objective: The aim of this study was to understand how companies that design, develop, and market electronic tracking devices for dementia care frame, through textual marketing content, the vulnerabilities and needs of persons with dementia and caregivers, the way in which electronic tracking devices respond to these vulnerabilities and needs, and the ethical issues and values at stake.

Methods: Electronic tracking device company websites were identified via a Google search, 2 device recommendation lists (Alzheimer’s Los Angeles and the Canadian Agency for Drugs and Technologies in Health), and the 2 recent reviews of wander management technology by Neubauer et al and Ray et al. To be included, websites must be official representations of companies (not market or third-party websites) developing and selling electronic tracking devices for use in dementia care. The search was conducted on December 22, 2020, returning 199 websites excluding duplicates. Data synthesis and analysis were conducted on the textual content of the included websites using a modified form of the Qualitative Analysis Guide of Leuven.

Results: In total, 29 websites met the inclusion criteria. Most (15/29, 52%) companies were in the United States. The target audience of the websites was largely caregivers. A range of intertwined vulnerabilities facing persons with dementia and their caregivers were identified, and the companies addressed these via care tools that centered on certain values such as providing information while preserving privacy. Life after device implementation was characterized as a world aspired to that sees increased safety for persons with dementia and peace of mind for caregivers.

Conclusions: The way electronic tracking device content is currently conveyed excludes persons with dementia as a target audience. In presenting their products as a response to vulnerabilities, particular values are linked to design elements. A limitation of the results is the opaque nature of website content origins. How or when values arise in the process of design, development, and marketing is unknown. Therefore, further research should explore the process companies use to identify vulnerabilities, how values are decided upon and integrated into the design of products, and the perceptions of developers regarding the ethics of electronic tracking devices.

(JMIR Aging 2022;5(4):e38865) doi:10.2196/38865

KEYWORDS
dementia; wandering; electronic tracking devices; bioethics; locators; monitors; surveillance devices; management; technology; care tool; caregiver; device; vulnerable; elderly
Introduction

Background

Dementia is an umbrella term for a variety of diseases or conditions that cause progressive loss of cognitive function to the extent that daily life and independence become impaired [1]. Persons with dementia will increasingly come to rely on others for their needs and will typically require some form of institutionalized care [2]. Persons with dementia have expressed their strong desire to remain in their own homes and communities for as long as possible [3,4]. Of the many symptoms of dementia, wandering is a particular problem regarding maintaining a person with dementia safely in their home. Wandering involves movements that have a “frequent repetitive, temporally disordered and/or spatially disordered nature...some of which are associated with eloping, eloping attempts or getting lost unless accompanied” [5]. A person with dementia who elopes while wandering faces a high risk of serious injury or death, with time missing predicting higher likelihood of mortality, making swift recovery paramount [6,7]. Even one incident of wandering can precipitate placement into institutionalized care—despite a person with dementia’s active wish to remain at home [8]. The consequences of wandering extend to those closest to persons with dementia as well; for instance, family, friends, and caregivers may undergo psychological stress, physical fatigue, or other adverse effects during a wandering episode [9,10], and communities often expend large sums of resources to search for a wandering person with dementia [11].

Wandering behavior makes persons with dementia vulnerable to various harms. Vulnerability is a common human condition—to be human is to be vulnerable. This general or common vulnerability is, as ten Have [12] states, “a general characteristic of the human condition.” Our vulnerability is constituted on the realities of our embodied existence, “we are characterized by a general fragility or precariousness because we have a finite, mortal body and because we are unavoidably socially related and dependent on others” [12]. Although all persons are vulnerable, certain individuals, groups, and populations experience a special vulnerability and are “more subject to possible harm and violence than others” [12]. In dealing with dementia and its challenges to physical, psychological, and social capabilities, persons with dementia and their families face a variety of particular vulnerabilities that are exacerbated or lessened by their social context (eg, level of accessible care and social support) [13].

In working to address the vulnerabilities of wandering, technological tools have remained a constant potential panacea. A review found that 83 technologies were being implemented into 26 types of wander management devices, with the most common being electronic tracking devices—also known as locators, monitors, or surveillance devices [14]. Electronic tracking devices are technological artifacts that facilitate the monitoring, locating, or logging of a person with dementia’s temporal location. For example, a watch with a built-in GPS can be used with software to track a wearer’s location in real time or save their location history for later analysis. The idea of using electronic tracking devices to manage dementia-related wandering is not new; however, advances in technology miniaturization and general lowering of technology costs combined with governmental initiatives in assistive technology for older adult care have resulted in an overall increase in electronic tracking devices being brought to market [14-17].

The use of electronic tracking devices is not without ethical concern [18]. Many authors have discussed various concerns related to electronic tracking device use in care practices, including their potential to harm persons with dementia through, for instance, increasing social isolation by reducing contact with caregivers [19,20] or helping facilitate unjust control of persons with dementia [21,22]. More traditional clinical-ethical concerns are also discussed, such as the fraught reality of informed consent within care for persons with dementia who experience wandering symptoms [23] or how a caregiver should best respond to persons with dementia who are resistant to electronic tracking device use [19]. Most normative literature focuses on the use of electronic tracking devices [24], although many ethical questions are related to electronic tracking device design. For instance, privacy hinges on design decisions made during the development process, what data are collected, how they are used, and how they are communicated to caregivers, a strong motivation exists for undertaking research regarding electronic tracking device use [19]. Most normative literature focuses on the use of electronic tracking devices [24], although many ethical questions are related to electronic tracking device design. For instance, privacy hinges on design decisions made during the development process, what data are collected, how they are used, and how they are communicated to caregivers, a strong motivation exists for undertaking research regarding electronic tracking device use [19].

Despite the importance of technology design, relatively little research has been conducted regarding companies responsible for the design, development, and marketing of electronic tracking devices. One avenue of gaining insight into companies’ implicit and explicit ethically related insights is to explore their public positions present within their websites and web-based marketing material. Doing so may reveal how they approach key stakeholders, the problems they face, and the ethical issues involved. In addition, internet use by older adults, in general, has seen steady increases [34], and informal caregivers report turning to the internet to find out information about dementia as a disorder, how they can provide better care, and what professional resources are available to them [35]. As a place of interaction between companies, persons with dementia, and caregivers, a strong motivation exists for undertaking research on how companies present their products on the web. However, to date, little research has focused on this subject. The analysis by Vermeer et al [17] of websites based in the United Kingdom, the Netherlands, and Sweden that market electronic tracking devices for use within dementia care is the only current example. Notably, this study had a relatively wide scope via the inclusion of third-party retail websites (eg, Amazon), and no particular...
emphasis was placed on ethical questions as an environmental scanning methodology was used for data analysis.

Objectives

That no study to date has performed an ethical analysis of websites of companies that design and develop their own products is a gap in the literature when viewed in light of the importance of the ethics of technology design, but this gap is more severe when placed in the context of vulnerability. Companies exist within the same social milieu as persons with dementia and their caregivers, one that comprises a multitude of stakeholders, including individuals, companies, and governmental bodies [36]. By placing their products as a means of mitigating the burdens of dementia-related wandering, they insert themselves into the vulnerability of persons with dementia and caregivers, and how they design their products and market them has an effect, whether positive or negative, on the stakeholders involved. With this background, we aimed to gain an understanding of how companies view the vulnerability of persons with dementia and caregivers as well as how they position their products to address these vulnerabilities by performing a qualitative analysis of the written website content for ethically related concepts and themes [37,38]. Hence, our research questions were as follows: how do companies that design, develop, and market electronic tracking devices for dementia care frame, through website textual content, (1) the vulnerabilities and needs of persons with dementia and caregivers, (2) the way in which electronic tracking devices respond to these vulnerabilities and needs, and (3) the ethical issues and values at stake.

Methods

Design

We conducted a search for websites inspired by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement using the Google search engine and then included websites from electronic tracking device recommendation lists of patient advocacy groups and from 2 recent systematic reviews of wander management technology [39]. We chose to develop the search strategy based on PRISMA so that our search process would be conducted in a transparent and widely recognized manner. Next, we performed a qualitative analysis of the included websites’ content using a modified form of the Qualitative Analysis Guide of Leuven (QUAGOL) [40,41].

Search Strategy

Overview

A total of 3 sources were used to identify potential websites. The primary source was the results from a Google search conducted according to PRISMA guidelines. The second and third sources consisted of the electronic tracking device recommendation lists provided by Alzheimer’s Los Angeles (a patient advocacy organization) and the Canadian Agency for Drugs and Technologies in Health and the recent systematic reviews of wander management technology by Neubauer et al [14] and Ray et al [42] (Figure 1). Both the reviews by Neubauer et al [14] and Ray et al [42] were identified after performing a search on PubMed for reviews of electronic tracking devices in the management of dementia-related wandering. The recommendation lists and reviews were included both as additional avenues for identifying potential websites and to serve as a check of the primary Google search.
**Initial Website Identification**

A Google search was the main method for identifying relevant websites. The search engine settings were as follows: the region was set to Belgium, the language was set to English, and 30 results were shown per page. The search string consisted of the following: “dementia” tracking OR locator device wandering -site:pinterest.com -site:amazon.com.

During early piloting, it became clear that web pages from Amazon and Pinterest were obscuring relevant results. As any web page from these websites would be excluded based on the inclusion and exclusion criteria (Textbox 1), the advanced search operator -site: was used to exclude these web pages from the results [43].

The Google search was conducted on December 22, 2020, returning a total of 250 websites. Of these, only the first 72% (180/250) of websites, or the first 6 results pages, were included as the remaining web pages were increasingly irrelevant or defunct. Other website content analysis studies have limited themselves to 50 results [44,45] or 3 results pages [46].

Additional websites were identified from the electronic tracking device recommendation lists provided by Alzheimer’s Los Angeles [47] and the Canadian Agency for Drugs and Technologies in Health [48] and by the reviews by Neubauer et al [14] and Ray et al [42]. Relevant references to electronic tracking devices, companies, or websites were followed up with and added to those found with the Google search (Figure 1). The recommendation lists and reviews were included for triangulation and as a check of the Google search results. An indication that the Google search string was not returning relevant websites would be if no websites from the recommendation lists or reviews appeared in the search results. An overlap of 12 websites was found in both the Google search and recommendation lists and reviews, indicating that our search was returning relevant results. After duplicates were excluded, a total of 199 websites were identified.

![Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart overview of the website identification process [14,42]. CADTH: Canadian Agency for Drugs and Technologies in Health.](https://aging.jmir.org/2022/4/e38865)
Websites were archived on February 15, 2021.

Data extraction involved viewing the archived copy of the websites as it was at a particular time [52]. All 29 included websites were archived on February 15, 2021.

To create an offline archive of the included websites, we used SiteSucker (Rick Cranisky) [51], a computer program for Macintosh OS X that downloads a “[website’s] files to a local folder...while preserving link structure” [52]. This method allows for the preservation of “an entire website in a coherent form” as it was at a particular time [52]. All 29 included websites were archived on February 15, 2021.

Data extraction involved viewing the archived copy of the websites using the Safari web browser (Apple Inc), identifying web pages that contained elements of the inclusion criteria, and exporting them as PDF files.

Characteristics of the companies, websites, and electronic tracking devices were collected via a data collection form (Multimedia Appendix 1). Descriptive statistics were applied to describe the form characteristics of companies, websites, and electronic tracking devices [53]. Textual data on characteristics regarding the content of website information were qualitatively analyzed using a modified form of the QUAGOL [40,41]. Originally developed for qualitative data analysis in original empirical studies, the QUAGOL has also come to be used in analysis of other data sources such as argument-based texts [40,54-57]. The QUAGOL consists of 2 parts, each containing 5 stages. The first part is a thorough preparation of the coding process, and the second part is the actual coding process via a qualitative software program [41]. We modified the QUAGOL in line with previous studies [18,58] by only completing the first 5 stages of the preceding part. The first 5 stages are characterized by an iterative process of diverging deeper into the data, with constant movement between the various stages to draw out insights. Thus, although the 5 stages are presented in a sequential order, in reality, one moves between each stage as well as returns to previous stages throughout the data analysis. In the first stage, all researchers thoroughly read the data to gain a holistic understanding of each website. During the second stage, 1 researcher (JH) developed initial reports for each website that captured their essence by outlining vital characteristics and key points, with focus being placed on ethically related content. Ethically related content was identified based on themes and concepts present in a previously completed review of argument-based ethics literature [18] and those stemming from the relational care ethics approach as operationalized in the dignity-enhancing care model of bioethics [37,38,59]. Although brief paraphrasing of key quotes was included, initial reports remained as close to the data as possible. The third stage entails a movement from initial report to conceptual scheme (Multimedia Appendix 2). Relevant concrete examples from the data are teased out and abstracted into concepts relevant to gaining insight into the research question. It is here that the concepts reported in the results take rudimentary form. To ensure that the concepts were being

### Website Screening

A 2-step process was conducted to screen the websites according to predefined inclusion and exclusion criteria (Textbox 1). The first step was a wide application of the inclusion criteria. A keyword search was conducted on each of the 199 websites to locate references to (1) dementia or Alzheimer disease, (2) tracking device technology, and (3) wandering. If the initial web page did not contain any keywords, a second search was conducted using the advanced search operator site in the Google search engine [43]. This operator allows for the identification of every instance of a keyword on a website. For example, by searching for dementia site:www.Company1.com, every location in which the word dementia appears on any web page within Company1.com would be returned. After the first step, 23.1% (46/199) of the websites generally met the inclusion criteria.

The second step of website screening involved stringently applying the inclusion and exclusion criteria after a close reading of the website content. This involved ensuring that the technology underpinning the device met the definition of an electronic tracking device and that some level of marketing indicated that the device was meant for use in dementia care. A total of 57% (26/46) of the websites met all the inclusion criteria and, with the addition of 3 websites from the authors’ previous awareness, a total of 29 websites were included (Figure 1).

### Data Extraction and Synthesis

Archiving websites as they existed at the time of data extraction is crucial to preserving the broader context in which data are situated and allowing for a precise recall of website content [49,50]. To create an offline archive of the included websites, we used SiteSucker (Rick Cranisky) [51], a computer program for Macintosh OS X that downloads a “[website’s] files to a local folder...while preserving link structure” [52]. This method allows for the preservation of “an entire website in a coherent form” as it was at a particular time [52]. All 29 included websites were archived on February 15, 2021.

Data extraction involved viewing the archived copy of the websites using the Safari web browser (Apple Inc), identifying

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Official website of a company</td>
</tr>
<tr>
<td>Product is advertised for use in managing wandering in persons with dementia</td>
</tr>
<tr>
<td>Product is an electronic tracking device (ETD, which is defined as a technological artifact that facilitates the monitoring, locating, or logging of the temporal location of persons with dementia).</td>
</tr>
<tr>
<td>Website language is English</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market websites, third-party sellers, blogs, and other noncompany websites</td>
</tr>
<tr>
<td>Product is advertised for other population groups (eg, children with autism)</td>
</tr>
<tr>
<td>Product is not an ETD</td>
</tr>
<tr>
<td>Website language is non-English</td>
</tr>
</tbody>
</table>
appropriately and accurately developed, each researcher was given a partially overlapping set of initial reports to independently turn into conceptual schemes. It was found that each author identified the same or similar main concepts, indicating a high level of consensus. During the fourth stage, conceptual schemes were tested against the data for accuracy and appropriateness, with researchers rereading the data and initial reports to see if anything was missing, underreported, or in need of modification. The fifth stage involved merging all individual schemes into a global scheme through a process of forward-backward movement between each scheme to facilitate the identification of common themes and concepts. It is during this stage that the final conceptual categories were tested and refined before being reported in the Results section of this paper. Throughout the entirety of this work, the research team met regularly to discuss the ongoing process, and an archive of reports, conceptual schemes, and merged scheme drafts has been retained.

Table 1. Company characteristics (N=29).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Companies, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Country of company headquarters</strong></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>15 (52)</td>
</tr>
<tr>
<td>Australia</td>
<td>4 (14)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>2 (7)</td>
</tr>
<tr>
<td>New Zealand</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Denmark</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Ireland</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Finland</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Spain</td>
<td>1 (3)</td>
</tr>
<tr>
<td>India</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Singapore</td>
<td>1 (3)</td>
</tr>
<tr>
<td><strong>Founding date of company</strong></td>
<td></td>
</tr>
<tr>
<td>1980 to 1990</td>
<td>4 (14)</td>
</tr>
<tr>
<td>1990 to 2000</td>
<td>3 (10)</td>
</tr>
<tr>
<td>2001 to 2010</td>
<td>5 (17)</td>
</tr>
<tr>
<td>2011 to 2020</td>
<td>9 (31)</td>
</tr>
<tr>
<td>Unknown</td>
<td>8 (28)</td>
</tr>
<tr>
<td><strong>Company type</strong></td>
<td></td>
</tr>
<tr>
<td>Nonprofit</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Private for profit</td>
<td>23 (79)</td>
</tr>
<tr>
<td>Mix (public-private collaboration)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Unknown</td>
<td>2 (7)</td>
</tr>
</tbody>
</table>

**Results**

**Form Characteristics of the Companies, Websites, and Electronic Tracking Devices**

**Company Characteristics**

First, we detected a general line of company form characteristics—their target audiences, used communication and marketing tools, applied language types, and specific electronic tracking device characteristics.

As for the company characteristics, the headquarters of the included electronic tracking device companies were found throughout North America, Europe, Asia, and Oceania. By company organization, most (23/59, 79%) of the included companies were for profit, a few (3/29, 10%) were nonprofit, and 3% (1/29) were public-private collaborations (Table 1).

**Website Characteristics**

Predominantly, 3 audiences were targeted by the website content: informal caregivers, formal caregivers (including institutions), and public safety agencies (eg, law enforcement or the fire department; Table 2). Only 3% (1/29) of the companies explicitly targeted persons with dementia themselves.

In terms of target users, electronic tracking devices advertised for use with persons with dementia were also often advertised for use with other stakeholder groups such as persons living with various cognitive disabilities (eg, autism) and, to a lesser extent, older adults, children, and employees.

The websites used various tools to share information about their products (Table 2). Most (22/29, 76%) companies had a social
media presence, approximately half (12/29, 41%) made use of multimedia resources (eg, video), and all (29/29, 100%) company websites used text-based tools. These information tools were used to convey multiple types of information. Multimedia videos were used to show how a product worked (through guides, overviews, and how-tos) and why they made a difference (testimonials). All (29/29, 100%) websites used text to describe their products, their benefits, and the advantages they conferred to the target audience. Approximately half (15/29, 52%) contained text-based testimonials. A number (14/29, 48%) of sites contained a blog or news section, which served as a space for posting articles related to dementia, wandering, electronic tracking devices, and other general interest stories, and 10% (3/29) of the websites provided a section for academic and professional resources.

The websites made use of differing types of language to describe electronic tracking devices (Table 2), varying from a rather cold and technical kind of tracking language (eg, tracker device) to more neutral language (eg, monitor and locator) to more warm and personal care language (eg, personal alarm and assistive device). These types of language were not mutually exclusive, and some companies described their products using a mix of language types.
Table 2. Website characteristics (N=29).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Companies, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target audience</strong></td>
<td></td>
</tr>
<tr>
<td>Informal caregivers</td>
<td>28 (97)</td>
</tr>
<tr>
<td>Formal caregivers (including institutions)</td>
<td>6 (21)</td>
</tr>
<tr>
<td>Public safety agencies</td>
<td>7 (24)</td>
</tr>
<tr>
<td>Persons with dementia</td>
<td>1 (3)</td>
</tr>
<tr>
<td><strong>Target users</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Persons with dementia</strong></td>
<td></td>
</tr>
<tr>
<td>Persons with dementia exclusively</td>
<td>6 (21)</td>
</tr>
<tr>
<td>Persons with dementia nonexclusively</td>
<td>28 (97)</td>
</tr>
<tr>
<td><strong>Persons without dementia</strong></td>
<td></td>
</tr>
<tr>
<td>Persons living with various cognitive disabilities (eg, autism and Down syndrome)</td>
<td>17 (59)</td>
</tr>
<tr>
<td>Older adults</td>
<td>9 (31)</td>
</tr>
<tr>
<td>Other (eg, children, employees, assets, and pets)</td>
<td>6 (21)</td>
</tr>
<tr>
<td><strong>Information tools</strong></td>
<td></td>
</tr>
<tr>
<td>Social media</td>
<td>22 (76)</td>
</tr>
<tr>
<td>No social media</td>
<td>8 (28)</td>
</tr>
<tr>
<td>Multimedia (eg, video or audio)</td>
<td>12 (41)</td>
</tr>
<tr>
<td>No multimedia (eg, video or audio)</td>
<td>12 (41)</td>
</tr>
<tr>
<td>Textual</td>
<td>29 (100)</td>
</tr>
<tr>
<td><strong>Information types</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Multimedia</strong></td>
<td></td>
</tr>
<tr>
<td>Testimonials</td>
<td>4 (14)</td>
</tr>
<tr>
<td>How-tos or guides or product overview</td>
<td>8 (28)</td>
</tr>
<tr>
<td><strong>Textual</strong></td>
<td></td>
</tr>
<tr>
<td>Product description</td>
<td>29 (100)</td>
</tr>
<tr>
<td>Testimonials</td>
<td>15 (52)</td>
</tr>
<tr>
<td>Blog</td>
<td>7 (24)</td>
</tr>
<tr>
<td>News</td>
<td>7 (24)</td>
</tr>
<tr>
<td>Academic or professional resources</td>
<td>3 (10)</td>
</tr>
<tr>
<td><strong>Language used to describe ETDs</strong></td>
<td></td>
</tr>
<tr>
<td>Tracking or tracker device</td>
<td>15 (52)</td>
</tr>
<tr>
<td>Tag</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Mobile locator tracking unit</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Location-tracking app</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Monitoring device</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Locator</td>
<td>8 (28)</td>
</tr>
<tr>
<td>Wearable</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Safety app</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Safety device</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Personal emergency location device</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Personal emergency response system</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Characteristics</td>
<td>Companies, n (%)</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Assistive device</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Personal alarm</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Wandering alert</td>
<td>1 (3)</td>
</tr>
</tbody>
</table>

*The companies may have multiple characteristics.

The companies focus on persons (eg, children) with Down syndrome who do not have dementia.

ETD: electronic tracking device.

**Electronic Tracking Device Characteristics**

With regard to the electronic tracking device characteristics, the most common electronic tracking device form factor was a fob or tag, followed by wristbands (Table S1 in Multimedia Appendix 3). Most electronic tracking devices used a combination of GPS and mobile network technology, with some also including additional technology. Most device prices ranged between US $1 and US $500. Some electronic tracking devices required a subscription to operate, most costing between US $1 and US $50 per month.

**Characteristics Regarding Content of Website Information**

Second, we found that the electronic tracking device company websites were largely organized around three overarching dimensions describing (1) the vulnerabilities that persons with dementia and their caregivers have to face, (2) the care tools that electronic tracking devices possess that address these vulnerabilities, and (3) a conception of the world that is aspired to and comes about from using the presented electronic tracking device. In the following sections, we will describe these 3 dimensions in more detail.

**Expressed Vulnerabilities**

**Overview**

During the data analysis process, it became clear that the companies identified a range of vulnerabilities facing both persons with dementia and their formal or informal caregivers. These vulnerabilities stemmed from dealing with dementia-related wandering and from using electronic tracking devices themselves. They can be grouped into the following categories: bodily, psychological, social, existential, moral, and technical vulnerabilities (Table 3).
Table 3. Overview of company-identified vulnerabilities (N=29).

<table>
<thead>
<tr>
<th>Vulnerability</th>
<th>Companies, n (%)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bodily vulnerabilities</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Persons with dementia</strong></td>
<td></td>
</tr>
<tr>
<td>Wandering</td>
<td>26 (90)</td>
</tr>
<tr>
<td><strong>Caregivers</strong></td>
<td></td>
</tr>
<tr>
<td>Limited in body</td>
<td>3 (10)</td>
</tr>
<tr>
<td><strong>Psychological vulnerabilities</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Persons with dementia</strong></td>
<td></td>
</tr>
<tr>
<td>Loss of independence</td>
<td>26 (90)</td>
</tr>
<tr>
<td>Feeling insecure or unsafe</td>
<td>6 (21)</td>
</tr>
<tr>
<td>Feeling like a burden on family</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Psychological or emotional stress</td>
<td>4 (14)</td>
</tr>
<tr>
<td><strong>Caregivers</strong></td>
<td></td>
</tr>
<tr>
<td>Psychological or emotional stress</td>
<td>21 (72)</td>
</tr>
<tr>
<td>Lack of knowledge</td>
<td>6 (21)</td>
</tr>
<tr>
<td>Loss of independence</td>
<td>5 (17)</td>
</tr>
<tr>
<td><strong>Social vulnerabilities</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Persons with dementia</strong></td>
<td></td>
</tr>
<tr>
<td>Erosion of social life (eg, isolation)</td>
<td>3 (10)</td>
</tr>
<tr>
<td><strong>Caregivers</strong></td>
<td></td>
</tr>
<tr>
<td>Erosion of relationship with person with dementia</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Need for external support or help</td>
<td>4 (14)</td>
</tr>
<tr>
<td><strong>Existential and moral vulnerabilities</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Persons with dementia</strong></td>
<td></td>
</tr>
<tr>
<td>Erosion in quality of life</td>
<td>8 (28)</td>
</tr>
<tr>
<td><strong>Caregivers</strong></td>
<td></td>
</tr>
<tr>
<td>Faced with difficult decisions</td>
<td>6 (21)</td>
</tr>
<tr>
<td><strong>Technical vulnerabilities</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Persons with dementia</strong></td>
<td></td>
</tr>
<tr>
<td>Stigma derived from ETDs(^b)</td>
<td>7 (24)</td>
</tr>
<tr>
<td>Physical discomfort from ETDs</td>
<td>5 (17)</td>
</tr>
<tr>
<td><strong>Caregivers</strong></td>
<td></td>
</tr>
<tr>
<td>Exacerbated vulnerabilities from difficult-to-use ETDs</td>
<td>17 (59)</td>
</tr>
</tbody>
</table>

\(^a\)The companies may have multiple vulnerabilities.

\(^b\)ETD: electronic tracking device.

**Bodily Vulnerabilities**

Bodily vulnerability relates to potential physical risks and harms. All (29/29, 100%) companies identified the physical risks related to wandering as the principal vulnerability facing persons with dementia. Becoming lost can lead to serious injury or even death. This vulnerability is the primary reason that companies have developed electronic tracking devices and, in some sense, is the cause of all other identified vulnerabilities. Caregivers’ bodily vulnerabilities were also described, albeit to a lesser extent. The companies pointed to caregivers’ physical limitations (such as increasing fatigue or exhaustion) that can be reached when they are required to monitor, care for, or otherwise attend to a person with dementia at all times.

**Psychological Vulnerabilities**

According to the company websites, the predominant vulnerability facing caregivers is psychological or emotional stress. As persons with dementia experience more severe symptoms of dementia and begin to wander, caregivers are...
confronted with stress-inducing situations. Fearing for the safety of their loved one, they may feel unconfident in current safety arrangements but lack the knowledge to rectify or improve the situation. As the care needs of a person with dementia grow, caregivers may become increasingly focused on providing care to the exclusion of personal interests or pursuits, making them vulnerable to losing their own independence.

Persons with dementia, too, are psychologically vulnerable, with many companies underscoring loss of independence. This loss can be external in origin through, for instance, limitations placed by caregivers, but it may arise internally as well; persons with dementia who are aware of their potential for wandering may feel insecure or unsafe and, thereby, self-limited. The loss of independence is seen as a contributor to feelings of isolation and loneliness and, as care needs continue to rise, of being a burden upon one’s family. These specific examples coincide with companies alluding to the general psychological or emotional stress that can occur during routine dementia care.

Social Vulnerabilities

Various social vulnerabilities were found within the websites’ content. Persons with dementia may experience an erosion in their social life, both because of the course of dementia and as a result of potential restrictions aimed at reducing the physical and psychological risks of wandering. Informal caregivers, too, may see their social lives harmed. As they devote increasing time to care responsibilities, they may set aside their own social engagements and responsibilities; in the case of children caring for parents, life plans may be delayed or changed.

Existential and Moral Vulnerabilities

Existential vulnerabilities entail the risk of harms related to one’s existential self and one’s own experience of being able to live a meaningful life. In this regard, the company websites referred to the risk of being institutionalized against one’s wishes, which is closely associated with the potential for the person with dementia to experience a general degradation in quality of life. Moral vulnerability deals with the difficulties surrounding moral decisions. Many companies acknowledged that caregivers are vulnerable when faced with difficult care decisions, often running into difficulties in balancing the aim to keep a person with dementia safe with the wish to maintain a person with dementia’s independence and autonomy in leading their own life.

Technical Vulnerabilities

The companies were very cognizant of the new vulnerabilities that electronic tracking devices introduce in persons with dementia and their caregivers. Difficult-to-use technology may exacerbate the psychological and emotional stress of persons with dementia and caregivers. Furthermore, if an electronic tracking device is not able to easily integrate into the daily life or habits of persons with dementia and their caregivers, it may lead them to missing the full benefits of an electronic tracking device or, in a worst-case scenario, outright abandon the technology. For example, “choosing a GPS transmitter designed for shoes is not good if the person with dementia has the habit of taking their shoes off” (company 12). The form factor of the device itself opens persons with dementia up to new vulnerabilities. Physical discomfort may increase agitation or stress in persons with dementia, and they may even injure themselves if they attempt to remove an unwanted or unfamiliar device. Device esthetics may also contribute to the stigmatization of persons with dementia.

Care Tools of the Electronic Tracking Device: Addressing the Vulnerabilities

Overview

The second overarching dimension that was present in the company websites represents the various ways in which the care tools that electronic tracking devices possess address the aforementioned vulnerabilities. By outlining specific device designs, functions, or capabilities, the companies positioned their electronic tracking devices as the right care tool for helping both persons with dementia and their caregivers address or even overcome their vulnerabilities. Accordingly, we detected several categories of tools that serve as a response, viz. providing information, supporting communication, creating communities of care, and focusing on adjusted and user-friendly design (Table 4).
Table 4. Overview of electronic tracking device care tools (N=29).

<table>
<thead>
<tr>
<th>Care tool</th>
<th>Companies, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Providing information</strong></td>
<td></td>
</tr>
<tr>
<td>High level of information availability</td>
<td>16 (55)</td>
</tr>
<tr>
<td>Medium level of information availability</td>
<td>4 (14)</td>
</tr>
<tr>
<td>Low level of information availability</td>
<td>5 (17)</td>
</tr>
<tr>
<td><strong>Supporting communication</strong></td>
<td></td>
</tr>
<tr>
<td>Two-way</td>
<td>8 (28)</td>
</tr>
<tr>
<td>One-way</td>
<td>3 (10)</td>
</tr>
<tr>
<td>SOS function</td>
<td>10 (34)</td>
</tr>
<tr>
<td><strong>Realizing communities of care</strong></td>
<td></td>
</tr>
<tr>
<td>Personal care community</td>
<td>5 (17)</td>
</tr>
<tr>
<td>Formal care community</td>
<td>10 (34)</td>
</tr>
<tr>
<td><strong>Adjusted and user-friendly design</strong></td>
<td></td>
</tr>
<tr>
<td>Easy-to-use design</td>
<td>17 (59)</td>
</tr>
<tr>
<td>Form factor</td>
<td>11 (38)</td>
</tr>
</tbody>
</table>

*The companies may have multiple care tools.

**Providing Information**

In total, 3 levels of information availability were offered by marketed electronic tracking devices. At the highest level, the companies granted caregivers near-total access to the real-time location of persons with dementia. The motivation behind this decision varies. Information plays an important role in providing peace of mind to caregivers and improving overall care quality by providing caregivers with the tools necessary to keep the person with dementia safe. Increased information about a person with dementia increases a caregiver’s situational awareness and enables them to proactively protect the person with dementia by monitoring, adjusting, and intervening when necessary. In addition, with greater information, direct control is given back to caregivers. Being able to modify various parameters, such as the geographical zone where a person with dementia can walk freely, empowers caregivers in managing their loved one’s condition. Devices marketed toward institutional care settings focused on better resident management, for example, by showing the location of all residents on a facility map.

At the middle level of information availability were devices that allow a caregiver to view a person with dementia’s location only if certain thresholds are met, for instance, when a person with dementia leaves a defined geographical zone, a set ambient temperature is reached, or on a fixed time schedule. The motivation for such a mediated flow of information is based upon a balance between the ethical values of autonomy and bodily integrity. The companies strived to uphold a person with dementia’s privacy and freedom (autonomy) while also attempting to keep them safe (bodily integrity). The companies advertised that their product gave persons with dementia “more privacy, greater autonomy and more reassurance.” An additional motivation is the desire to optimize battery capacity, with periodic rather than constant updating increasing time between charges.

The lowest level of information availability consisted of devices whose underlying technology cannot facilitate real-time tracking or localizing. Similar to those in the higher levels, these devices can raise an alarm when a person with dementia leaves a defined geographical zone, with a radio receiver then being used to hone in on the person with dementia. The choice to use this technology centers on producing a simple, reliable, and affordable product. As company 4 stated, “telemetry tracking is not rocket science.” Such technology also comes standard with a very long battery life (≥6 months), meaning caregivers do not need to remember to charge or replace batteries often. An additional benefit of using this technology is that privacy is more readily protected as a person with dementia can only be tracked if the identification number for the receiver is known.

**Supporting Communication**

Many companies whose devices offered high- or middle-level information availability highlighted communication as an important care tool. Two-way communication is a feature that allows a caregiver to speak with a person with dementia as if over the phone, allowing for greater interpersonal connection as well as enabling a caregiver to speak with a person with dementia during an emergency. Often, the person with dementia does not need to do anything as the electronic tracking device automatically answers the call. Another feature in some devices is one-way listening. With this, a caregiver is able to listen in and hear what is happening in a person with dementia’s surroundings. A communication tool unique to the person with dementia is an SOS emergency button. When a person with dementia presses the SOS button, an emergency notification or alert is sent to caregivers or a monitoring service indicating that something is wrong.

**Creating Communities of Care**

Creating communities of care was also present within company website content as an answer to address existing vulnerabilities.
The companies built tools into their electronic tracking devices that were geared toward supporting informal caregivers by enabling them to create a personal care community. Some electronic tracking devices are usable only within formal care communities such as institutional care facilities, emergency services (eg, police department and fire department), or local nonprofit chapters. The motivation behind embedding an electronic tracking device into such a community varies. The institutional knowledge cultivated and stewarded by formal care communities may provide unique benefits to persons with dementia and caregivers as enrollment means becoming “a part of a community that is dedicated to their safety and well-being” (company 14). In addition, the companies can ensure that their products are not being misused, for example, using an electronic tracking device on someone without a cognitive condition or medical need.

**Adjusted and User-friendly Design**

The companies emphasized that their products were easy to operate and integrate into daily life. This is seen as important for meeting the needs of persons with dementia and caregivers as complicated and difficult technology may increase stress and frustration, and technology that is hard to integrate into life has a high chance of abandonment. A number of electronic tracking device design elements are marketed as contributing to ease of operation. Software design such as font size or an easy-to-use mobile app or website contribute to easier operation. So, too, does physical design, with companies focusing on the frequency and method in which an electronic tracking device will need to be charged, the shape and feel of the device, and the manner in which it is worn.

A device that a person with dementia will not wear is not useful. Physical form factor, such as comfortability or style, is seen as important for electronic tracking device acceptance by persons with dementia. Sleek, stylish, or discreet devices are also put forward to protect persons with dementia from being stigmatized, experiencing indignity, or becoming targets for criminal activity. Company 13 pointed to the benefits of a small, lightweight, and unobtrusive device: “Don’t embarrass Mom with garish panic jewelry.”

A key point of design for improved integration is customizability. Not all persons with dementia and caregivers have the same vulnerabilities, problems, or lifestyles and, therefore, having an electronic tracking device that can be adapted to suit individual contexts better meets real-world needs. Customization can range from the manner in which an electronic tracking device is worn, the length of time that movement history is stored, frequency of updates or notifications, defined geographic “safe” zones, and more. This customizability also extends to formal care institutions, with many companies offering to tailor products and services to the needs of particular institutions.

**A World Aspired to**

Overview

The third overarching dimension that was present in the company websites represents a conception of the world that is aspired to and comes about from using the presented electronic tracking device. Many companies offered a vision of life after the implementation of their product. This “world aspired to” is how the companies anticipated their electronic tracking devices would change the daily lives of persons with dementia and their caregivers (Table 5).
Table 5. Anticipated benefits and recognized limitations of electronic tracking device (ETD) use (N=29).

<table>
<thead>
<tr>
<th>Anticipated benefits</th>
<th>Companies, n (%)^a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Persons with dementia</strong></td>
<td></td>
</tr>
<tr>
<td>Increase safety</td>
<td>23 (79)</td>
</tr>
<tr>
<td>Maintain or increase independence</td>
<td>11 (38)</td>
</tr>
<tr>
<td>Improve psychological or emotional state</td>
<td>9 (31)</td>
</tr>
<tr>
<td><strong>Caregivers</strong></td>
<td></td>
</tr>
<tr>
<td>Increase peace of mind</td>
<td>21 (72)</td>
</tr>
<tr>
<td>Increase knowledge</td>
<td>9 (31)</td>
</tr>
<tr>
<td>Enable higher-quality care</td>
<td>7 (24)</td>
</tr>
<tr>
<td>Increase caregiver independence</td>
<td>4 (14)</td>
</tr>
<tr>
<td><strong>Both persons with dementia and caregivers</strong></td>
<td></td>
</tr>
<tr>
<td>ETD is affordable and cost-effective</td>
<td>6 (21)</td>
</tr>
<tr>
<td>Improve interpersonal relationship</td>
<td>3 (10)</td>
</tr>
<tr>
<td><strong>Recognized limitations of ETDs</strong></td>
<td></td>
</tr>
<tr>
<td>Technical limitations</td>
<td>7 (24)</td>
</tr>
<tr>
<td>Role and fit limitations</td>
<td>6 (21)</td>
</tr>
<tr>
<td>Price limitations</td>
<td>6 (21)</td>
</tr>
<tr>
<td>Ethical limitations</td>
<td>4 (14)</td>
</tr>
</tbody>
</table>

^aThe companies may appear in multiple categories.

**Bodily Safety and Psychological Peace of Mind**

For persons with dementia, the most saturated aspiration put forward by the companies was increased bodily safety. The implementation of electronic tracking devices will result in reducing the risks associated with wandering as a person with dementia can be easily kept track of and located during an emergency. As a result of bolstered bodily integrity, the companies envisioned a range of new benefits for persons with dementia. Psychologically and emotionally, persons with dementia will see their independence maintained or increased. As company 16 described, “[t]he Safer Walking GPS Locator exists to provide a safeguarding measure so that a person living with dementia can be encouraged to continue to get outside and walk, live well and enjoy social interaction and independence.” In addition, persons with dementia may experience a general improvement in their psychological and emotional state being confident in their situation, feeling safer or more secure, and not having “the unpleasant feeling of being a burden or strain on their relatives” (company 12).

In the same way, the companies aspired to a world where a caregiver choosing to integrate an electronic tracking device into their care environment would result in greater peace of mind. Peace of mind is a common refrain that coincides with language that sees worry, anxiety, fear, panic, and other psychological and emotional stresses reduced or eliminated. As company 19 stated, “[r]elax and know your loved ones are just a click away. No panic, no worry.” A number of companies also saw their electronic tracking device providing caregivers with more knowledge about their care receiver. This increase in knowledge is related to the various metrics that different electronic tracking devices can track, log, and report. In turn, some companies pointed to data analytics as a way for caregivers, both professional and informal, to “understand how the client is doing” and “establish a baseline of usual, expected behavior” (company 22). These examples helped build toward the aspiration that electronic tracking devices will bolster the relationship between caregiver and person with dementia as well as enhance the ability of caregivers to provide higher-quality care. Finally, some companies saw electronic tracking devices as a way to provide caregivers with opportunities to gain back their own independence through, for example, pursuit of personal interests or moments of respite. As company 12 stated in their company manifesto, “our focus is on ‘freedom’ both for people with dementia and their relatives.”

**No Ideal World**

Although, in many cases, companies had an idealized vision of their products’ impact on stakeholders, some recognized that the future they aspire to is not an ideal world where all worries and vulnerabilities will be duly addressed. Rather, the context and situation will be bounded by the limitations of the devices themselves. The limitations of electronic tracking devices include their role and fit, price, and technological underpinnings. In addition to these, some companies highlighted ethical limitations that persist through device adoption (Table 5).

Electronic tracking devices are envisioned as a powerful tool, aide, and support to caregivers of persons with dementia, capable of providing numerous benefits; however, their role is precisely
this—a tool. These devices are not meant to serve as a replacement for the watchful eyes of a human caregiver. Specifying further, many companies were candid about the potential misfit between their device and the needs of a specific caregiver and person with dementia. Not everyone is a good candidate for an electronic tracking device and, although the companies had strived to develop flexible products that could meet any everyday scenario, ultimately, some conceded that other devices may serve caregivers and persons with dementia better; as company 2 stated, “we’re not the only game in town, so don’t hesitate to consider other options.”

The companies were very aware that their target customers (often being on a fixed income) are sensitive to the price of electronic tracking devices. Most companies described their products as being “cost effective” or “affordable,” underscoring the value for price that their package of care tools offered. Coinciding with value propositions, many companies offered resources aimed at helping caregivers procure funding for devices, whether this be through petitioning health insurance companies to reimburse or outright purchase a device, grants from a local charity, or a payment plan.

The worlds that the companies aspire to were often optimistic about the impact of electronic tracking device technologies. This was despite their many limitations. The companies were quick to highlight the technological problems of competitors, with only a few identifying potential limitations within their own choice of technology.

Most companies did not explicitly broach the topic of ethics. A few did explicitly discuss ethics by acknowledging that, despite the benefits that electronic tracking devices bring, ethical conflicts will persist, such as issues related to privacy. On a web page dedicated to “Ethical and technical challenges of GPS transmitters,” company 12 stated the following under the heading “privacy and autonomy”:

The benefit of using GPS transmitters are so great that they to some extent outweigh the disadvantages. The location system may help ensure that the person with dementia does not wander. The main drawback is that most GPS transmitters monitor and record every step the person takes. The person with dementia is therefore monitored 24/7 and not only when needed. So it might be difficult to persuade the person with dementia to use a location system. In addition to concerns over monitoring, people with dementia may often reject things they are not familiar with. Even though you may have gone to a lot of trouble to introduce the location system, some people with dementia will still refuse to have anything to do with it. If this is the case, the best way of showing care is to think about safety rather than having a bad conscience about concealing the location device.

Many companies advised concealing a device on a resistant person with dementia; for instance, company 13 highlighted that its device could easily be slipped into a pocket, stashed in a car, or tucked into a purse or backpack (“don’t embarrass Mom with garish panic jewelry”). Others touted the concealability of their product. Company 12 wrote the following:

“In the early stages of Alzheimer’s and dementia, the ability to hide the technology inside a shoe preserves the privacy and dignity of the wearer.”

Despite the emphasis on safety, privacy is still seen as an important value to uphold and, as seen previously, many companies intentionally designed their device to maximize the privacy of persons with dementia while maintaining their safety. Therefore, even though electronic tracking devices have great potential in helping meet the needs of persons with dementia and caregivers, conflicts will remain a problem in the envisioned world of electronic tracking device companies. Another concern is the willful misuse of electronic tracking devices. As mentioned previously, some companies anticipated that their product might be used inappropriately and had taken steps to prevent this. By doing so, these companies defined the boundaries of what they saw as ethically permissible regarding the use of their devices.

Characteristics Regarding the Narrative Style of Websites

Next to form and content characteristics, we also found that the manner in which the companies delivered their content could be placed on 3 narrative continua. The first is between an idealistic and realistic narrative, the second is between a technical and human narrative, and the third is between a company-centric and cocreation narrative.

Idealistic to Realistic

Companies using an idealistic narrative spoke about their electronic tracking devices’ capabilities and impact on persons with dementia and caregivers in an idealized manner, using language that did not recognize the real limitations or realities of their electronic tracking devices. For example, company 6 stated that their device had the highest level of accuracy, “ensuring your loved ones and belongings are safe & secure.”

When limitations were mentioned, idealistic websites might use hedging language to mitigate perceived impact. Company 13 hedged their device’s reliance on cell networks for high-accuracy location services by saying the following: “When cell service fails [company 13’s device] uses satellite coverage as a back-up, though accuracy may be reduced.”

Limitations in this back-up coverage, nevertheless, remained unmentioned.

On the other side of the continuum, those companies using a realistic narrative provided a more balanced assessment of their products’ strengths, weaknesses, and potential impact on persons with dementia and caregivers. This assessment was distinct from disclaimers or limitations of liability found within contractual documents (eg, terms of service). Company 17 provided a paradigmatic example:

We would like to say there is a 100% guarantee [of locating a missing loved one]; however, there is no such thing, regardless of the technology or method utilized. It must always be remembered you are dealing with people and each situation is different...The best protection is to have trained personnel with the latest equipment available, respond and conduct these searches.
Realistic websites might still portray their devices as being capable of providing great benefit to persons with dementia and caregivers but temper idealized expectations.

**Technical to Human**

The second continuum, technical to human, centers on the approach to addressing the problems of wandering. Websites that used a technical narrative approached wandering as a problem for which technical solutions exist, emphasizing the technical specifications and abilities of their devices as the solution to the problems of wandering. Company 3 is an example of this, with the technical aspects of their electronic tracking device being the center of their marketing content:

> [Company 3’s device] provides a way to monitor someone in the home who is at risk of wandering. The system alerts the caregiver when the wanderer goes beyond a set range and provides tracking capability up to 1 mile so that the wanderer can be located and returned.

A human narrative embeds the problems of wandering within the social and relational contexts of persons with dementia and their caregivers. It is not just a technical issue but a human problem. Accordingly, companies that used the human narrative discussed the psychological, relational, and emotional experiences that persons with dementia and caregivers may go through while dealing with wandering and, in this way, acknowledged the complexities of this situation. A clear example of this is the following excerpt from company 1:

> Patients with Alzheimer’s and dementia are often frustrated by the worry and concern their families show. Especially in the early stages of memory loss and confusion, families often find themselves struggling with wanting to keep their loved ones safe, while at the same time not stifling their freedom and independence. It’s a difficult and emotional time for everyone involved, but with the Company 1 GPS Safety Device for Dementia you can find the peace of mind you need to help maintain independence for as long as possible without compromising safety.

**Company-Centric to Cocreation**

The final continuum, company-centric to cocreation, centers on how companies articulated some of the elements that contributed to the design process. Company-centric narratives centered on companies bringing their specialized knowledge to bear on the problems of wandering. Whether other stakeholders were included in the process of design was not mentioned. Company 4 stated that “[w]e invented at-risk people tracking in 1986. We started it all. Advantage: Experience.” On the other end of the continuum lie narratives of cocreation. These companies emphasized that the design process included stakeholders such as persons with dementia, caregivers, older persons, family members, and patient organizations. For example, company 12 repeatedly stated that their product was “developed in collaboration with relatives, caregivers and people with dementia.”

---

**Discussion**

**Principal Findings**

**Overview**

This is the first study to examine the ethically relevant web-based material of those companies responsible for the design and development of electronic tracking devices. As such, it offers an initial exploration of how those organizations publicly approach ethics in relation to electronic tracking devices in dementia care, revealing that, in general, the companies recognized vulnerabilities facing not only persons with dementia but caregivers as well and, in many cases, acknowledged the 2 as being closely intertwined (eg, bodily vulnerability of persons with dementia contributes to psychological vulnerability). Although the companies’ responses to these needs differed both in care tools offered and values advanced, they affirmed the importance of the safety and autonomy of persons with dementia and caregivers’ psychological well-being as values central to their devices. Against this background, 2 specific points are particularly impactful. The first deals with the missing place of persons with dementia within the website content, and the second deals with the presence of a relationship between values and electronic tracking device design.

**The Place in Content of Persons With Dementia**

A major insight from this study is the place persons with dementia occupy within electronic tracking device companies’ website content. Beyond perhaps a few web pages, persons with dementia were largely relegated to the status of third party, being the topic of the conversation rather than a participant. This was evident in the target audience of websites that catered to formal or informal caregivers. For example, the marketing of care tools was from the perspective of what they offered caregivers, that is, how they assisted caregivers’ care practices. This focus on caregivers over persons with dementia is encapsulated in common sentence formulations where the caregiver served as subject and persons with dementia served as related possession—a care tool benefits you (ie, caregiver) and your person with dementia, as seen, for example, in the following quotation from company 9 (emphasis ours): “Designed to help you and your loved one feel safe, in control and have peace of mind.”

This conclusion builds upon previous research—which found no person with dementia-centric marketing content within retail websites based in the Netherlands, the United Kingdom, and Sweden [17]—by demonstrating that the relegation of persons with dementia to a third party occurs at the level of organizations responsible for the development of electronic tracking devices. That companies do not create content directed at persons with dementia is striking given the larger push in dementia care to create opportunities of empowerment for persons with dementia through care practices, shared and supported decision-making models, and a movement toward person-centered care [60]. Even though the companies touted their products as enhancing the autonomy and freedom of persons with dementia while maintaining their safety, the actual choice to use these devices appears to be completely left out of their hands and placed solely
within the purview of caregivers. This also raises questions regarding the role of companies in the informed consent process, particularly when selling directly to informal caregivers.

In addition to being a third party within the marketing material, a close examination of the offered care tools shows that most tools acted upon persons with dementia while largely empowering caregivers. With the notable exception of those devices that offered SOS functionality or person with dementia–initiated phone calls, most tools involved caregivers acting upon persons with dementia, for instance, one-way listening or tracking metrics such as location. This emphasis on caregivers may lead to questions about who this device is truly meant for, contributing to the lack of consensus on the nature of electronic tracking devices, which have been suggested to be a form of assistive technology [61], surveillance technology [62], telecare [63], smart wearables [42], and monitoring technology [64], among others.

It is known that persons with dementia have a strong desire to maintain their autonomy in decision-making concerning electronic tracking devices [65,66]. In this regard, a key takeaway from this study is that electronic tracking device companies should invest in making their marketing more inclusive of persons with dementia by centering their focus on the relationship between persons with dementia and caregivers. This reorientation toward the relational both brings persons with dementia into the conversation about their own future care and bolsters the relationships that undergird and enable this very autonomy [58]. Thus, rather than simply moving content focus from one artificially isolated audience to another, a more accurate picture of the real-life relational context in which decisions about electronic tracking devices occur is presented. This relational context is evident, albeit in truncated form, within the identification of interrelated vulnerabilities (eg, bodily and psychological) and within those websites using a human narrative as they often emphasized the very real relational and emotional conflicts that caregivers experience in wanting to balance the safety and autonomy of the person with dementia in their care. Furthermore, a move toward the relational dimension would greatly benefit the pursuit of person-centered care as it incorporates relational ethics within itself [67].

**Values and Electronic Tracking Device Design**

An additional insight pertains to the arc from identified vulnerabilities to care tools to a world aspired to. It is evident that the companies linked particular values to design decisions. For example, those companies putting forward electronic tracking devices that provided a medium level of information availability stressed the privacy of persons with dementia as a core value that should be placed above others such as constant information availability. Such decisions to place relative importance on certain values was found in the other levels as well. Companies with low level of information availability stressed reliability and dependability, and those with a high level of information availability stressed information, situational awareness, and control of care. In deliberately choosing to uphold certain normative values, the companies also made the decision (whether explicitly or implicitly) to not address or to de-emphasize other values. Thus, to uphold privacy requires a sacrifice in the amount of real-time information available to caregivers, and to uphold maximum information availability requires a sacrifice in the privacy of persons with dementia. Companies are choosing to give weight to certain values and link them to product design elements and, therefore, their devices are intrinsically value-laden artifacts.

The connection between design decision-making and particular values is not surprising given the abundance of literature focusing on the ethics of design. Indeed, a diversity of theories and methods has been put forward to infuse ethics into the design process—among others, value-sensitive design [28,68,69], responsible research and innovation [70,71], reflective design [30], constructive technology assessment [72], and technology mediation [73].

**How Do Values Emerge Within Electronic Tracking Devices**

Knowing that companies link values to design, an important question to ask is how, exactly, do these particular values come to be embraced? Some preliminary insights can be gleaned from electronic tracking device company websites. First, it is evident from those companies using a narrative of “co-creation” that stakeholders are involved in the process of development in some capacity. Second, engagement with academic research is also present. Most notably, company 2 had a web page with links to relevant academic publications, and company 12 repeatedly highlighted its collaboration with academic institutions (being a public-private endeavor). An additional indicator of academic engagement is the large overlap in company-identified vulnerabilities and values that are also present within academic literature. Company emphasis on caregiver conflict between balancing the aim to keep a person with dementia safe with the wish to maintain a person with dementia’s autonomy, for instance, reflects the central conflict most preoccupying the normative literature regarding electronic tracking devices in dementia care [18].

Although these early insights demonstrate that some companies engaged with stakeholders and the academic literature in the process of bringing their product to market, it is premature to draw too detailed conclusions regarding the link between value and design. The internal processes that gave rise to the company website content remain opaque and, as such, many questions are left unanswered. For example, it cannot be said that those companies that did not mention stakeholder involvement on their website excluded stakeholders entirely from the design process. Similarly, it is unknowable from the websites who made decisions or when, in the course of development, they were made (eg, were they a decision of design or post hoc marketing strategy?).

**Future Research Orientations**

However, what these insights do offer is a way to orient future research. It is clear that the companies identified a series of vulnerabilities facing persons with dementia and caregivers as a result of wandering and that, in crafting devices to address these vulnerabilities, they identified certain care tools and values deemed necessary to bring about a certain conception of life. Future research should seek to better understand both the internal
company processes that give rise to certain value decisions and the perceptions of electronic tracking device developers regarding the ethics of electronic tracking device design, development, and use. The former may uncover the extent to which stakeholders are involved in the development process, including who developers consider to be relevant stakeholders as well as the formal use of any approaches to design ethics, and the latter may increase understanding of how developers conceive autonomy, personhood, and the lived reality of persons with dementia—all important factors that contribute toward the construction of the envisaged user of their device (eg, whether idealized or realistic [74]) and the values developers design for.

Undertaking such research would reveal much about how companies navigate the ethical landscape that their products will operate within [73], ultimately contributing to greater harmony between industry, health care professionals, persons with dementia, and informal caregivers.

Strengths and Limitations
This study is the first to provide a nuanced analysis of electronic tracking device companies’ website content, showing how organizations responsible for the design and development of electronic tracking devices portray their products within the context of dementia-related wandering. The strength of these results is bolstered by the study’s strong methodological rigor. First, the process to identify websites used a particularly robust methodology, which is a multipronged strategy based on the PRISMA guidelines. Second, we used a method that preserved the websites in their entirety as they were at a particular time, mitigating the known data preservation problem associated with internet-based research [52]. Not only did this allow for a uniform and constant data collection process, but it also provided an opportunity to dive deeper into websites to ensure all relevant data were extracted without fear of data either disappearing or being altered. The richness of the data collected from this process benefited from reproducible procedures of the QUAGOL data analysis methods, which necessitated prolonged engagement with the data. From its initial full reading to the creation of individual schemes and then merged schemes, the QUAGOL requires a researcher to consistently engage with and to return to the original data. This data analysis process is traceable via an audit trail of conceptual schemes, researcher journals, and meeting notes (Multimedia Appendix 2). Third, data analysis and synthesis involved 3 researchers performing independent analysis at certain points with a high degree of consensus, strengthening the reflexivity and rigor of the results reported.

There are a few limitations to this study. First, only accepting English-language websites introduces a potential risk of bias and reduces potential transferability of the results. Second, because of the nature of the websites, the web pages included in this study may have since undergone significant changes or become defunct. Any in-depth research of websites will potentially be a step behind as the data are artificially frozen in time. Third, as indicated in the discussion, there is a limit to the complexity and depth of the results that can be derived from company websites.

Conclusions
Although persons with dementia are the focus of electronic tracking device use, they are not the focus of electronic tracking device company websites. Website material is more akin to a conversation between companies and caregivers about persons with dementia than with persons with dementia. This relegation of persons with dementia to a third party is an important conclusion, one that should serve not only as a sign that electronic tracking device companies should focus on developing materials centered on the relationship between caregivers and persons with dementia but perhaps also as a starting point for critical reflection on the enterprise that gives rise to these devices and the level of person with dementia involvement in their design, development, and use. In particular, serious reflection should be given to the question of how persons with dementia can be further involved in the use of electronic tracking devices when they are presently not a focus of marketing material. Indeed, a second important conclusion is the presence of three conceptual categories that form a triptych (ie, they are closely associated and better appreciated as a whole): identified vulnerabilities, care tools, and a world aspired to. As presented, vulnerabilities facing persons with dementia and caregivers are addressed through specific care tools by means of bringing a particular world to life after electronic tracking device implementation. These 3 concepts are further linked, in some fashion, to stakeholders; academic research; and, most vitally, values. What remains unknown is how these conceptions come to be. How are vulnerabilities identified? What process has led to the decision to center an electronic tracking device on a particular value such as privacy? By what means are stakeholders or the results of academic research included in this process? Future research should turn to understanding how the content of electronic tracking device company websites came to fruition. Doing so would provide a deeper foundation for the ethical evaluation of electronic tracking devices used in dementia care.

Acknowledgments
This study was supported by a grant from Research Foundation—Flanders (Fonds voor Wetenschappelijk Onderzoek—Vlaanderen; G0A5822N). The funder had no role in the study design; collection, analysis, and interpretation of the data; or writing of the paper.

Authors’ Contributions
JH and CG shared in the conception and design. JH completed data collection. JH, CG, and YD completed data analysis and interpretation, drafting, and editing. All authors approved the paper.

https://aging.jmir.org/2022/4/e38865
Conflicts of Interest

None declared.

Multimedia Appendix 1
Data collection form used to collect company, electronic tracking device, and website characteristics.

References


43. Russell DM. Advanced search operators. Google. 2019. URL: https://docs.google.com/document/d/1ydVaJlL1EYbWtlf9TPIBTE5IBADkQfZrQaBZxqXGs/edit [accessed 2019-10-01]
44. McKenna RA, Rollo ME, Skinner JA, Burrows TL. Food addiction support: website content analysis. JMIR Cardio 2018 Apr 24;2(1):e10 [FREE Full text] [doi: 10.2196/cardio.8718] [Medline: 31758778]


47. Topfer LA. GPS locator devices for people with dementia. In: CADTH Issues in Emerging Health Technologies. Ottawa, ON, Canada: Canadian Agency for Drugs and Technologies in Health; 2016:147.


Abbreviations

**PRISMA:** Preferred Reporting Items for Systematic Reviews and Meta-Analyses

**QUAGOL:** Qualitative Analysis Guide of Leuven

©Jared Howes, Yvonne Denier, Chris Gastmans. Originally published in JMIR Aging (https://aging.jmir.org), 11.11.2022. 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.
Review

Effectiveness of Technology Interventions in Addressing Social Isolation, Connectedness, and Loneliness in Older Adults: Systematic Umbrella Review

Eric Balki1, BSc, MA, MSc, PhD; Niall Hayes2, BSc, MSc, PhD; Carol Holland1, BSc, PhD

1Centre for Ageing Research, Faculty of Health and Medicine, Lancaster University, Lancaster, United Kingdom
2Nottingham Trent University, Nottingham, United Kingdom

Corresponding Author:
Eric Balki, BSc, MA, MSc, PhD
Centre for Ageing Research
Faculty of Health and Medicine
Lancaster University
Furness Building, Hazelrigg Ln,
Lancaster, LA1 4YG
United Kingdom
Phone: 44 1524 592127
Email: e.balkhi@lancaster.ac.uk

Abstract

Background: The global population of older adults (aged >60 years) is expected to triple to 2 billion by 2050. Proportionate rises in older adults affected by loneliness and social isolation (or social connectedness) are expected. Rapid deployability and social changes have increased the availability of technological devices, creating new opportunities for older adults.

Objective: This study aimed to identify, synthesize, and critically appraise the effectiveness of technology interventions improving social connectedness in older adults by assessing the quality of reviews, common observations, and derivable themes.

Methods: Following the guidelines of PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), 4 databases (PsycINFO, PubMed, Embase, and MEDLINE) were searched between February 2020 and March 2022. We identified reviews with adults aged ≥50 years in community and residential settings, reporting outcomes related to the impact of technologies on social disconnectedness with inclusion criteria based on the population, intervention, context, outcomes, and study schema—review-type articles (systematic, meta-analyses, integrative, and scoping)—and with digital interventions included. Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) was used to measure the strength of outcome recommendations including the risk of bias. The reviews covered 326 primary studies with 79,538 participants. Findings were extracted, synthesized, and organized according to emerging themes.

Results: Overall, 972 publications met the initial search criteria, and 24 met our inclusion criteria. Revised Assessment of Multiple Systematic Reviews was used to assess the quality of the analysis. Eligible reviews (3/24, 12%) were excluded because of their low Revised Assessment of Multiple Systematic Reviews scores (<22). The included reviews were dedicated to information and communications technology (ICT; 11/24, 46%), videoconferencing (4/24, 17%), computer or internet training (3/24, 12%), telecare (2/24, 8%), social networking sites (2/24, 8%), and robotics (2/27, 8%). Although technology was found to improve social connectedness, its effectiveness depended on study design and is improved by shorter durations, longer training times, and the facilitation of existing relationships. ICT and videoconferencing showed the best results, followed by computer training. Social networking sites achieved mixed results. Robotics and augmented reality showed promising results but lacked sufficient data for informed conclusions. The overall quality of the studies based on GRADE was medium low to very low.

Conclusions: Technology interventions can improve social connectedness in older adults. The specific effectiveness rates favor ICT and videoconferencing, but with limited evidence, as indicated by low GRADE ratings. Future intervention and study design guidelines should carefully assess the methodological quality of studies and the overall certainty of specific outcome measures. The lack of randomized controlled trials in underlying primary studies (<28%) and suboptimal methodologies limited our findings. Robotics and augmented or virtual reality warrant further research. Low GRADE scores highlight the need for high-quality research in these areas.

Trial Registration: PROSPERO CRD42022363475; https://tinyurl.com/mdd6zds

https://aging.jmir.org/2022/4/e40125

JMIR Aging 2022 | vol. 5 | iss. 4 | e40125 | p.127
(page number not for citation purposes)
**Introduction**

**Background**

The use of technology to support older adults against feelings of loneliness and social isolation provides novel opportunities that have grown in the field of aging, as technology demonstrates that information and communications technology (ICT) use and training [1] and robotics conflate in the provision of programs and activities to facilitate social connectedness.

Social isolation and loneliness in older adults have been extensively researched. Many studies showed that the prevalence of these problems increases with age. For example, the prevalence of loneliness among young adults, early to middle-aged adults, and late to middle-aged older adults are 39.7%, 43.3%, and 48.2%, respectively [2]. The current global population of people aged ≥60 years is expected to triple to 2 billion by 2050 [3]. The number of people aged >50 years experiencing loneliness is expected to reach 2 million by 2025-2026, a 49% increase in 10 years [1]. Loneliness and social isolation are different concepts but are interlinked and can be considered the constructs of social disconnectedness [4]. Social isolation is objectively defined as the deprivation of relationships and social interactions, whereas loneliness is a subjective sense of not meeting one’s social needs [5]. Socially disconnected individuals are vulnerable to social isolation and loneliness because they have small social networks and low participation rates in social activities [6]. Fafchamps and Shilpi [7] defined social isolation as “deprivation of social connectedness and an inadequate quality and quantity of social relations at different levels of interactions (individual, group, community and broader social environment)” [6].

Socially disconnected older adults are also vulnerable to a range of health disorders, including infection [8], high blood pressure [9], impaired cognitive function [10], depression [11], stress associated elevation of hypothalamic-pituitary-adrenocortical activity [12], cardiovascular disease [13], diminished immunity [14], and mortality [15]. In addition, loneliness elevates the risk of dementia [16] and accelerates the progression of Alzheimer disease [10]. As the population proportion of older adults increases, negative health outcomes are expected to rise along with social isolation, and loneliness is likely to increase along with negative health outcomes [17].

Rapidly deployable technologies, along with socioeconomic changes that have reduced the cost of technology, have increased the accessibility of technological devices, creating new opportunities for older adults [18]. Internet-based technology interventions for social disconnectedness have grown over the past decade [19]. Digital communication technologies can improve the lives of older adults by facilitating their social relationships. Technologies such as email, social networking sites (SNSs), videoconferencing, and mobile instant messaging (MIM) apps have been shown to improve self-rated health and lower the incidence of loneliness, chronic illnesses, and depressive symptoms in older adults [20]. They also supplement the social benefits of physical interactions by reinforcing existing connections or providing routes to new connections, further reducing loneliness levels. Frequent users of technology and the internet can also access health information and social support for psychosocial problems. However, many studies on technology intervention ignore confounding factors, such as age, gender, living arrangements, economic status, education level, cognitive status, and daily living activities [21,22], which may influence the effectiveness of the intervention and the robustness of the findings. The small number of high-quality studies in this arena limits the generalizability of the results.

Several reviews have summarized works on technology interventions for older adults experiencing loneliness [23,24], but their value is diminished by the plethora of unclear evidence, heterogeneity of both populations, measures and methodologies, diverse outcomes, scattered focus, and broad topics. As the existing reviews are heterogeneous in content, lacking the investigation of outcome measures used and discussions on causation, they cannot reach generalizable conclusions.

For a standardized systematic report on these reviews, we must assess the quality of the reviews and find common observations and derivable themes. An umbrella review method can provide a focus for areas where there are competing interventions and amalgamate evidence from multiple quantitative and qualitative reviews [25]. To our knowledge, an umbrella review exploring the types and effectiveness of intervention technologies for social connectedness has not been published.

**Aims**

To bridge this gap in the literature, we aimed to explore the findings and limits of current knowledge on the impact of technology interventions on social disconnectedness in older adults. We also emphasize areas requiring further research. In a comprehensive umbrella review, we synthesized the various categories and types of the used technology interventions, discussed their effectiveness and limitations, and finally explored their potential and need for further research. Finally, we amalgamated all the evidence from the umbrella review and used Grading of Recommendations, Assessment, Development, and Evaluations (GRADE) to make recommendations for interventions targeting social connectedness. This review attempts to answer the following questions:

1. What technology interventions are used to influence social connectedness in older adults?
2. How effective are these technology interventions in improving social connectedness in older adults, and what aspects make them effective?
Methods

This umbrella review followed the standardized procedures [12,26,27] of systematic reviews. The protocol followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) systematic review protocol guidelines [28] and the Joanna Briggs Institute (JBI) methodology for umbrella reviews [12].

Search Strategy

The search strategy involved controlled vocabulary searching; phrase searching; and applying Boolean logic, limits, and filters. A comprehensive systematic search of 4 databases (PsycINFO, PubMed, Embase, and MEDLINE) was conducted between February 2020 and March 2022. The reference lists were also examined for additional reviews. The following search terms were used: “ageing,” “aging,” “older adults,” “reviews,” “2000-22,” and synonyms for “social isolation and loneliness,” “social connectedness,” and “technology interventions.” As an example, Textbox 1 shows the search terms and search strategy applied to the PubMed database. Search terms can be found in Multimedia Appendix 1.


<table>
<thead>
<tr>
<th>Search terms used</th>
<th>Search strategy applied</th>
</tr>
</thead>
<tbody>
<tr>
<td>SU (technology or computer or Internet) and TI (review or meta-analysis or metasynthesis) and SU (older OR aging OR aged OR elderly OR senior) and (social isolation OR loneliness OR social connectedness)</td>
<td>Limiters—Published Date: 200001-2021;211231; Language: English; Publication Type: Academic Journal; English Language; Year of Publication: 2000-22; Publication Year: 2000-22; Publication Type: Peer Reviewed Journal; English Language; Exclude Dissertations Search modes—Boolean or Phrase Sort by best Match</td>
</tr>
</tbody>
</table>

Inclusion and Exclusion Criteria

The inclusion criteria were formulated using the population, intervention, comparison or context, outcomes, and study schema [29,30]. Table 1 describes the inclusion criteria under which the studies were selected for this review.

Table 1. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>PICOS framework</th>
<th>Inclusion criteria and reasons</th>
<th>Exclusion criteria and reasons</th>
</tr>
</thead>
<tbody>
<tr>
<td>P—participants</td>
<td>Persons aged &gt;50 years, who are living in community or residential settings with no major neurocognitive impairments</td>
<td>Participants aged &lt;50 years</td>
</tr>
<tr>
<td>I—interventions</td>
<td>Interventions using any form of information and communications technology, smart communication devices, internet-based communication systems, information systems, video games, technological devices, and robots or technological pathways allowing for social interaction. These interventions must be specifically targeted at impacting or improving social connectedness in older adults</td>
<td>Nontechnology interventions, smart devices for home, or telehealth technologies not designed to impact social connectedness (eg, diabetes-measuring devices)</td>
</tr>
<tr>
<td>C—context</td>
<td>Community settings, independent living, and participants in nursing and care homes</td>
<td>Hospital settings, mental and physical illnesses, and disease or illness-specific cases</td>
</tr>
<tr>
<td>O—outcomes</td>
<td>Quantitative or qualitative outcome data or results focusing on social isolation or loneliness or social connectedness</td>
<td>Reviews lacking descriptions of outcome data</td>
</tr>
<tr>
<td>S—study</td>
<td>Review articles of any type using a systematic, qualitative, or quantitative method, including narrative, quantitative, and qualitative comparative studies. Articles must describe a clear intervention and include qualitative and quantitative comparative studies</td>
<td>Reviews with no technology intervention, no clear outcomes, or no systematic review processes. Reviews earlier than 2005 were not included because technology interventions before this time would not be directly comparable with ones of the present day</td>
</tr>
</tbody>
</table>

Selection Process

The abstracts and titles of all potentially relevant articles were screened. Full texts were then evaluated, and duplicates were removed. Uncertainties were discussed among the research team members to reach a consensus. Relevant data of the included articles were summarized in tables and checked for accuracy by a second investigator (CH).

Analysis

The data analysis was based on a thematic synthesis with an inductive, iterative process consisting of 3 main stages: (1) free line-by-line review of the results, synthesis tables, and discussion sections of the included papers; (2) organization of themes into related areas; and (3) the identification, development, and refinement of detailed descriptions of factors that impacted the effectiveness of technology interventions [31]. All measures used were specified, and the statistical results (if
provided) were summarized. The technology types were listed along with their effectiveness, and the authors’ conclusions were also summarized.

**Quality Assessment**

The methodological qualities of the reviews were assessed using the Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) [32] quality rating tool for reviews. The 11-item R-AMSTAR includes 11 questions (Multimedia Appendix 2 [19,20,23,24,33-52]) whose scores are summed to give the overall quality score of a systematic review. The R-AMSTAR tool provides a quantifiable assessment of systematic reviews and a measurement of their methodological quality. The maximum possible score is 44.15. Any review scoring <22 was excluded as it lacked 1 or more critical R-AMSTAR definitions. For example, the review might not assess the scientific quality of the studies or might apply a poor method for combining study findings [53].

**Grading of Evidence**

The overall certainty of the evidence was evaluated using the GRADE method, which analyzes the risk of bias (imprecision, inconsistency, indirectness, and publication bias) and assesses the quality of the included evidence, which we used to make recommendations [54]. Initially, we categorized the evidence based on the inclusion or exclusion of randomized controlled trials (RCTs), followed by the inclusion or exclusion of observational studies. We then considered whether the studies had serious limitations or important inconsistencies in the results, or whether uncertainty about the validity of the evidence (the extent to which the participants, interventions, and outcome measures are similar to those of interest) was warranted (Multimedia Appendix 2). Limitations in study quality found in the R-AMSTAR appraisal, important inconsistency of results, or uncertainty about the directness of the evidence lowered the grade of evidence. For instance, if all available studies have serious limitations, the grade will drop by a level, and if all studies have very serious limitations, the grade will drop by 2 levels. The quality of evidence is also reduced by imprecise or sparse data and an imprecise understanding of social concepts.

**Results**

**Overview**

The article elimination process is summarized as a flowchart in Figure 1. The initial search extracted 972 publications. Further, 91 articles were identified after checking the reference lists. After excluding duplicates and irrelevant publications, articles were screened using the population, intervention, comparison or context, outcomes, and study schema inclusion criteria (Table 1). The commonest reasons for exclusion were interventions targeted at specific mental and physical illnesses (138/972, 14.2%) and interventions not matching the prespecified definition (95/972, 9.8%). A total of 90 full-text reviews were further passed through a 3-step screening process (title, abstract, and full-text based) for eligibility and inclusion in the qualitative synthesis of this review. Finally, 24 reviews based on technology interventions were eligible for the synthesis.
Quality Assessment
Among the 24 selected articles, 3 (12%) articles with R-AMSTAR scores <22 were excluded because they failed a priori systematic review processes (lacked clarity in scope or purpose, had a priori–defined participant population, had unclear outcomes of interest, lacked clarity on interventions, involved nonspecific subgroup analyses, and lacked meaningful hypotheses). The 21 remaining reviews were of moderate quality, with none meeting all of the R-AMSTAR criteria.

Data Extraction
Data from the 21 reviews were extracted using a piloted, standardized data extraction form that captures and summarizes findings. As both technology interventions and extracted outcome data were heterogeneous, they were deemed inappropriate for a quantitative synthesis using meta-analytic techniques. Instead, a narrative synthesis summarizing the effectiveness of interventions was implemented. Under the methodological considerations of umbrella reviews, the results were reported descriptively in tabular form (Multimedia Appendix 3 [19,20,23,24,33-49]) along with their associated characteristics (number of articles, databases used, participants, types of interventions, study design, measures used, summary of results, authors’ conclusions, and summary and review methods). Multimedia Appendix 3 provides details of the 21 reviews in this study.

Study Characteristics
The 21 selected reviews included 16 (76%) systematic reviews (reviews of evidence on a clearly formulated question and the use of systematic and explicit methods to identify, select, and critically appraise the relevant primary research), 2 (10%) integrative reviews (reviews that summarize past empirical or theoretical literature to provide a comprehensive understanding), 2 (10%) scoping reviews (preliminary assessments of the potential size and scope of the available research literature), and 2 (10%) meta-analyses (statistical analyses combining the results of multiple scientific studies). Most of the reviews covered the beneficial impact of technologies on loneliness, whereas others focused on social isolation, connectedness, and quality of life. General ICT was the most commonly applied intervention technology. The publication period was from 2005 to 2022, but 19 of the selected reviews were published within the last 7 years. Of the 21 reviews, 1 (5%) review focused on assistive technology for communication. Overall, 19% (4/21) of reviews focused on general interventions for social connectedness but examined technologies such as general ICT and videoconferencing, and 10% (2/21) of reviews focused on communication technologies for social connectedness in older adults. In all, 38% (8/21) of reviews investigated the impact of general internet and computer technologies on social isolation and loneliness. Of 21 reviews, 1 (5%) review examined the impact of smart technologies on social connectedness, and another (1/21, 5%) study reported the impact of health promotion technologies on social isolation and loneliness.
all, 10% (2/21) of reviews explored the ability of general ICT to improve the quality of life. Of 21 reviews, 1 (5%) review examined interventions to reduce social isolation and loneliness during the COVID-19 pandemic; 2 (10%) reviews focused on the impact of SNS on loneliness, another (1/21, 5%) examined interventions for preventing loneliness in nursing homes, and another (1/21, 5%) evaluated the benefits of telehealth in alleviating loneliness and increasing medication compliance. Here, telehealth was implemented through video health care professional visits to older adults. The 21 reviews covered a total of 326 underlying primary studies on technology interventions. It is worth pointing out that we were not able to confirm the presence of gray literature or studies that looked at technology interventions in the reviews.

The interventions discussed in the reviews were general ICT (11/21, 52%), videoconferencing (4/21, 19%), computer and internet training (3/21, 14%), telecare (2/21, 10%), SNS (2/21, 10%), and robotics (2/21, 10%). The reviews reported mixed results. Positive effects of ICT on loneliness were the most commonly reported, followed by the positive impacts of ICT on social isolation or connectedness. Reviewing data from the underlying primary studies in the reviews, the most effective intervention mode for social connectedness was identified as general ICT, followed by videoconferencing and robotics (Table 2).

### Table 2. Effectiveness versus ineffectiveness of different intervention modes on social connectedness, identified in the underlying primary studies of the review papers (n=321).

<table>
<thead>
<tr>
<th>Study intervention</th>
<th>Effective, n (%)</th>
<th>Ineffective, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D or augmented reality (N=1)</td>
<td>1 (100)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Video gaming (N=4)</td>
<td>3 (75)</td>
<td>1 (25)</td>
</tr>
<tr>
<td>Videoconferencing (N=14)</td>
<td>11 (78)</td>
<td>3 (22)</td>
</tr>
<tr>
<td>Robotics (N=22)</td>
<td>16 (73)</td>
<td>6 (27)</td>
</tr>
<tr>
<td>Telecare (N=34)</td>
<td>22 (65)</td>
<td>12 (35)</td>
</tr>
<tr>
<td>SNS(^a) (N=61)</td>
<td>30 (49)</td>
<td>31 (51)</td>
</tr>
<tr>
<td>Computer training (N=66)</td>
<td>39 (59)</td>
<td>27 (41)</td>
</tr>
<tr>
<td>ICT(^b) (N=119)</td>
<td>86 (72)</td>
<td>33 (28)</td>
</tr>
</tbody>
</table>

\(^a\)SNS: social networking site.

\(^b\)ICT: information and communications technology.

### Results From Systematic Reviews With Meta-analyses

Among the 21 selected reviews, only Choi et al [20] and Bornemann [33] performed meta-analyses of homogenous data (Multimedia Appendix 3). Choi et al [20] reported a significant pooled decrease in loneliness after implementing technology interventions (Z=2.085; P=.04). However, Bornemann [33] concluded a nonsignificant decrease in loneliness after reviewing 5 out of 7 studies included in the review by Choi et al [20] (Z=.44; P=.37)—that is, the same 5 studies yielded different pooled meta-analysis results in the 2 reviews. This divergence indicates potential biases in the analytic approaches; for instance, Bornemann [33] excluded some studies included in Choi et al [20], and some of their findings were inconsistent with the narrative conclusions of their included studies. Bornemann [33] questioned the validity of some of the data acquired by Choi et al [20]. Although this review does not cross-examine these findings, we clarified that a study included in Choi et al [20] should have been excluded, as it was not an ICT intervention study. We decided that although the statistical conclusions of Bornemann [33] were correct, Choi et al [20] raised some valid points. Multimedia Appendix 4 gives the levels of certainty in the quality assessment of outcomes developed within the GRADE framework. Low-quality assessments in different categories are mainly attributable to the elements of the study design, poor study quality, inconsistency, and indirectness.

### Categories of Technology Interventions

Of the 21 studied reviews, 14 (67%) dealt with general ICT (which was a catch-all term defining a diverse set of technological tools and resources used to transmit, store, create, share, or exchange information), 4 (19%) with videoconferencing, 3 (14%) with computer and internet training, 2 (10%) with telecare, 2 (10%) with robotics, 2 (10%) with SNS, 3 (14%) with gaming, and 1 (5%) with 3D augmented reality (AR). Among the primary studies, general ICTs were the most commonly adopted interventions (with 119 studies), followed by computer training, SNS, telecare, and robotics (Table 3). Although some of these categories overlapped, we differentiated them as they were distinguished in the original reviews.
Table 3. Frequencies of intervention categories in the primary studies (N=321).

<table>
<thead>
<tr>
<th>Intervention Category</th>
<th>Frequency, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D or augmented reality</td>
<td>1 (0.3)</td>
</tr>
<tr>
<td>Video gaming</td>
<td>4 (1.2)</td>
</tr>
<tr>
<td>Videoconferencing</td>
<td>14 (4.4)</td>
</tr>
<tr>
<td>Robotics</td>
<td>22 (6.9)</td>
</tr>
<tr>
<td>Telecare</td>
<td>34 (10.6)</td>
</tr>
<tr>
<td>SNS&lt;sup&gt;a&lt;/sup&gt;</td>
<td>61 (19)</td>
</tr>
<tr>
<td>Computer training</td>
<td>66 (20.6)</td>
</tr>
<tr>
<td>ICT&lt;sup&gt;b&lt;/sup&gt;</td>
<td>119 (37.1)</td>
</tr>
</tbody>
</table>

<sup>a</sup>SNS: social networking site.

<sup>b</sup>ICT: information and communications technology.

Outcome Measures Used

All the reviews reported large numbers and diverse outcome measures of primary studies. Besides constructs of social disconnectedness (loneliness, social support, social contact, number of confidants, social networks, social connectedness scales, social isolation, and social well-being), many studies assessed factors such as quality of life, self-esteem, stress, and depression. Although not directly related to social disconnectedness, these factors may affect or be affected by social disconnectedness and may be useful to include outcome measures alongside social connectedness. A minority of the reviews also reported outcome measures of empowerment.

When analyzing these quantitative primary studies, the reviews commonly applied validated tools, such as the University of California Los Angeles (UCLA) Loneliness Scale (or a modified version) and the De Jong Gierveld Scale [4]. The UCLA was the most tested dependent variable. Among various other measures were the Social Support Scale by Schuster and Hunter [34], Social and Emotional Loneliness Scale [55], and Multidimensional Scale of Perceived Social Support by Zimet et al [56]. Social connectedness was sometimes measured using the holistic Social Connectedness Scale by Lee and Robin [57], which is regarded as a comparatively reliable measure.

The definitions and uses of outcome measures differed across the reviews. A total of 62 outcome indicators of social connectedness were used in the primary studies. Most reviews did not report on the lack of intervention effects (including the absence of significance values); moreover, the primary studies adopted a mixture of validated and nonvalidated outcome measures, making such reporting difficult. Consequently, they could not conclude whether the primary studies had validatable statistically significant outcomes.

Social Concepts Used

The social concepts used for determining outcomes varied in range and diversity. In many reviews, the source papers did not define social participation or social isolation but instead evaluated these factors as general or neighboring concepts [19,35-37]. Loneliness was evaluated more consistently than social participation and social isolation but was sometimes incorrectly interchanged with social isolation. Most studies assessed loneliness on standardized scales, notably the UCLA Loneliness Scale [35,36,38,39].

A few of the reviews highlighted that inconsistency and lack of specific definitions hindered the grouping and evaluation of their chosen papers [19,37,39]. Morris et al [39] described social connectivity as a multidimensional concept that is difficult to define, conceptualize, and measure. They elaborated that outcome measures, such as the UCLA Loneliness Scale and Perceived Social Support Scale, identify only single aspects of social connectedness.

Cattan et al [37] also noted a complex association among social isolation, loneliness, and living alone, which was difficult to describe in their reviewed studies. Rarely among the review studies, Cattan et al [37] attempted to distinguish living alone from social disconnectedness and suggested that living alone be measured independently as a concept of physical isolation.

Ibarra et al [40] correctly defined loneliness as “a subjective measure referring to the ‘unpleasant’ lack of and quality of social relationships.” By contrast, isolation is an objective measure referring to few or no social relationships, although their study clarified the difference between social isolation and loneliness.

Gardiner et al [36] and Williams et al [38] adopted the less frequently used concept of social facilitation for creating mechanisms through which older adults can interact with peers. From an alternative perspective, they measured the facilitation of social connections. The article by Williams et al [38] was especially relevant, as it examined interventions during the COVID-19 pandemic. Facilitation may lead to effective interventions that reduce social isolation and loneliness, without violating COVID-19 shielding and social distancing measures.

In conclusion, different definitions and measurements of loneliness, social isolation, and social connectedness have led to diverse findings and wide variations across and within disciplines, defying a coherent picture of the research. Although some of the more recent studies and reviews have addressed this heterogeneity, reliable and succinct findings will remain elusive without further investigations.
Group Interventions Versus One-to-One

Many interventions implemented in the individual papers of the reviews were broadly divisible into group and one-to-one interventions. In general, group interventions were more frequently implemented than one-to-one interventions, although both types were effective [24,37,40,41]. Cattan et al [37], who reviewed 3 computer group interventions, reported that group interventions with educational and social activities are particularly effective.

The imbalance between the group and one-to-one interventions impairs comparisons between the 2 types and conclusions regarding their comparative successes. Nevertheless, some of the reviews pointed out the possible advantages and limitations of these intervention approaches. Poscia et al [41] noted that group interventions might beneficially create a sense of security and belonging, although the real effect of the intervention might be obscured by interactions among the group members. Individual interventions might create deeper, more personal bonds and boost confidence in social engagements. Ibarra et al [40] further observed that one-to-one interactions limited participants’ contact with family, friends, and acquaintances, whereas group interventions encouraged them to interact with new people and potentially expand their networks, thereby increasing their number of new social connections.

Overall, group interventions appear to improve social disconnectedness, but the insufficient number of one-to-one interventions prevents an objective comparison and firm conclusions of the best interaction type. However, the GRADE assessment of the quality of evidence suggested a very low advantage of group interventions over one-to-one interventions (Multimedia Appendix 4).

Effectiveness of Technology Interventions as an Overarching Category

Technology interventions that enhance social connectedness include general ICT, video games, robotics, and the Personal Reminder Information Social Management system (a custom-designed experimental SNS for older adults). Less conclusive evidence exists for the beneficial effects of SNS [20,24,25,37,41,42].

Overall, technologies appear to positively affect loneliness, social isolation, and other psychosocial aspects of older adults’ lives. Khosravi et al [42] examined 8 technology types and found that most technologies, in some formats, can increase social connectedness in older adults.

When technologies were intended to strengthen existing connections, their positive impacts on loneliness and social isolation were more consistent [24,40,41]. Ibarra et al [40] found that technologies are fundamental to long-distance interaction and are thereby necessary for expanding social networks, improving existing ties, and increasing social connectedness. However, they noted that how technology is availed, the limitations and opportunities of technology, and their effects on the success of the intervention are all unclear. Some reviews [20,35,43,44] included a psychosocial outcome of interest, such as social isolation, life satisfaction, loneliness, or depression. It was found that interventions significantly reduce loneliness but are ineffective against depression [35,43,45]. Damant et al [45] found a significant correlation between internet use and depression, suggesting that although the literature reports a significant correlation between loneliness and depression, technology can exert divergent impacts on these 2 psychosocial variables. However, Khosravi and Ghapanchi [43] reported that technology interventions can potentially reduce depression through engagement in social interaction, hinting that social isolation impacts more strongly on depression than does technology.

Choi and Lee [58] presented a detailed statistical evaluation of 8 RCT studies investigating the impacts of various technology interventions on loneliness. They found a statistically significant decrease in loneliness in the intervention group compared with the control and usual care groups (P<.07 and P<.001, respectively). However, there were no statistically significant differences in loneliness among the members of the intervention groups before and after the intervention (P>.05).

Individual reviews reported less conclusive outcomes of the overall technology use. The results of Morris et al [39] ranged from positive to no impact on loneliness, and Damant et al [45] noted a negative association between “social involvement and participation” and older adults’ use of technology, thereby indicating that the more socially involved people were, the less they tended to use technology. They found that high internet use was associated with high levels of loneliness. Interestingly, Chen and Schulz [35] found a positive effect of technology on social connectedness, this impact usually diminished in studies spanning >6 months. The time frame of studies investigating the effectiveness of technology was also a recurrent theme in other studies. The diminished effect is potentially linked to fatigue from using the intervention or inconsistency in the study approach over time.

Specifically, the following technology interventions appear to reduce social isolation but lack rigorous statistical support for a positive effect: robotics, telecare, and SNS [34,36,42,45]. Overall, 86% (18/21) of reviews examined the impact of technology intervention on loneliness. The reviews covered 324 primary studies involving 66,565 participants. Of the 18 reviews, 15 (83%) reported a positive effect of technology on loneliness; the remaining 3 (16%) studies found a 0 or negative effect. From the reviews, it can be concluded that technology interventions exert an overall positive influence on social isolation and loneliness (social disconnectedness), but their effectiveness depends on the design of the study. Longer training times, shorter study durations, and facilitation of existing relationships tended to increase the effectiveness of the intervention. The quality of evidence supporting the effectiveness of technology interventions on social connectedness (GRADE assessment) was moderate to low.

General ICT

This section explores the findings of general ICT interventions reported in the reviews. General ICT is an umbrella term for generic technology devices, services, applications, and internet platforms [59]. ICT includes internet-based networks, mobile phones, computers, tablets, and any software requiring an
internet connection. Interventions in this category include interactions via internet use (eg, discussions and forums), emails, video chats and conferencing, SNS, virtual spaces, classrooms, and messaging services. Some reviews mentioned systems tailored for older adults, such as the customized touch screen video-chat system described by Ibarra et al [40]. Computers with a mouse and keyboard as input devices were preferred, closely followed by tablets and mobile phones (the latter appeared as the most popular device in recent reviews). Other interventions used customized television sets and touch screen computers. Khosravi et al [42] and Khosravi and Ghapanchi [43] reported studies on Personal Reminder Information Social Management (a customized social networking platform). In most of the reviews, general ICTs were regarded as a single category, although videoconferencing and SNS were often placed in separate subcategories.

Many of the reviewed studies found that ICT interventions not only significantly reduce loneliness but also exert a positive impact on other aspects of social isolation, providing social support and connectedness, communication with family and friends, and ICT-accessible information sources [19,20,35,42,43]. Some reviews hinted that ICT facilitates the acquisition of information through the internet, either through interactions with other people or through finding relevant information on the web, which helps reduce loneliness [35,38,60]. Indeed, Morris et al [39] found that social connectedness especially benefits from technologies with web-based programs incorporating items such as health information, support groups, chat rooms, or discussion boards.

Damant et al [45] alone reported on studies with less promising results. In a study, only a small number of older adults maintained contact with their families via the internet. These participants were reluctant users with the sole purpose of keeping in touch with their grandchildren. In another study, they found no significant correlation between internet or email use and contact with family and other people. Both studies revealed no significant correlation between computer use or training and loneliness. Some of the studies reviewed by Damant et al [45] reported exacerbated loneliness through ICT use. It appears that ICT can positively reinforce existing social networks but has a limited impact on building new ones.

Only 2 reviews provided a homogenous meta-analysis. Both reviews reported positive impacts of general ICTs on social disconnectedness. In total, these reviews included 119 primary studies: 86 reporting a positive impact on social isolation or loneliness and 33 reporting unclear or no impact. The studies agreed that increasing the frequency of general ICT use enhances social connectedness, improving the ease with which older adults can interact and maintain contact with others, thus reinforcing social connections with friends and family. The evidence that frequent ICT use facilitates the creation of new relationships or contacts is much weaker, further supporting, in part, the conclusions of Damant et al [45].

Together, these results suggest that general ICT can facilitate established connections and might supplement or replace older communication methods. Its role in establishing new connections is uncertain. Our results suggest that when considering ICT interventions (at least for older adults), it is important to distinguish between their ability to maintain relationships, potential ability to deepen relationships, and inability to help create new relationships. The GRADE strength of the ICT category, although only moderate, was the highest among the categories because a large number of primary studies, including RCTs, were reviewed in this category, and there was consensus and clarity on the outcome measures.

Social Networking Sites

Although SNS is a subcategory of ICT, it warrants its own heading because 33% (7/21) of reviews discussed separate finding on SNS. The reviews gave mixed results. Whereas some studies supported the use of SNS in reducing loneliness, a sizable number showed no impact or even an increase in loneliness after SNS use [19,42,46]. Both Chen and Schultz [35] and Wiwatkonupakarn et al [46], who reviewed high-quality RCT studies on the use of SNS, reported inconclusive impacts of SNS on loneliness. They found some support for sites such as Facebook, which provides games that can be played with others over a network, thus fostering social interaction and alleviating loneliness. The mixed findings in these reviews might be explained as follows: although older adults embraced the use of SNS to support their social relationships and help them overcome loneliness, they did not regard these sites as a replacement for face-to-face contact. Participants preferred to use SNS for searching for and disseminating information rather than socializing. Morris et al [39] reported positive effects of smart technologies similar to SNS, especially when they incorporated health information, support groups, chat rooms, or discussion boards. Their findings support a role of SNS in knowledge-seeking and support-acquisition scenarios, with consequent impact on loneliness.

These findings may partly depend on the type of SNS, as different types of SNS support different features. For example, Facebook may promote socialization more effectively than YouTube, whereas YouTube may better facilitate explicit knowledge acquisition and information transfer than Facebook. Ibarra et al [40] discovered that participants favored off-the-shelf solutions, such as Facebook and About-My-Age (an SNS for older adults). Users of these sites commented on their decreased loneliness and easy control of the sites. The sheer volume of users on these platforms might assist older adults in finding relevant information, including information on how to use the platforms, thus creating a positive feedback loop.

On the downside, SNS use raises several concerns: privacy, lack of perceived usefulness, and possibly demographic factors [19,47]. Newman et al [47] noted an interesting connection between educational attainment and SNS use: SNS users tended to be White, employed, educated, and married. They also found attitude differences toward technology use among sociodemographic groups based on gender (women) and age (older people).

Overall, 61 primary studies examining SNS were found in the reviews: 31 reporting positive impacts of SNS on social isolation and loneliness and 30 reporting unclear or no impacts of SNS. Therefore, the effectiveness of SNS is inconclusive. The results suggest that older users can obtain support, acquire knowledge,
and maintain their existing relationships through SNS. In terms of combating social disconnectedness and establishing new relationships, SNSSs are less effective and can be detrimental at times. However, the effectiveness of SNS in developing new relationships, fostering and maintaining existing ones, and acquiring knowledge and support has not been explored in depth, and the idiosyncrasies of SNSSs must be unraveled in further research. The strength of evidence (GRADE assessment) of the reviews in this category is low because of indirectness, missing information, and publication bias.

**Videoconferencing**

Overall, videoconferencing appeared to exert a positive impact on loneliness and social connectedness. The visual aspect of this intervention seemed especially appealing to older adults [24,34-36,38,40,44,49]. In total, 3 reviews reported on videoconferencing between family members and their established contacts. All reviews described a statistically significant reduction in loneliness [39,41,45]; however, videoconferencing was more effective in facilitating established connections than in building new ones. Moreover, videoconferencing showed a weak impact on information gathering. For instance, Chen and Schultz [35] reported that videoconferencing did not significantly provide informational support (information communication for problem-solving assistance) or instrumental support (tangible goods, services, and aid), which may improve social connectedness [35]. Ibarra et al [40] mentioned 1 study in which Skype used for educational purposes did not change participants’ loneliness levels and another study in which Skype combined with computer training better reduced loneliness levels than did Skype alone. These reviews suggest that videoconferencing is effective for maintaining established connections, such as those with family members, but is less effective for other purposes, such as education and information seeking, which may indirectly impact social connectedness.

Gardiner et al [36] and Ibarra et al [40] mentioned the importance of appropriate hardware and design in videoconferencing. They reported that technical, financial, and design issues are potential barriers to the wider uptake of this technology.

When used in health support, videoconferencing yields mixed results. The intervention often decreases the loneliness and social isolation of residents in care and nursing homes, but a few studies have found no difference from the baseline [34,35,43]. More clearly, participants in these settings benefit from videoconferencing contact between family and friends, with beneficial effects on loneliness. Interestingly, Husebo and Storm [48] found that virtual visits by clinicians reduced the social isolation of residents in care homes, suggesting that videoconferencing can enhance the perception of independence by providing easy access to services. In general, videoconferencing appears to reduce loneliness in residential, nursing, and clinical care settings, although the specific aspects of the intervention that ensure its success have not been elucidated.

Overall, 14 primary studies in this subcategory were found in the reviews. Of these studies, 11 reported a positive impact on social isolation or loneliness. Owing to reviews such as by Schuster and Hunter [34], with clear outcomes and the inclusion of RCTs, the GRADE strength of evidence in this subcategory was moderate to low. The use of standardized outcome measures would have strengthened the GRADE rating.

**Mobile and Instant Messaging**

Among the studied reviews, only Ibarra et al [40] alone described studies on MIMs such as WhatsApp and Line (messaging services). In 1 study, WhatsApp was used more extensively than email by relatives; however, a lack of responses can increase the perception of loneliness. Ibarra et al [40] hinted that as WhatsApp and similar applications are easy to use and allow the sharing of pictures, they exert a positive impact on social disconnectedness. However, the evidence was insufficient for concluding the impact of MIMs on social connectedness and loneliness; moreover, the few primary studies suggest that MIM explorations are only emerging at this stage. Given the lack of information found in the reviews, the GRADE strength of the evidence in this category was very low.

**Computer and Internet Training**

In total, 13 reviews evaluated the impact of computer and internet training on various guises. All reviews found a positive impact of these interventions on social connectedness and loneliness [20,24,36,39,41,43,45]. In 4 of these reviews, loneliness reduction was found by the authors to be statistically significant [39,41,45]. However, all these studies investigated group training, suggesting that positive impacts were contributed (at least partly) by interaction with others in the group. Indeed, Damant et al [45] found a study in which group training increased the perceived support of friends and another study in which loneliness levels were reduced when email and web-based forums formed part of the training regime.

Mixed results were also obtained for this category. Baker et al [19] reviewed 2 studies on ICT training. 1 finding no correlation between the training and social connection and the other concluding that ICT training can enhance social networks. Although the authors did not elaborate on this discrepancy, the very different time frames of the 2 studies (12 months vs 8 weeks) may have affected the results. Indeed, whenever mixed results were found, the training time appeared to be a contributing factor, with shorter training times more likely to yield inconclusive results [24,36,42]. Furthermore, Choi and Lee [58] reported that in most studies, older adults enjoyed using technology and significantly increased their frequency of use, suggesting that minimal training was required.

Unusually, among the reviews, Williams et al [38] found that overall computer training produced no effect on social isolation. Overall, ICT training showed a higher ability to reduce loneliness in longer-duration studies than in shorter-duration studies.

As some reviews did not differentiate between the impacts of training and subsequent use, any assumptions would be dubious. Morris et al [39] noted a combined result, in which interactive web-based programs, discussion forums, and training mainly enhanced social connectedness; only 1 study reported inconclusive results. The effect of training was often confounded
with the effect of the mechanism (such as group-based training), making it hard to differentiate and properly evaluate whether computer training on its own was having an effect. The GRADE strength of the evidence was low, emphasizing the need for assessing the full potential of computer training in social connectedness.

### Telecare

Telecare was among the less frequent interventions in the review studies, but when included, it appeared to reduce social isolation and loneliness [42,48,49]. Husebø and Storm [48] comprehensively investigated telecare services for older adults. After reviewing 12 primary studies covering this area, they found that virtual visits by clinicians can reduce social isolation and loneliness in older adults compared with no contact. Other benefits included self-management of medication and self-care, which can postpone admission to long-term care or substantial in-home care. In all areas, telecare both directly and indirectly affected participants’ perceived social isolation and loneliness. In 4 of the studies, older adults interacted with others experiencing similar issues. These interactions were highly valued and enabled the development of deeply empathetic connections [59]. By contrast, Damant et al [45] found no conclusive evidence of enhanced social connectedness among older adults using videoconferencing (14 studies).

Although none of the authors described the key features of successful telecare interventions, an emergent theme from successful primary studies was a high frequency of contacts. Interventions designed for regular and frequent contact were apparently more successful than interventions delivered on demand (eg, when a resident needed clinical attention). Overall, 34 primary studies in the analyzed reviews covered this category. The impact of telecare on social connectedness was inconclusive, and uncertainty was further increased by the poor reporting of the results. Consequently, the GRADE strength of evidence in this area was very low.

### Robotics

Robotics is a cutting-edge field and was mentioned in only 6 reviews. Some studies found that a pet robot provides the same level of benefit as animal-assisted therapy, which is known to reduce loneliness and social isolation [35,36,42]. Ibarra et al [40] mentioned that older adults feel embarrassed when conversing with a virtual pet, although this discomfort might have been exacerbated by audio problems and latency in messages. Choi and Lee [58] provided an excellent systematic review covering animal robots, humanoid robots, and mobile robots. They identified a notable development trend in robotic interventions from simpler animal robots to complex, multifaceted web-based social platforms that offer emotional support and promote social participation, cognition, physical activity, nutrition, and sleep. In most of their examined studies, robotic interventions decreased loneliness and social isolation. Although no other study has looked at the impact of virtual pets on loneliness, this seems to be a promising area that needs further research, with the potential of virtual or robotic pets offering a distinct advantage of social affordance compared with animal-assisted interventions.

Khosravi et al [42] and Antunes et al [44] examined conversational agents designed for companionship and video communication, enabling older adults to connect with family members and friends and offering “talk therapy.” Overall, these agents improved social interaction and reduced the loneliness of participants. With the ongoing development of pseudo–artificial intelligence (AI) technology and the advent of voice-assisted agents, such as Alexa and Siri, conversational agents are promising solutions and need to be further explored.

Khosravi and Ghanachi [43] concluded that robotic technologies increase the perception of being socially connected and hence, exert a positive impact on social and emotional well-being. However, the perception of not being socially isolated differs from the actual reduction in social isolation, which depends on real person connections. On the adapted effectiveness scale, robotic technologies scored 1.8 out of 3.0. Although these reviews indicate that social connectedness can be increased through robotics, this category is still new, and further studies on AI conversational agents and other robotic interventions are required. Therefore, the GRADE strength of evidence in this category is moderate to low.

### Gaming

According to Khosravi et al [42], Video gaming devices such as Wii, which capture natural physical activities, achieve a greater reduction in loneliness and better social interaction than typical video games. Chen and Schultz [35] and Williams et al [38] found that Wii strengthens social interaction and reduces loneliness; however, web-based gaming was outside the scope of these studies. Choi and Lee [58] reported 3 studies in which video games and exercises were combined into an exercise game, enabling communication with others. This game reportedly reduces loneliness during exercise. However, the GRADE confidence in the effect of gaming is very low because solid evidence is lacking.

### 3D and AR

Similar to robotics, 3D environments have been newly introduced as a loneliness-reduction intervention technique and are rarely reported. Khosravi et al [42] reported that most studies on 3D environments included a small number of participants, suggesting a need for further research. Although the underlying studies reported a positive impact of 3D environments on loneliness, the weak methodology and reporting of findings cast doubt on their validity. This category has been underexplored and requires further research. Current developments in 3D worlds, Facebook’s foray into Metaverse, and AR developments by prominent companies such as Google and Microsoft should accelerate the design of 3D interventions for older adults. Owing to a lack of evidence, the GRADE confidence in the effects of 3D environments and AR is very low.

### Usability Impact on Effectiveness of Technology

There were few reviews that examined the usability of technology and its impact on the effectiveness of interventions. Some reviews identified a link between usability and acceptance of technology; more accessible devices were distinctly more likely to be embraced by users than less accessible devices.
[19,40,48,58]. Even when usability was not a formal outcome, the studies observed participants’ initial feelings of uncertainty and fear of using technology. These trepidations were overcome with time, familiarity, and sufficient training [19,40]. Ibarra et al [40] reported that touch screen computers were especially effective in reducing loneliness and social isolation, highlighting the importance of an easily accessible system or interface. Husebø and Storm [48] noted that when introducing technology to older adults, a usable and simple design that considers the likely interactions of older adults with technology is essential. Choi and Lee [58] identified 6 studies in which the use of and attachment to ICT interventions increased over time along with the average density of social networks.

However, systematic reviews typically neglect the human-computer interaction components of intervention technology. Moreover, standardized measures of usability (eg, the System Usability Scale) for intervention studies have not been defined [19,40]. The use and adoption of technology by older adults largely depends on the learning ability of the individual and the perceived difficulty of use. To ensure that technology can effectively reduce loneliness in older adults, these potential barriers should be examined appropriately.

Overall, the reported studies showed that whether technology can reduce loneliness depends on its usability. An intervention perceived as difficult to use by older adults cannot be effective. This aspect must be further investigated to improve the success of technology interventions.

Owing to a lack of evidence, the GRADE confidence in the effect of usability on the success of intervention technologies is very low.

Summary Recommendations

On the basis of the results, Table 4 summarizes the key recommendations extracted for technology interventions targeting social isolation, connectedness, and loneliness.

We have also summarized the key recommendations for study design targeting social isolation, connectedness, and loneliness in Table 5.
### Table 4. Summary of key recommendations for technology interventions.

<table>
<thead>
<tr>
<th>Category</th>
<th>Key recommendations</th>
<th>Certainty of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>General ICT(^a)</td>
<td>- Simple technology interventions can be more successful than complex ones. Usability is a potentially important outcome.</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>- ICT is not recommended for increasing either the quantity or quality of communications or helping to establish new relationships. It is recommended for maintaining and enhancing existing relationships and access to services (such as health-related services).</td>
<td></td>
</tr>
<tr>
<td>SNS(^b)</td>
<td>- SNS is not recommended as an intervention for loneliness and isolation as SNS use has often been shown to worsen loneliness.</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>- SNS is useful in knowledge and support acquisition scenarios, which can themselves reduce loneliness. Research shows that SNSs are generally more successful in these scenarios than in making new connections.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Privacy is an important concern among older adults and needs to be considered when designing an intervention.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Usability is potentially a very important theme and needs to be factored into the study design.</td>
<td></td>
</tr>
<tr>
<td>Videoconferencing</td>
<td>- Videoconferencing reduces loneliness by providing social support and improving the existing conditions in health care–type situations.</td>
<td>Moderate low</td>
</tr>
<tr>
<td></td>
<td>- Financial investment (eg, cost of computer hardware) needs to be considered when planning a videoconferencing intervention.</td>
<td></td>
</tr>
<tr>
<td>MIM(^c)</td>
<td>- MIM is recommended for rapid deployment as it is easy to use, and applications such as WhatsApp additionally allow the sharing of pictures, which can improve social connectedness.</td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td>- MIM can replace email, but designers must be wary because any lack of responses can increase the perception of loneliness.</td>
<td></td>
</tr>
<tr>
<td>Computer and internet training</td>
<td>- Longer training periods are recommended with shorter-duration studies (as highlighted above) as they have been the most effective.</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>- For reducing loneliness, group-based training is more effective than one-to-one training.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- The study design should reflect whether the training or use of the intervention causes reduction in loneliness.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- RCTs(^d) are particularly important in the study design as they determine precise effect sizes.</td>
<td></td>
</tr>
<tr>
<td>Telecare</td>
<td>- Frequency of contact combined with telecare solution influences the success of an intervention. Interventions designed for regular frequent contact are more successful than interventions delivered on-demand; for example, when a resident needs clinical attention.</td>
<td>Very low</td>
</tr>
<tr>
<td></td>
<td>- Videoconferencing groups such as group counseling can help to reduce feelings of anxiety, isolation, and loneliness and provide emotional and social support; however, designers must understand that some participants do not immediately feel at ease with others, especially in a group setting.</td>
<td></td>
</tr>
<tr>
<td>Robotics</td>
<td>- Pet robots can provide the same advantages as animal-assisted therapy in reducing loneliness and social isolation; study designs can mimic previous studies in this area.</td>
<td>Moderate low</td>
</tr>
<tr>
<td></td>
<td>- Conversational agents provide companionship through social interaction, enabling older adults to connect with family members and friends (social presence). These agents can be effective and are recommended for intervention studies.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- RCTs are recommended in the study design of robotic interactions, especially as this area is understudied.</td>
<td></td>
</tr>
<tr>
<td>Gaming</td>
<td>- Video gaming devices such as Wii, which capture natural physical activities, are recommended as they reduce loneliness and provide better social interactions than typical video games.</td>
<td>Very low</td>
</tr>
<tr>
<td>3D and augmented reality</td>
<td>- Too few of the existing studies provide robust recommendations, and further longitudinal and cross-sectional RCT studies are needed in this area.</td>
<td>Very low</td>
</tr>
</tbody>
</table>

\(^a\)ICT: information and communications technology.  
\(^b\)SNS: social networking site.  
\(^c\)MIM: mobile instant messaging.  
\(^d\)RCT: randomized controlled trial.
Table 5. Summary of key recommendations for technology interventions.

<table>
<thead>
<tr>
<th>Category</th>
<th>Key recommendations</th>
<th>Certainty of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group vs one-to-one</td>
<td>Studies should be designed as group-based interventions, as they appear to better facilitate social connectedness than one-to-one interventions.</td>
<td>Very low</td>
</tr>
<tr>
<td>Effectiveness of technology interventions</td>
<td>Certain types of technologies (information and communications technology and video-conferencing) are particularly suitable as interventions for social isolation and loneliness. For best results, studies should be designed to strengthen existing bonds, especially the connections between family members (eg, grandchildren).</td>
<td>Moderate low</td>
</tr>
<tr>
<td>Frequency of use</td>
<td>Frequency of use is encouraged (greater use increases the effect size).</td>
<td>Very low</td>
</tr>
<tr>
<td>Training</td>
<td>Training, especially in the use of technology, is encouraged as it improves the success of the study.</td>
<td>Very low</td>
</tr>
<tr>
<td>Duration</td>
<td>Shorter-duration studies are recommended (shorter studies achieve better results than longer-duration studies).</td>
<td>Low</td>
</tr>
<tr>
<td>Outcome measures</td>
<td>The impact of intervention is stronger on social isolation than on loneliness, and studies should be designed to look further on how to impact loneliness. Use of standardized measures such as University of California Los Angeles Loneliness Scale and the Lubben Social Network Scale is recommended.</td>
<td>Low</td>
</tr>
<tr>
<td>Mechanisms</td>
<td>Mechanisms by which interventions reduce social isolation through the design of studies, including the gaining of social support, engagement in activities of interest, the making of new connections, and search for new information, should be clearly defined at the outset.</td>
<td>Very low</td>
</tr>
<tr>
<td>Usability</td>
<td>Intervention studies should adopt standard measures of usability (eg, System Usability Scale) because the adoption of technology by older adults largely depends on learnability and perceived difficulty of use. These barriers often prevent technology from reducing loneliness in older adults.</td>
<td>Very low</td>
</tr>
</tbody>
</table>

Discussion

Principal Findings

This umbrella review, as highlighted in the analyzed reviews, found that different studies adopted a vast diversity of outcome measures and nonstandard definitions of loneliness and isolation [20,33,35,39,42], and therefore, heterogeneity, lack of clarity, and lack of consistency across reviews have influenced the interpretations of their findings. The strengths of the evidence for effectiveness ranged from very low (robotics, telecare, 3D or AR, and video games) to moderate (ICT). These low ratings were attributed to the poor overall quality of evidence, study design, and outcomes. However, our umbrella review showed that despite the heterogeneous quality and diverse scope of existing reviews, which prohibit the drawing of generalizable conclusions, technology can effectively target social disconnectedness in older adults [61,62].

An umbrella review following the JBI methodology [12,26] was warranted because the types of reviews, levels of evidence, and outcomes of different reviews range widely in quality, from meta-analyses to qualitative syntheses, and the availability of a wide range of reviews allows our umbrella review to comprehensively consolidate the current state of evidence on interventions for social connectedness. As highlighted in the analyzed reviews, different studies adopted a vast diversity of outcome measures and nonstandard definitions of loneliness and isolation [20,33,35,39,42], and therefore, heterogeneity, lack of clarity, and lack of consistency across reviews have influenced the interpretations of their findings. Many of the review authors included social isolation and loneliness interchangeably when selecting their intervention studies, failing to recognize that each condition is a component of social disconnectedness. This confusion weakens the recognition of differing results, as loneliness is generally more resistant to interventions than social isolation. Although some loneliness measures (eg, UCLA and De Jong Gierveld Scale) have been regularly adopted, the Lubben Social Connectedness Scale was applied in only 9 of the primary studies. This scale, which assesses an individual’s psychological sense of belonging, might better reflect the interaction among different dimensions of social connectedness than commonly adopted measures [39,60]. Most of the primary studies developed their measures or used less common measures, such as the Self Anchoring Scale, Social Network Structure, Social Supportive Behavioral Scale, and Social Connectedness Index.

The designs and qualities of the reviewed primary studies varied widely. Several reviews included RCTs and pilot, qualitative, and quantitative studies. In addition, the studies reviewed by
Khosravi et al [42] were conducted across the health domain. The primary studies in each review are typically nonoverlapped, indicating that the reviewers’ searches did not capture all relevant studies and sometimes omitted important studies and assessments of bias risk.

The findings of many underlying primary studies in the reviews were compromised by poor study designs, leading to conflicting information. For example, when reviewing the effects of computer and internet training on loneliness, Chen and Schulz [35] reached an inconclusive verdict. Choi et al [20] reported a significant impact of the intervention, and Bornemann [33] demonstrated no significant effect of the intervention. Moreover, the effect size calculated by Bornemann [33] differed from the more accurate calculation by Choi et al [20], although both reviews shared 5 primary studies in their meta-analyses.

The reviewers generally agreed on the effectiveness of group-based interventions. Reviews examining the designs of the reviewed studies noted group-based interventions yielded positive effects on social disconnectedness [24,36,37,40,41]. The different effects of group interventions can be attributed to the social interaction value of being in a group rather than the actual intervention [36,37]. When the intervention was delivered over a longer duration, the effect of the group activity diminished over time, and the intervention became less effective. Interventions with a participatory, productive, and collaborative focus [36], especially educational [37], appeared to realize an effective group-based intervention.

The reviews varied in scope, from assessments of the effectiveness of interventions, such as videoconferencing, to overviews of studies published in the field. The inclusion criteria and quality assessments of the primary studies also differed among the reviews, diminishing confidence in their findings. Our study confirmed a low quality of evidence in this field, whereas improved technology interventions for older adults are increasingly demanded by both policymakers and health professionals. Although the existing guidelines can encourage standardization of systematic reviews, these guidelines were largely ignored by researchers; accordingly, the strength of the reviews is diminished, which in turn led to the quality of evidence GRADE scores also being generally low.

The scope of the reviews varied from a specific focus on the effectiveness of a targeted intervention (such as computer training) to an overview of the published studies in the field. The inclusion criteria for the primary studies and their quality assessment depended on tools used for rating rigor and bias. Such variations cast doubt on the conclusions of these reviews. This review confirms the lack of high-quality evidence in the field and highlights the failure to adhere to the existing guidelines. Standardization of systematic review reporting is expected to strengthen confidence in the review conclusions.

Unlike their younger counterparts, older adults often lack the skills, functional capacity, and accessibility to adopt digital technology [63], which has led to the so-called “digital divide” among populations. However, these expansive categories are not mutually exclusive to older adults. In resource-restricted settings, they also incorporate gender differences, age, economic status, cultural practices, and educational qualifications [63] and can play an important role in reducing the existing digital divide between younger and older adults. Most of the reviews did not adequately consider these differences, presuming a general dearth of resources for older adults. Also important are the usability and design of the intervention, which were notably absent in the primary studies. The individual circumstances of older adults (including finances, environment, and access to resources) may influence the success of interventions. When usability was examined (as in some reviews), it was done without the use of standardized usability measures, but usability did influence the effectiveness of the intervention; therefore, further exploration of this area is vitally important.

To improve the quality of results, interventions should be tailored to match the specific needs of older adults, and sufficient training should be provided for using the interventions. This tailoring requires the involvement or participation of participants in training in a variety of formats [24,41]. As usability issues can reduce the effectiveness and uptake of an intervention, neglecting usability as an outcome measure reduces confidence in a holistic discussion of the effectiveness of an intervention. Thus, the potential impact of technology on social connectedness in older adults requires further investigation.

**Comparison With Prior Work**

Our umbrella review is one of the few works that have looked at technology interventions for social connectedness for loneliness, following a well-established systematic approach such as the JBI umbrella review method. In examining other works, we came across reviews that focused on interventions generally [64-67] as opposed to technology interventions, which we noted was a tendency to bundle technology interventions with other common interventions, such as cognitive enhancement work groups, adult day center attendance, gender-based social groups, activities such as befriending and mentoring programs, and animal-assisted therapy. The problem with this approach is that it dilutes attention away from the different types of technology interventions and how they individually affect social connectedness and compares them as a category to other types of interventions. The fallacy of this approach for technology interventions is that it only presents high-level evidence and comparisons of diverse interventions, preventing readers from understanding the deeper nuances that make certain technology interventions more successful than others. Indeed, many of reviews of reviews are scoping reviews [66,67], corresponding to a more expansive inclusion criterion, which limits more substantive findings about specific interventions and rather presents a broader scope of general findings. Our umbrella review, on the other hand, by focusing specifically on technology interventions, extends the understanding of specific technology interventions by reporting and rating the evidence. By doing so, we were able to identify current evidence gaps related to the understanding of the types of technologies, study design, and their impact on social connectedness, loneliness, and isolation. We were also able to identify weaknesses in the reviews and areas for future research.

**Strengths and Limitations**

Most of the reviews demonstrated a need for stronger evidence on the effectiveness of technology interventions that reduce
lonelessness. Weak methodologies have limited the ability of reviews to establish conclusive remarks on their effectiveness [35,42]. Many outcome measures have greatly limited comparisons, which affected the interpretation of the results.

The present review may also have been biased by accepting only English-language publications. However, many of the shortcomings and limitations of this umbrella review stem from the underlying problems of the primary papers included in the reviews. Among the common shortcomings were small-scale implementations with small sample sizes, low levels of evidence, and short periods of assessment.

Another recurring limitation was the inconsistent definitions of social concepts. Social concepts such as loneliness, social isolation, and social connectedness were formally defined, but the authors did not use these definitions consistently; instead, they were often used interchangeably, inherently confounding measurements of these outcomes. The reviews were generally heterogeneous in focus (eg, addressing loneliness and depression) and discussed various interventions and syntheses of outcomes (eg, meta-analyses, qualitative reviews, and mixed methods). Accordingly, the present review interchanges the terms social connectedness and social disconnectedness to describe combinations of singular aspects such as social isolation and loneliness. Nevertheless, the methodology was the greatest limitation. Finally, the absence of gray literature in the reviews may have increased publication bias and led to the lack of inclusion of evidence for interventions that are not typically indexed in bibliographic databases. Future systematic reviews should consider including gray literature in the included studies.

The methodological limitations of the reviewed studies impaired the internal validity and usefulness of the reviews for technical and policy decision-making, as highlighted by the reviewers [20,24]. The reviews reported on diverse methodologies, including the use of nonstandardized outcome measures, which broaden the perspective but risk biasing the conclusions. Furthermore, as interventions vary widely in nature, direct comparisons are difficult, and the definitions of technology interventions are rather narrow in some studies [39].

The reviewed quantitative studies collected their data with questionnaires using scales developed for the study purpose. The reliability and validity of these nonstandardized scales are difficult to evaluate. Most reviews pointed out the suboptimal methodological quality of studies in this field, particularly the scarcity of RCTs (<28% of studies) and the dominance of quasi-experimental studies, which challenge the delivery of robust conclusions.

Therefore, the results of this review should be interpreted with caution.

Suggestions for Future Research and Policy Implications

Various technology interventions in different formats offer many ways to engage older adults. However, usability was rarely discussed in the reviews and was not assessed as an outcome measure. Although the existing guidelines encourage the standardization of systematic reviews, they have not been followed with the required rigor. Equally, the underlying primary studies of the reviews failed to address causation in a rigorous study design, and their heterogeneity limited their generalizability. It appears that there is a need for more studies on the multidimensional impact of technology on social connectedness, along with the assessment of other measures that may be interacting with technology use (such as educational attainment, psychological resilience, and age-friendliness of environments). Robotics is a relatively new technology that has emerged to be promising, but there are very few studies in this domain. Research on mobile technology interventions for social isolation is also encouraged as mobile phone technology provides opportunities for increasing the uptake of technology interventions targeting loneliness in older adults. Our results on the grading of evidence revealed that the strength of evidence was generally low to very low, indicating that the efficacy of the interventions is unclear and that more rigorous research is needed.

Our review provides insights into strategies to reduce loneliness and isolation for older adults using technology interventions, with implications for future research, policy, and practice. Attention to social connections needs to be incorporated into existing preventative efforts for chronic diseases in older adults. Chronic illnesses develop slowly over decades. Since social connectedness is known to impact multiple mechanistic pathways in both the development and progression of disease, it warrants attention in primary, secondary, and tertiary prevention efforts. Given the lower economic costs of technology interventions for individuals, families, employers, and the broader health care system, we urge health care and health policy professionals to prioritize the investigation of technology interventions for social connections in prevention efforts.

Conclusions

This umbrella review consolidates the state-of-the-art knowledge on the types of technology interventions that influence social connectedness in older adults and their effectiveness. The data were collected from the last 2 decades. Technology purportedly enables long-distance interactions, allowing older adults to become socially connected, obtain support, expand their social networks, and strengthen their existing ties. Some important themes that would improve the effectiveness of technical interventions for older adults emerged from the literature, namely group interventions, short-duration training and study programs, the use of general ICT, and videoconferencing. These implementations are more effective for maintaining existing connections than for building new ones. Certain technologies, such as robotics (including virtual pets), AI-based conversational agents, and MIMs, show promising potential but have been underexplored.

We attempted to determine which technology interventions can effectively improve social connectedness. The following conclusions emerged from our study. Reports on the effectiveness of computer and internet training on loneliness and social isolation provided mixed and inconclusive results. General ICT and internet-mediated communications were shown to reduce loneliness and social isolation in most studies, although the results apparently depend on the frequency of use.
and the time frame of the study, with shorter studies being more successful than longer ones. ICT interventions help socially isolated older adults through a range of mechanisms, including gaining social support, providing connections to the outside world, introducing new friends, and boosting self-confidence. All of these mechanisms must be studied hand in hand to gain a complete understanding of these processes. Finally, in our GRADE evaluation, most of the evidence was rated as moderate to very low, reflecting methodological issues, the small number of RCTs, diverse outcome measures and definitions, and mixed results. Such low scores highlight the need for high-quality research in this area.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Search terms used in database.
[DOCX File, 16 KB - aging_v5i4e40125_app1.docx ]

Multimedia Appendix 2
Revised Assessment of Multiple Systematic Reviews quality ratings.
[DOCX File, 27 KB - aging_v5i4e40125_app2.docx ]

Multimedia Appendix 3
Characteristics of the included reviews.
[DOCX File, 32 KB - aging_v5i4e40125_app3.docx ]

Multimedia Appendix 4
Grading of Recommendations, Assessment, Development, and Evaluations—certainty of evidence.
[DOCX File, 17 KB - aging_v5i4e40125_app4.docx ]

References


Abbreviations

- AI: artificial intelligence
- AR: augmented reality
- GRADE: Grading of Recommendations, Assessment, Development, and Evaluations
- ICT: information and communications technology
- JBI: Joanna Briggs Institute
- MIM: mobile instant messaging
- PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
- R-AMSTAR: Revised Assessment of Multiple Systematic Reviews
- RCT: randomized controlled trial
- SNS: social networking site
- UCLA: University of California Los Angeles

©Eric Balki, Niall Hayes, Carol Holland. Originally published in JMIR Aging (https://aging.jmir.org), 24.10.2022. 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.
The Impact of COVID-19 on Older Adults’ Perceptions of Virtual Care: Qualitative Study

Lama Abdallah; Paul Stolee, DPhil; Kimberly J Lopez, DPhil; Alexandra Whate, MSc; Jennifer Boger, DPhil; Catherine Tong, DPhil

1School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada
2Department of Recreation and Leisure, University of Waterloo, Waterloo, ON, Canada
3Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Catherine Tong, DPhil
School of Public Health Sciences
University of Waterloo
200 University Ave. West
Waterloo, ON, N2L3G1
Canada
Phone: 1 5198884567
Email: catherine.tong@uwwaterloo.ca

Abstract

Background: In response to the COVID-19 pandemic, older adults worldwide have increasingly received health care virtually, and health care organizations and professional bodies have indicated that virtual care is “here to stay.” As older adults are the highest users of the health care system, virtual care implementation can have a significant impact on them and may pose a need for additional support.

Objective: This research aims to understand older adults’ perspectives and experiences of virtual care during the pandemic.

Methods: As part of a larger study on older adults’ technology use during the pandemic, we conducted semistructured interviews with 20 diverse older Canadians (mean age 76.9 years, SD 6.5) at 2 points: summer of 2020 and winter/early spring of 2021. Participants were asked about their technology skills, experiences with virtual appointments, and perspectives on this type of care delivery. Interviews were digitally recorded and transcribed. A combination of team-based and framework analyses was used to interpret the data.

Results: Participants described their experiences with both in-person and virtual care during the pandemic, including issues with accessing care and long gaps between appointments. Overall, participants were generally satisfied with the virtual care they received during the pandemic. Participants described the benefits of virtual care (eg, increased convenience, efficiency, and safety), the limitations of virtual care (eg, need for physical examination and touch, lack of nonverbal communication, difficulties using technology, and systemic barriers in access), and their perspectives on the future of virtual care. Half of our participants preferred a return to in-person care after the COVID-19 pandemic, while the other half preferred a combination of in-person and virtual services. Many participants who preferred to access in-person services were not opposed to virtual care options, as needed; however, they wanted virtual care as an option alongside in-person care. Participants emphasized a need for training and support to be meaningfully implemented to support both older adults and providers in using virtual care.

Conclusions: Overall, our research identified both perceived benefits and perceived limitations of virtual care, and older adult participants emphasized their wish for a hybrid model of virtual care, in which virtual care is viewed as an addendum, not a replacement for in-person care. We recognize the limitations of our sample (small, not representative of all older Canadians, and more likely to use technology); this body of literature would greatly benefit from more research with older adults who do not/cannot use technology to receive care. Findings from this study can be mobilized as part of broader efforts to support older patients and providers engaged in virtual and in-person care, particularly post–COVID-19.

(JMIR Aging 2022;5(4):e38546) doi:10.2196/38546
virtual care; older adult; qualitative; COVID-19; elderly population; aging; telehealth; digital care; technology usability; patient perspective; technology access

Introduction

As a result of efforts to limit the spread of the virus that might occur through in-person appointments, the COVID-19 pandemic accelerated the shift to virtual health care. Virtual health care, subsequently, was widely adopted across Canada and beyond [1-5]. Simultaneously, policies at the institutional, national, and international levels flexed to accommodate recommendations on the use of virtual care within existing health care models [6,7]. Virtual care can be defined as “any interaction between patients and/or members of their circle of care, occurring remotely, using any forms of communication or information technologies, with the aim of facilitating or maximizing the quality and effectiveness of patient care” [8]. Virtual care is not limited to a particular technology or platform (eg, it can include the telephone) and is often used interchangeably with “telemedicine” or “eHealth” [8,9]. Prior to the COVID-19 pandemic, virtual care activities, although possible, were not common in Canada [10,11]. Although the COVID-19 pandemic sparked a dramatic increase in virtual care in Canada [1,2] and worldwide [4], questions remain about the quality and role of virtual care in practice [6,12], particularly with older patient populations.

Although Bhatia et al [1] found that older patients were the highest users of virtual care during the pandemic, Senderovich and Wignarajah [13] expressed concerns about the maintenance of the therapeutic alliance between physicians and older patients receiving virtual care (a therapeutic alliance being a patient-doctor relationship that supports positive health outcomes). Preparadigm research in the United Kingdom by Hammersley et al [12] found that older patients were less likely to choose virtual care than were younger patients. The experience of older patients with virtual care is thus of continued interest, both during and after the pandemic. Despite common misconceptions about older adults and technology, a national survey conducted in July 2020 found that 72% of older Canadians feel confident about their ability to use existing technologies, such as smartphones or video calls [14]. In the 3 months prior to the July 2020 survey, 52% of older Canadians accessed virtual care and 79% were satisfied with the virtual care received [14]; the bulk of the virtual care they received was over the telephone. Although studies have investigated the use of virtual care with older adults before the COVID-19 pandemic, this evidence is largely quantitative; there is a lack of qualitative data that reflect the perspectives and experiences of older Canadians accessing virtual care throughout the pandemic. Lopez et al’s [21] analysis of older adults’ use of technology during the pandemic found a notable increase, including broader adoption of videoconferencing software/video calls. Teti et al [22] emphasize the importance of reflecting qualitative data throughout the COVID-19 pandemic to understand how COVID-19 impacts populations as a social event as well as a medical pandemic. Qualitative approaches play a vital role in understanding social responses to pandemics, as they allow us to understand the lived experiences of those who are disproportionately impacted, including older adults [22]. The aim of this study was to use a systematic qualitative study to understand how older adults experienced virtual care during the pandemic and to include their perspectives on virtual care as an alternative or supplement to in-person care. Organizations, such as the Canadian Medical Association (CMA), have indicated that virtual care is “here to stay,” even if/when no longer necessitated by the pandemic. If virtual care is indeed here to stay, our interviews with older adults will contribute to broader discussions on how and when to use virtual care in a manner that reflects their experiences, wishes, and perspectives.

Methods

Study Design

This research is part of a larger study [21] in which we used a longitudinal qualitative study [23] approach to listen to older adults speak about their social connections and experiences of digital connectivity early (summer of 2020) and later (winter and early spring of 2021) in the COVID-19 pandemic. Our research team is situated in Ontario, Canada, and eligible participants included any older Canadian (aged 65 years or more) able to complete an English-language telephone/video interview and provide informed consent.

Ethical Considerations

We received ethics clearance from the University of Waterloo’s Office of Research Ethics (ORE #42265).

Recruitment

A purposive sampling strategy [24] was used to recruit a diverse sample of older adults (eg, rural/urban; community/assisted living; diverse abilities, socioeconomic profiles, genders, and ethnicities). Recruitment during the beginning of a global pandemic that was disproportionately impacting older adults was challenging, and we used several recruitment approaches to access diverse older adults. We recruited using social media (eg, Twitter), emails to large established groups with older adult members (blinded for review), telephone calls to older adults within our personal networks (ie, asking our personal contacts to share study materials within their networks), and promotion of our study via teleconferences with older adult participants. In total, 20 older adults completed the baseline in-depth interviews in the spring of 2020, which coincided with the first wave of COVID-19 in Canada. In the spring of 2021, follow-up interviews were conducted with 12 (60%) participants from the baseline sample, coinciding with the second wave of COVID-19 in Canada. Of the 12 participants, 8 (67%) did not participate in the follow-up interviews because of death (n=1, 12.5%), they could not be reached (n=3, 37.5%), or they declined to participate in a second interview (n=4, 50%). Recruitment for follow-up interviews coincided with a particularly challenging period of the pandemic (ie, stringent lockdowns; rising case...
counts and deaths, especially among older adults; and the darker, bleaker winter months); 3 (75%) of the 4 participants who declined to follow-up specifically expressed that this was because of the challenging period and timing. Participant characteristics are summarized in Table 1.

**Table 1. Participant characteristics.**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline (N=20)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td>14 (70)</td>
</tr>
<tr>
<td>Men</td>
<td>6 (30)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>66-94</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>76.9 (6.5)</td>
</tr>
<tr>
<td><strong>Living arrangement, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Alone</td>
<td>7 (35)</td>
</tr>
<tr>
<td>With spouse</td>
<td>10 (50)</td>
</tr>
<tr>
<td>With adult child</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Assisted living facility</td>
<td>2 (10)</td>
</tr>
<tr>
<td><strong>Community type, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>11 (55)</td>
</tr>
<tr>
<td>Suburban</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Rural</td>
<td>8 (40)</td>
</tr>
<tr>
<td><strong>Province, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Ontario</td>
<td>17 (85)</td>
</tr>
<tr>
<td>Alberta</td>
<td>3 (15)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>16 (80)</td>
</tr>
<tr>
<td>BIPOC*</td>
<td>4 (20)</td>
</tr>
</tbody>
</table>

*BIPOC: Black, Indigenous, and people of color.

**Data Collection**

All interviews were conducted over the telephone or via videoconferencing software. Baseline interviews lasted an average of 53 minutes (minimum 24 minutes; maximum 74 minutes); follow-up interviews lasted an average of 60 minutes (minimum 21 minutes; maximum 112 minutes). The interview questions (Multimedia Appendix 1) were developed in consultation with older adults from our “Seniors Helping as Research Partners” group and informed by our interdisciplinary research team, which includes experts in systems design engineering for older adults, recreational therapy, social gerontology, and designing health care systems for older patients. The first two-thirds of the interview focused on participants’ use of and access to technology, comfort with technology, etc, and the final third of the interview focused specifically on virtual health care. Interviews and analytic debriefs were digitally recorded, run through otter.ai transcription, and then cleaned and anonymized by research assistants using protocols established by our team. Anonymizing the transcripts included the assignment of a pseudonym for each participant. Additional details about the overarching study, recruitment, and data collection may be found here [21].

**Data Analysis and Strategies**

Our team-based analysis (ie, multiple members of the research team, drawing on different disciplinary perspectives to collectively analyzing the data; see Guest and MacQueen [25]) process used a framework analysis approach [26] that included the following steps:

- Step 1 (familiarization): Each transcript was read by 1 of 3 coauthors (LA, CT, and AW), who were the same coauthors who conducted the interviews.
- Step 2 (development of a coding framework): All coauthors used the initial read of the data, field notes, and debriefs to develop an initial set of thematic codes.
- Step 3 (indexing and charting): Three coauthors, (LA, CT, and AW) engaged in line-by-line coding [27].
- Step 4 (summarizing and synthesizing): The coding structure was further refined through team analysis meetings and shared coding memos to consolidate the most salient themes presented later.
Rigor strategies included reflexive memoing, an audit trail within NVivo (QSR International) [28], and team-based examination of the data and each step of the analysis [25]. We also reviewed our findings and interim analysis with 4 participants (ie, member checking and reviewing our interpretations of the data) via an online focus group that were recorded and transcribed to inform the analysis.

Results

Participant Details

We interviewed a total of 20 patients at baseline (14, 70%, women; 6, 30%, men) and 12 at follow-up (8, 67%, women; 4, 33%, men). Ages ranged from 66 to 94 years, with an average of 76.9 (SD 6.5) years. Most patients lived alone or with a spouse, in urban or rural communities of Ontario and Alberta. Of the 20 patients, 4 (20%) were Black, Indigenous, or people of color (BIPOC) and the remainder (n=16, 80%) were White.

In discussing their experiences with virtual care during the pandemic, older adults broadly shared 3 high-level themes: (1) their experiences accessing health care during the pandemic, (2) their perceived benefits and limitations of virtual care, and (3) their perspectives on when virtual care is acceptable and appropriate. In the quotes presented later, the suffixes included after the patient pseudonym and biographical information (B and F) refer to baseline and follow-up interviews, respectively. Participants often shared their perspectives on virtual care in the first interview and in the second replied, “Like I said last time…”; thus, more of the presented quotes are from the baseline interviews than from the follow-up interviews. There was not a notable change in participants’ perspectives on virtual care across the 2 time points.

Experiences With Health Care During the Pandemic

Participants described their health care experiences during the pandemic in terms of issues with accessing care, and their pandemic experiences of in-person and virtual care. Although we did not specifically probe for issues with accessing care, many participants mentioned that they had not seen their primary care providers for months; some had not contacted their providers since the start of the pandemic:

No, because I haven’t been in touch with them since well... [Richard, 76 years, male, B]

Since March. [Joan, 66 years, female, B]

No, I haven’t seen my doctor this year. [Richard, 76 years, male, B]

At baseline, Susan, aged 82 years, expressed that older adults are reluctant to access in-person care because they are weighing the risk of contracting COVID-19 against the risk of missing an appointment:

And, and the other thing is that people hesitated maybe too much sometimes to go to the hospital. Like you said, people full of coughing in an emergency room. But, you know, there are situations where people might have delayed going and they needed to go...Yeah, they're really...it’s assessing the risk. Like, you know, maybe I’m better to stay home than to get COVID. [Susan, 82 years, female, B]

The disruption and discontinuity of care resulting from the COVID-19 pandemic caused participants to feel fear and anxiety about the frequency and quality of care they received. However, in comparison to the risk of contracting COVID-19 when accessing in-person care, they seemed to fear the virus more than the potential complications that could result from missing so many in-person appointments.

Participants felt that they accessed in-person care less regularly than they would have prior to the pandemic. In-person care was mostly accessed for emergencies, specialist services (eg, oncology, physiotherapy), and services that could not be accessed online (eg, diagnostic imaging, blood work). When in-person care was accessed, some aspects of the care were organized virtually. At baseline, Nancy, aged 66 years, had an x-ray performed in-person, with the results of the x-ray communicated virtually:

I've had one X-ray and that's about it, I think. [Nancy, 66 years, female, B]

Okay. And then you’ve got the results of the x-ray over the phone? [AW]

That's right, yeah. [Nancy, 66 years, female, B]

Virtual Care

When asked about their experiences accessing virtual health care during the pandemic, most participants were able to discuss a time when they accessed care virtually either at baseline or follow-up. Virtual care usually involved phone calls for completing intake, scheduling, and accessing appointments; text or email messages for sending photos of health concerns; and emails/phone calls for receiving requisitions and test results. Few participants accessed virtual care in the form of video calls; most virtual care had been provided over the telephone, with some referrals or results (eg, of bloodwork) being communicated over email. In general, participants were satisfied with the virtual care they received from their family physicians. Participants also felt their relationships with regular providers were not negatively impacted by the pandemic; participants maintained their patient-doctor relationships virtually despite changes in care delivery and frequency. Participants were less comfortable with certain tasks being performed virtually, such as being prescribed a new medication, diagnosis, or meeting with a specialist for the first time.

Perceived Benefits of Virtual Care

When prompted about the benefits of virtual care, all participants identified at least 1 positive aspect of virtual care compared to in-person care. The most common perceived benefit was convenience, which was discussed by most participants. Other commonly cited benefits were improved safety due to the avoidance of unsafe situations associated with in-person care (eg, contracting a virus) and the efficiency of the health care provider. Table 2 summarizes the perceived benefits of virtual care identified by participants.

Table 2 summarizes the perceived benefits of virtual care identified by participants.
Participation identified many aspects of virtual care that they perceived to be more challenging or less effective compared to in-person care. Many identified limitations they believed would impact others (eg, the challenges less tech-savvy older adults would face while accessing and using technology, lack of access to technology for all older adults) but maintained that virtual care was ideal for themselves and had few negative aspects. The most cited limitations of virtual care describe a lack of nonverbal communication (eg, facial expressions and body language) and limited opportunities for physical examination. Other limitations included challenges with older adults accessing and using technology, challenges with patients’ and doctors’ ability to express themselves verbally (eg, in telephone-only appointments), negative impacts on care coordination and continuity, and the potential exacerbation of the social isolation of older adults (ie, for some isolated older adults, in-person visits to primary care are an essential piece of their limited social lives). One participant expressed concerns about accommodations for older adults who require language interpretation services while accessing health care. Participants’ perceived limitations are presented in Table 3.

Participants highlighted an important caveat to our interpretation of the data: a benefit of virtual care for one older adult can be a limitation of virtual care for another. For example, although one person may appreciate the efficiency of a virtual care appointment, another may deeply miss the interpersonal and social interactions that accompany an in-person visit.
Table 3. Summary of the perceived limitations of virtual care use for older adults.

<table>
<thead>
<tr>
<th>Limitation</th>
<th>Description</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonverbal communication and body language</td>
<td>Usual forms of virtual care (ie, phone calls and video calls) eliminate nonverbal communication, including facial expressions and body language.</td>
<td>• “...you do miss some of the eye contact and the body language, and I make the point that when people communicate, they often talk about the words only being about 7%, the tone being 38% of the...body language being 55%, so email or a phone, you might get the tone but you don’t get the body language and that’s, and, sometimes that’s very important. I know, how many times that I noticed body language, that I would ask another question, and bingo, the real problem would come out, where it wouldn’t have come up if you hadn’t been able to observe the body language” [James, 75 years, male, B³]</td>
</tr>
<tr>
<td>Technology</td>
<td>Virtual care can be difficult to access for older adults who either do not have access to sufficient technology at home or do not know how to use the technology they have to engage in virtual care.</td>
<td>• “I worry about that. And for people who don’t have access! I mean everybody doesn’t have a computer at home, or they have a computer, but they barely don’t know how to use it. I talk to people, and they say, well, you know, my son will help me, my daughter will help me, my grandkids will help me. But other than that, they don’t know how—they don’t use it. Or they might use it just for...phone conversation...you know, for e-mail kind of stuff, and that’s it. And so, they don’t get it...they don’t get to use. They don’t have real access, and now—the library is being closed now. People have even less access.” [Helen, 77 years, female, B]</td>
</tr>
</tbody>
</table>
| Verbal communication                    | Participants are concerned about the ability of both patients and doctors to express themselves in virtual visits. They mentioned that people with cognitive or hearing difficulties may find accessing virtual care especially difficult due to the challenges of understanding. | • “May I didn’t explain it well enough to them. I’m not a nurse, you know, and I just know how it feels to me, I probably don’t have an experience and that, whatever is happening to me, this time when I would call, you know. So, it might be my terminology, my reporting might not be as good as they might need.” [Sharon, 82 years, female, B]  
• “Yes, and if they have hearing problems, that might be, a deterrent too because...or cognitive problems where they, have problems understanding” [Shirley, 77 years, female, B]  
• “We have a tendency to save the important questions ‘til the end. That’s a known fact, is that people are having a physical or whatever, you go through all the steps. And as they’re walking out the door they say, ‘Oh, by the way, I’ve been having chest pains’.” [Helen, 77 years, female, B] |
| Physical examination                    | Although participants believed that many health issues can be successfully discussed virtually, several participants also expressed concerns over the lack of physical and tactile examination during virtual appointments. The general concern behind this was that doctors would be more prone to accidentally missing something if the patient was not physically in front of them. | • “I just want to say that it is very limited. Sometimes, when you have a problem and you’re seeing a doctor, you want him to look, with his own eyeballs to see the actual thing. You...to see your skin, in the real thing not...not done through a camera, and you want him to poke you, you know, or feel. There’s so much in an examination, that should be done tactile, as opposed to only visual. Only visual, you miss so much without the tactile attached to it.” [Richard, 76 years, male, B] |
| Care continuity and coordination        | Participants felt that because virtual care usually meant that their provider was physically seeing them less often than they would with in-person care, the continuity of care and ability of physicians to coordinate care activities suffered. | • “The cons would be, perhaps a lack of follow up sometimes, because you’re not seeing anything done. Like if I go to his office and he gives me a referral to someone or if he’s...you just don’t see that referral happening through you, you see him doing it, and sometimes that doesn’t happen as quickly as it could.” [Nancy, 66 years, female, B] |
| Social isolation and health care as a social experience | Health care visits can be vital social experiences for older adults. Switching to a virtual format removes much of the social activity and personality from appointments. | • “I think if I really needed to see the doctor, she would let me go in and talk to her. And I think I still need that if I do need it. And that would be most likely for an emotional situation more than anything” [Judith, 75 years, female, F]  
• “And also, for some people the doctor’s visit is one of your social experiences. The more you live alone, like I live alone, these kinds of contacts are part of your...socialization, your contacts...like going to the library, going to the doctor, these are all things where people have contact with others. So, if you make these things more virtual, you cut back on people’s contacts with the outside world.” [Helen, 77 years, female, B] |

³B: baseline.
²F: follow-up.
Perspectives on When Virtual Care Is Acceptable and Appropriate

Subthemes related to participant perspectives on the acceptability and appropriateness of virtual care included receptiveness to virtual care in some scenarios, preferences for the future of virtual care, and the willingness of older adults to adapt to virtual care and the supports required for them to do so.

The Future of Virtual Care and the Preference for a Hybrid Model (Some In-Person)

When participants were asked how they would like the health care system to operate postpandemic, participants presented 2 main preferences: approximately half expressed their preference to return to a health care system that provides the majority of services in-person, while the other half preferred to retain some aspects of the COVID-19 era virtual care and reintroduce aspects of in-person care to create a hybrid system of health services.

Participants who preferred to return to an in-person model of health care were not necessarily opposed to the use of virtual care. Some agreed that, although virtual care was useful during the COVID-19 pandemic, they would prefer to access in-person care whenever possible:

> Something like prescription renewals will be convenient to have them continue through the pharmacy, to my doctor, that would be very convenient. But aside of that, I'd rather see my physician in person. [Nancy, 66 years, female, B]

Many participants expressed support for a hybrid health care model that includes aspects of both virtual and in-person care:

> But if it went back, it went back but in a modified way, like it's not all one or all the other. It's not all phone or all office, like it could be a mix. [Patricia, 82 years, female, B]

Helen, aged 77 years, expressed that although she preferred a hybrid model of care, it would need to be carefully organized and implemented to be effective:

> But it needs to be carefully thought of. And I've always been hesitant about virtual care, 'cause I don't want to see that as an instead...yes, virtual care yes, but it has to be in addendum. It has to be something in between. It's very useful to check up on something. [Helen, 77 years, female, B]

Although many participants supported a future that incorporated aspects of both virtual and in-person care models, they were concerned about how this would be funded at a system level, as well as enable access to the system using technologies with which older adults feel more comfortable. Although many participants noted a need for technology training and supports for older adults, several noted that efforts aimed at improving virtual care should also be focused on training for providers, not just patients:

> I think the...the...to take advantage of those types of situations I think technology use, I think somebody should actually encourage the GP and their registered...their nurses or their receptionist to be more proficient in these technologies. I think seniors when there's a need, they'll do anything to learn it. [Geraldine, 72 years, female, B]

Although many participants were reconciled to virtual care being a major component of their health care in the future, they saw a need for related training and support for both providers and patients.

Discussion

Principal Findings

We identified 3 high-level themes in our interviews with older adults about their virtual care use during the pandemic. Older adults shared (1) their experiences with virtual and in-person health care during the pandemic, (2) their thoughts on the benefits and limitations of virtual care, and (3) their opinions on when virtual care is most appropriate. Consistent with the results of the AGE-WELL [14] survey, most of the participants in our study experienced some form of virtual care access, primarily via telephone or online, with fewer participants having accessed care via video. Participants expressed reluctance to attend in-person visits during the pandemic, with in-person care accessed mainly in emergencies or for services that were not available virtually.

Comparison With Prior Work

Prepandemic studies of virtual care (eg, [18]) have found both benefits and limitations; this was also the case for the participants in this study. Importantly, most participants felt they were able to maintain their patient-doctor relationship despite the change in the mode of care delivery, thus alleviating some of the concern raised by Sendorovich and Wignarajah [13] about maintaining the quality of the therapeutic alliance. Our study participants described the convenience of virtual visits as well as increased safety, including the avoiding of unsafe travel conditions, similar to findings by Elliott et al [29]. Participants also felt virtual care would be more time efficient for the provider, but we note that some studies of virtual care have not found cost-saving benefits [30]. In contrast, our study participants recognized the lack of physical, hands-on examinations in a virtual care appointment, as was also found in studies by Breton et al [31] and Mao et al [32]. Participants also noted limitations in terms of compromises in both verbal and nonverbal communication. This is consistent with the finding of Hamersley et al [12] that there was less information sharing in virtual visits, although these authors also noted that...
the virtual care visits they studied revealed somewhat greater efforts toward building rapport with the patient.

Some limitations of virtual care might be mitigated if video-and-voice appointments are used. Although video appointments are not a complete substitute for a tactile examination of the body, they enable synchronous visual examination, which may help alleviate patient concerns that providers may miss something. Video calls may also enable both patients and providers to interpret nonverbal cues and facial expressions more accurately. Conversely, it may be more difficult for providers to see the patient in a video call compared to a photo of an affected area (sent via email or text) due to the wide range of devices that older adults and clinicians use, variations in connectivity or access to a reliable internet connection among patients, and compatibility between devices. In addition, video visits make additional demands on the patient, who must be able to get online and manage the technology, which may be difficult due to disability or lack of experience with technology or a stable internet connection [32,33].

Limitations of This Study

First, our study is bound to a specific time and rooted in the perspectives of a small sample of older Canadians and as such may not be readily transferable to other settings. We recognize this as a limitation and that our findings only reflect the perspectives of the 20 interviewees. Future research with larger samples of older adults is warranted. Second, we recognize our sample is undoubtedly overrepresentative of individuals who have the interest, access, and privilege to engage in new technologies. Although we specifically sought out individuals from a range of cohorts, living arrangements, and ethnic groups, our recruitment strategies (which had to be mindful of social distancing) mostly connected us with privileged individuals who were already online, had access to email, and were able to complete a voluntary research interview (ie, they had the time and interest to do so and, at the very least, had a telephone). These advantages will be reflected in our results, and this body of literature would greatly benefit from more work with older adults who do not/cannot use technology to receive care. In the future, recruitment options that do not rely on newer technologies should be used to connect with individuals who are less tech-savvy (eg, radio, mail-based, and in-person recruitment).

Future Directions

Falk [34] has argued that virtual care may reduce inequities for some older persons, such as those living in remote communities, but at the same time might exacerbate inequities through avoiding direct service to these regions. Future research, including that of this team, must actively reach out to support those older adults on the underrepresented side of the digital divide [35-37], particularly as the United Nations calls for all nations to close these digital divides [38]. As suggested by participants in this study, support strategies should target both older adults and providers; Chen et al [39] found that training geriatric care professionals on virtual care technologies prior to the pandemic helped ease the transition to virtual care. Multiple virtual care resources have been designed for older adults in Canada, including appointment checklists (eg, [5]) and supportive liaisons to help navigate particular technologies, as implemented at Women’s College Hospital [40]. Technology-based interventions can also improve access for marginalized groups with less technical experience by simplifying user interfaces and workflows on virtual care platforms to increase usability [40,41]. Efforts should be made to collaborate with older adults when designing and implementing such strategies in order to maximize their usefulness and relevance [40,42,43]. This can be accomplished by engaging older adults in designing technology and virtual care systems, training providers, and research/program evaluation (eg, through advisory committees, participatory research, codesign, etc) [40].

Conclusion

The COVID-19 pandemic has been a major catalyst for the adoption of virtual care in Canada [6]. Our study confirmed that the potential benefits of virtual care for older adults are numerous; despite barriers to accessing virtual care, many older adults perceive benefits and are open to continued use of virtual care after the pandemic. Our study also found many limitations of virtual care, and a consensus that virtual care should be an addendum to the health care system, rather than its main delivery mechanism. These findings would thus call into question policies, such as the United Kingdom’s National Health Service (NHS) plan for digital-first primary care for every patient [44]. As we transition to a postpandemic world, older adults must be included in discussions on the design and implementation of virtual care options. Concerns related to privacy and confidentiality have been highlighted in other studies [6,45] but were not significantly present in our findings; this could be an explanation for why some older adults were less comfortable accessing virtual care.

In this study, we presented data from a small sample of older adults from Canada detailing their experiences with virtual care during the pandemic, their perceptions on the benefits and limitations of virtual care, and their willingness to engage in virtual care. Future dissemination of virtual care options should ensure that older adults’ views, preferences, and circumstances are considered and that accommodations are made for those whose use of virtual care is limited by disability or discomfort with the technology. The findings can also be used to inform future studies on the use of virtual care by older adults, as providers and patients continue to adapt to both the potential and pitfalls of this mode of care delivery.

Acknowledgments

To the older adult participants and our “Seniors Helping as Research Partners” group, thank you for all taking the time to share your experiences and perspectives. This project was funded by the University of Waterloo’s Network for Aging Research Catalyst Grant program (NARCATS2020). Ms Abdallah’s work was supported by the University of Waterloo’s Hallman Fellowship.
Data Availability

Anonymized data sets generated or analyzed during the study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide.

References


Abbreviations
BIPOC: Black, Indigenous, and people of color

©Lama Abdallah, Paul Stolee, Kimberly J Lopez, Alexandra Whate, Jennifer Boger, Catherine Tong. Originally published in JMIR Aging (https://aging.jmir.org), 20.10.2022. 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.
Usability and Acceptability of a Palliative Care Mobile Intervention for Older Adults With Heart Failure and Caregivers: Observational Study

Jennifer Paola Villalobos¹, MS; Sheana Salyers Bull¹, MPH, PhD; Jennifer Dickman Portz², MSW, PhD

¹Colorado School of Public Health, University of Colorado, Aurora, CO, United States
²Division of General Internal Medicine, University of Colorado, Aurora, CO, United States

Corresponding Author:
Jennifer Paola Villalobos, MS
Colorado School of Public Health
University of Colorado
13055 E 17th Ave
Mail Stop #F802
Aurora, CO, 80045
United States
Phone: 1 303 551 3823
Email: jennifer.p.villalobos@gmail.com

Abstract

Background: Heart failure is a leading cause of death among older adults. Digital health can increase access to and awareness of palliative care for patients with advanced heart failure and their caregivers. However, few palliative care digital interventions target heart failure or patients’ caregivers, family, and friends, termed here as the social convoy. To address this need, the Social Convoy Palliative Care (Convoy-Pal) mobile intervention was developed to deliver self-management tools and palliative care resources to older adults with advanced heart failure and their social convoys.

Objective: The goal of the research was to test the acceptability and usability of Convoy-Pal among older adults with advanced heart failure and their social convoys.

Methods: Convoy-Pal includes tablet-based and smartwatch tools facilitating self-management and access to palliative care resources. Older adults and social convoy caregivers completed an acceptability and usability interview via Zoom, including open-ended questions and the Mobile Application Rating Scale: User Version (uMARS). Descriptive analysis was conducted to summarize the results of open-ended feedback and self-reported acceptability and usability.

Results: A total of 26 participants (16 older adults and 10 social convoy caregivers) participated in the interview. Overall, the feedback from users was good (uMARS mean 3.96/5 [SD 0.81]). Both older adults and social convoy caregivers scored information provided by Convoy-Pal the highest (mean 4.22 [SD 0.75] and mean 4.21 [SD 0.64], respectively). Aesthetics, functionality, and engagement were also perceived as acceptable (mean >3.5). Open-ended feedback resulted in 5 themes including improvements to goal setting, monitoring tools, daily check-in call feature, portal and mobile app, and convoy assessment.

Conclusions: Convoy-Pal was perceived as acceptable with good usability among older adults with heart failure and their social convoy caregivers. With good acceptability, Convoy-Pal may ultimately lead to increased access to palliative care resources and facilitate self-management among older adults with heart failure and their social convoy caregivers.

(JMIR Aging 2022;5(4):e35592) doi:10.2196/35592

KEYWORDS

mHealth; older adult; symptom; heart failure; palliative care; app; digital health; cardiology; heart; Convoy-Pal; mobile; tablet; smartwatch; adult; aging
Introduction

Heart failure (HF) is the 4th leading cause of death from heart disease in the United States and is most prevalent among individuals aged 65 years and older (ie, older adults). According to data from 2015 to 2018, 7.5% of males and 3.9% of females aged 60 to 79 years have HF [1]. The prevalence of HF continues to rise over time as the population ages [2]. By the year 2040, the number of older Americans is expected to nearly double to an estimated 80.8 million [3]. As the prevalence of HF increases, the need for palliative care amplifies. Palliative care can be beneficial for patients with HF as well as their caregivers, families, friends, and loved ones, referred to here as the social convoy [4]. Palliative care offers a support system to help the social convoy cope during the patient’s illness and effectively control distressing symptoms experienced by patients with HF [5]. In general, symptom control and good communication are basic palliative care principles highly recommended to improve the quality of life for patients with HF [5]. Although relatively underexplored, digital health [6] innovations (ie, telehealth, wearable devices, and mobile health [mHealth]) provide modern opportunities for patients and their social convoy to engage in palliative care [7-11].

Although there is a need, few studies focus on HF-specific mHealth in palliative care or mHealth supports for the social convoy. A systematic review of mHealth in palliative care reports that the primary uses of mobile apps are for biological and clinical monitoring (75% of the apps), disease self-management (64% of the apps), and therapeutic patient education (50% of the apps) [12]. One pilot in the review targets patients with HF and has found that the use of the HF mobile app improves self-care management [13]. Another study involving HF patients and their informal caregivers shows mHealth may decrease risk of HF exacerbations and improve caregiver communication [14]. While there are early indicators that patients and caregivers benefit from mHealth, providers also express enthusiasm about the potential of mHealth in palliative care [15-17]. Palliative care providers recommend digital health innovations in the areas of telehealth, client health records, and personal health tracking [17].

Quality testing in mHealth includes acceptability and usability as standard and essential in the field. Acceptability testing is usually completed first, followed by usability testing. This type of testing, for example, allows researchers to increase confidence that subsequent research on the efficacy of a tool produces outcomes that ensure null or negative outcomes are not due to poor tool function. Essentially, acceptability testing in mHealth assists with determining the level of meaningful engagement with the app; otherwise, if not engaging, the app will not be used, which may affect retention over time [18]. Usability testing, on the other hand, highlights the need to adapt the apps to users’ needs to create more usable tools [19] and ensure an app can be used the way it was intended by the specific audience for the tool [20].

Given the limited access to HF-specific palliative care mHealth, the Social Convoy Palliative Care Mobile Intervention, Convoy-Pal, was developed in response to a need for self-care strategies for both older adults with HF and their social convoys. Convoy-Pal was co-designed with older adults, caregivers, and health care providers [21] under the clinical guidelines establish by the National Coalition for Hospice and Palliative Care [22]. As a step in the co-designing process, this study was to test the acceptability of Convoy-Pal among older adults with HF and their caregivers.

Methods

Convoy-Pal Platform

The authors are researchers at the mHealth Impact Lab [23] who contract with Routinify, Inc, [24], a vendor that delivers the Convoy-Pal intervention. Routinify offers a variety of software and hardware tools that are publicly available; costs vary based on the tools provided and can range from US $50 to $100 per patient. In this case, Routinify assists with the delivery of the Convoy-Pal intervention to older adults and their social convoys. However, Routinify is only permitted to deliver Convoy-Pal in contract with the mHealth Lab and is not engaged in the clinical research (ie, they are not involved in the study instruments, data collection, management, analysis, or designing the protocol).

Convoy-Pal is designed with the following care domains: physical, psychological, social, spiritual, near end of life, ethical and legal, and knowledge about palliative care overall. Convoy-Pal includes a palliative care assessment with self-monitoring and resource tools for each domain. For example, the near end-of-life aspects of care (Figure 1) includes information regarding grief support and self-care and provides an opportunity for life review activities. This also includes resources on how to communicate unaddressed concerns and identify a support group for social support. Convoy-Pal tools and content are designed to be delivered via WellAssist by Routinify, Inc. WellAssist is a personal point-of-care app and associated internet-connected medical devices. The app’s core is based on behavioral modifications in line with the overall plan of care. The app is designed so that all members within the social convoy can access and use the Convoy-Pal intervention.
Ethics Approval and Considerations

Study procedures were approved by the Colorado Multiple Institutional Review Board (number 18-0973). All participants electronically consented to participate. Convoy-Pal collects assessments, including smartwatch-captured vital information, regarding mental health and overall well-being data for the older adult patient and caregivers. It is ethically sound to obtain consent from older adults to share their health information with their caregivers as we did for this study; however, older adult patients also have the option not to share their health information with their caregivers if desired.

Recruitment

Potential patient participants were identified from the UCHealth University of Colorado Hospital health system’s electronic medical record. Potential participants were aged at least 65 years at the time of recruitment and had been hospitalized at the UCHealth Hospital more than 2 times for HF in the year prior (January 2020-2021). Participants were currently living in their homes and receiving follow-up HF care. We mailed a study invitation letter with an opt-out contact option. Patients who did not opt out were then contacted by phone for recruitment and asked to self-identify social convoy caregivers.

Data Collection

Two research coordinators (JPV and IM) interviewed participants via Zoom to gain feedback on Convoy-Pal. Participants were exposed to sections of Convoy-Pal throughout the interview process, which lasted between 40 minutes to 1 hour. Participants were shown the Convoy-Pal hardware, which consisted of the tablet, watch, and charging station, and the web-based system portal and mobile app during the interview. Participants were also shown Convoy-Pal features such as goal setting and planning, monitoring options, daily check-in and calling features, convoy caregiver assessments, and palliative care resources. During the exposure to Convoy-Pal, participants completed a self-report acceptability measure and were asked to provide open-ended feedback.

For self-report acceptability, participants completed the Mobile Application Rating Scale: User Version (uMARS) survey [25]. The uMARS survey comprises 4 objective quality subscales—(1) engagement with the app, (2) functionality and users’ perceived functioning of the app, (3) aesthetics, and (4) users’ perception of the quality of the information [25]—to determine app quality mean score. uMARS has 2 optional subscales that can be used depending on the aims of the research. These 2 subscales are the app subjective quality scale, which can be reported as individual items, and the perceived impact scale, which obtains information on the knowledge, attitudes, and behavior change toward improving health behavior [25]. All items are assessed on a 5-point scale, with a uMARS score of 5 considered excellent while a score of 1 is inadequate [26]. The uMARS is shown to have high interrater reliability for evaluating the quality of mHealth apps on well-being, for example [25,27].

For open-ended feedback, the tablet and smartwatch were introduced and displayed over Zoom to the participants. We used a semistructured interview guide (Multimedia Appendix 1) to ask the participants questions and their opinions regarding the hardware, goal setting and planning, monitoring options, daily check-in and convoy calling options, portal and mobile app for convoy, convoy assessments, and convoy resources. Notes, recommendations, and opinions from participants were archived into Qualtrics (Qualtrics) [28], data management software, as the interview was being conducted. The final data captured were stored and saved in Qualtrics with their study ID numbers.

Data Analysis

The uMARS survey data was analyzed using Excel (Microsoft Corp) calculation mechanisms and descriptive frequencies including mean scores for both caregivers and patients. Once
all the interviews were complete, the qualitative data were moved from Qualtrics to NVivo 12 (QSR International) [29], a qualitative research software package, for analysis. A preliminary codebook was created, incorporating explicit domains from the interview guide (deductive themes) by a research assistant (JPV). A descriptive qualitative approach [30] was then used to identify themes and subthemes [31,32]. The codebook and final data interpretation were discussed in a group with all authors. No member checking was conducted.

Results

Participants

We recruited 26 participants (16 patients and 10 caregivers) from the University of Colorado Denver and its affiliate, the University of Colorado Hospital. Patients and convoy caregivers participated together or separately depending on the patient’s ability to participate in the interview. Patients were primarily males (9/16, 56%), White (14/16, 88%), and had a mean age of 76 (SD 5) years. Caregivers were predominantly female (7/10, 70%), White (7/10, 70%), and had a mean age of 71 (SD 10) years. Patients were married (12/16, 75%) and had a postgraduate degree (8/16, 50%), with 44% (7/16) having an income of US $30,000 or more and 82% (13/16) owning an iPhone, Android, or a regular or basic phone (Table 1). Similarly, most caregivers were married (8/10, 80%) and had a college or postgraduate degree (7/10, 70%), with 50% (5/10) making US $30,000 or more; 40% (4/10) of caregivers chose not to answer the question regarding their income. All of the caregivers owned an iPhone, Android, or a regular basic phone (Table 1). Due to small cell sizes, demographic categories were collapsed and are not reported in the table.

Table 1. Participant technology use.

<table>
<thead>
<tr>
<th>Technology use</th>
<th>Patients, n (%) (n=16)</th>
<th>Caregivers, n (%) (n=10)</th>
<th>Total, n (%) (N=26)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell phone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic phone: iPhone, Android, or regular or basic phone</td>
<td>13 (82)</td>
<td>10 (100)</td>
<td>23 (84)</td>
</tr>
<tr>
<td>I do not have a cell phone</td>
<td>1 (&lt;1)</td>
<td>0</td>
<td>1 (&lt;1)</td>
</tr>
<tr>
<td>Did not respond</td>
<td>2 (13)</td>
<td>_a</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Digital activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Email</td>
<td>16 (100)</td>
<td>10 (100)</td>
<td>26 (100)</td>
</tr>
<tr>
<td>Look up information</td>
<td>16 (100)</td>
<td>10 (100)</td>
<td>26 (100)</td>
</tr>
<tr>
<td>Use social media</td>
<td>16 (100)</td>
<td>6 (60)</td>
<td>22 (84)</td>
</tr>
<tr>
<td>Post and share pictures or videos</td>
<td>15 (94)</td>
<td>9 (90)</td>
<td>24 (92)</td>
</tr>
<tr>
<td>Read or post comments</td>
<td>15 (94)</td>
<td>10 (100)</td>
<td>25 (96)</td>
</tr>
<tr>
<td>Play computer games</td>
<td>14 (88)</td>
<td>10 (100)</td>
<td>24 (92)</td>
</tr>
<tr>
<td>Video chat</td>
<td>16 (100)</td>
<td>10 (100)</td>
<td>26 (100)</td>
</tr>
<tr>
<td>Instant message or chat rooms</td>
<td>13 (82)</td>
<td>7 (70)</td>
<td>20 (76)</td>
</tr>
</tbody>
</table>

Not applicable.

Acceptability

Mobile Application Rating Scale: User Version

Overall, the acceptability feedback from users was good. The uMARS mean score was 4.00 (SD 0.78) among patients and 3.92 (SD 0.83) among caregivers, with an overall uMARS mean score of 3.96 (SD 0.81) among both groups (Table 2). Patients and caregivers showed the most concordance with Section D: information scale and the most discordance with Section C: aesthetics (Table 2). Further description of the mean, standard deviation, and minimum and maximum values for the subscales of the uMARS are provided in Multimedia Appendix 2.

Examining uMARS domain scores individually, we found that patients gave Section D: information the highest rating (mean 4.22, SD 0.75), followed by Section C: aesthetics (mean 4.13, SD 0.73), Section B: functionality (mean 3.87, SD 0.85), and Section A: engagement (mean 3.80, SD 0.79). Patients scored the app’s subjective quality scale a mean of 4.01 (SD 0.70) and the perceived impact of the app on the user’s knowledge, attitudes, and intentions related to the target health behavior a 3.64 (SD 0.96). Similarly, caregivers scored Section D: information the highest (mean 4.21, SD 0.64), followed by Section C: aesthetics (mean 3.89, SD 0.72), Section B: functionality (mean 3.82, SD 1.0), and Section A: engagement (mean 3.77, SD 0.96). The app subjective quality scale was rated mean 3.56 (SD 1.23) and perceived impact was rated mean 3.13 (SD 1.20) among caregivers.
Table 2. Mean, standard deviation, and range values for the subscales of the uMARS (Mobile Application Rating Scale: User Version).

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Caregivers</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Section A: engagement</td>
<td>3.80 (0.79)</td>
<td>2.25-5.00</td>
</tr>
<tr>
<td>Section B: functionality</td>
<td>3.87 (0.85)</td>
<td>2.00-5.00</td>
</tr>
<tr>
<td>Section C: aesthetics</td>
<td>4.13 (0.73)</td>
<td>3.33-5.00</td>
</tr>
<tr>
<td>Section D: information</td>
<td>4.22 (0.75)</td>
<td>2.50-5.00</td>
</tr>
<tr>
<td>Total</td>
<td>4.00 (0.78)</td>
<td>—</td>
</tr>
</tbody>
</table>

^aNot applicable.

**Open-Ended Feedback**

Five main themes were identified after receiving open-ended feedback: goal setting, monitoring tools, daily check-in call feature, portal and mobile app, and convoy assessment. Representative quotes for themes and additional subthemes are reported in Table 3.

Table 3. Participant feedback (N=23).

<table>
<thead>
<tr>
<th>Theme</th>
<th>Subthemes</th>
<th>Representative quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal setting</td>
<td>Obtainable goals</td>
<td>“Questions should be addressed monthly not weekly.” [72-year-old participant]</td>
</tr>
<tr>
<td></td>
<td>Comment section</td>
<td>“Provide fill-in-the-blank options.” [75-year-old participant]</td>
</tr>
<tr>
<td>Monitoring tools</td>
<td>Added features</td>
<td>“Would like to see prompting feedback if things are not okay.” [84-year-old participant]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Add ways to detect stroke and falls down the stairs.” [74-year-old participant]</td>
</tr>
<tr>
<td>Reminders</td>
<td>Checklist</td>
<td>“Design a checklist of all of the medication a person takes for specific medication notifications versus getting general messages.” [71-year-old participant]</td>
</tr>
<tr>
<td>Portal and mobile app</td>
<td>Thresholds</td>
<td>“Would like to see thresholds on the graphs to determine who should be consulted.” [74-year-old participant]</td>
</tr>
<tr>
<td>Convoy assessment</td>
<td>Wrong approach</td>
<td>“if [a caregiver] is in crisis mode, they will not fill out the questions...this is not beneficial for patients who need extra help and support.” [72-year-old participant]</td>
</tr>
</tbody>
</table>

**Goal Setting**

Participants expressed the need for the goal-setting section to provide realistic and obtainable goals. For example, it was expressed that goal setting should be addressed monthly, not weekly. Additionally, participants expressed that they would like an option to add a comment box to include other action items and or commentary for their goals.

**Monitoring Tools**

Participants agreed that the monitoring tools were helpful for people with HF and other chronic conditions. Feedback to improve the monitoring tools included the addition of other features, such as feedback prompting concerning vitals, electrocardiogram measures, fall detection, stroke indicators, and heart palpitation monitoring.

**Reminder and Call Feature**

Most of the participants liked the daily check-in and call feature. One participant said “...the feature is good for people who live alone and want to keep in contact via FaceTime with their loved ones” (72-year-old participant). A few participants who disliked the daily check-in feature expressed that some patients might feel burdened by the frequency of check-ins. Others indicated the feature was redundant as they could schedule reminders and calls through their personal phone instead. One suggestion included designing a checklist of all of the medication a person takes and getting notifications on those specific medications versus just getting a general message.

**Portal and Mobile App**

The majority of the participants liked the portal and mobile app. The participants appreciated that the portal, charts, and layout of the mobile app were clear and concise. Participants also liked the opportunity to share access to personal data with family members. Feedback from 2 participants included adding thresholds to the graphs to determine who, such as a provider or family member, should be consulted, and adding instructions on who to call with concerns.

**Social Convoy Assessment**

Participants provided many recommendations when asked their opinions about convoy assessment. Many participants were hesitant about the caregiver assessments due to time, burden, and specific assessment topics. For example, in the domain of emotional assessment, a participant said “if [a caregiver] is in [in] crisis mode they will not fill out the questions...this is not beneficial for patients who need extra help and support” (Table 3). Another participant expressed that “they would not answer those questionnaires truthfully because they were not raised to share emotions growing up” (Table 3). Additionally, a
participant expressed that some assessments should be addressed in person and not via the tablet. Other themes that arose from the feedback included ensuring assessments are HIPAA (Health Insurance Portability and Accountability Act) compliant and appropriate language is used. For example, one participant responded, “The ‘I feel sad’ language might not be appropriate for people because they are not readily going to admit that they are sad” (71-year-old participant). Participants also suggested that the assessments should not take too long to complete.

**Discussion**

**Principal Findings**

Convoy-Pal is designed to increase access to palliative care resources and self-management in the setting of HF. Acceptability testing is essential because it results in a better quality product. In this acceptability and usability study, the Convoy-Pal is considered acceptable and a good quality app, based on the uMARS scores among older adults with HF and their caregivers. Older adult patients and caregivers also provided recommendations for improving Convoy-Pal, which included adding comment sections, designing a checklist for medications, including thresholds on graphs for interpretation, and adding features such as fall detection. Based on this feedback, authors will update and continue to assess Convoy-Pal for usability and feasibility.

Although there are few high-quality HF mobile apps [33] assessed for acceptability, functionality, and efficacy [33,34], our findings are supported by other empirical studies. For example, palliative care patients found a mobile mortality risk tool acceptable to use [35]. Additionally, using wearables for monitoring palliative care was also feasible [36,37], a tool that Convoy-Pal offers with its smartwatch. Similar to the feedback provided for Convoy-Pal, a commentary article, systematic meta-review, and qualitative study [15,38,39] report the need to track relevant information, receive education pertinent to health for older adults, and provide information sharing such as medication use.

Aside from HF apps specifically, digital health interventions overall have the potential to improve the accessibility and effectiveness in palliative care, as reported by a recent systematic meta-review [38]. Palliative care is one area where technologies are increasingly being deployed. Although leveraging existing resources for palliative care is one approach, mHealth interventions targeting palliative care enable patients increased access to this resource without spending time or traveling to locations [40]. mHealth palliative care allows older adults to participate in and govern their care. For example, they do this by self-reporting symptoms and needs, which improves communication with providers and caregivers [38,41,42]. Traditional palliative care resources do not provide a self-governing element in this unique way.

Additionally, HF mobile interventions rarely target the social convoy or palliative care domains. Therefore, Convoy-Pal would contribute to the advancement of palliative care and HF mHealth while also advancing a team approach to information sharing and targeting family- and caregiver-specific issues [43]. Convoy-Pal has the potential to support older adults with HF and their social convoy in the management of physical, psychosocial, and spiritual concerns.

**Limitations**

First, due to university and state COVID-19 restrictions, research assistants were not able to meet with older adult and caregiver participants to physically interact with Convoy-Pal on the Routinify tablets or complete assessments in person. Researchers therefore collected acceptability and usability data by displaying Convoy-Pal and all its features remotely to participants for about 1 hour through Zoom. Second, the uMARS survey, for this reason, was modified by our team to reflect the following 2 optional responses for all subscales of the survey: (1) “Optional: Missing due to lack of time with app” and (2) “Optional: Did not feel comfortable answering.” If the participant did not feel comfortable answering the uMARS questions due to their belief that there was not enough time to explore Convoy-Pal, then they could select either optional response. The minor modifications made to the uMARS had not previously been tested and therefore may have reduced the validity of the original items. Physical interaction with the hardware may have yielded additional user feedback. Last, the study was further limited by small sample size and a single health system as well as lack of diversity representative of the local community. Acceptability and usability of Convoy-Pal may differ in other regional areas and varying access to health care.

**Conclusion**

HF is a leading cause of death in the United States, and mHealth provides opportunities for patients and their social convoy to participate in palliative care. In our study, 16 older patients and 10 caregivers were interviewed and asked to complete a uMARS assessment and provide open-ended feedback. Overall, older patients and their caregivers perceived Convoy-Pal as acceptable with good usability. Although in-person usability testing is needed, Convoy-Pal was perceived acceptable and may ultimately increase access to palliative care resources and facilitate self-management among older adults with HF and their caregivers.

**Acknowledgments**

The authors thank Routinify, Inc, for its assistance with the development and assessment of the Convoy-Pal. We also thank research assistant Irazema Mino for assistance conducting some qualitative interviews. This study was funded by grant K76AG059934 from the National Institute on Aging.
Data Availability
The data sets generated and analyzed during this study are not yet publicly available. The full study protocol and data sets will be shared to a public data repository upon the conclusion of the Convoy-Pal feasibility study.

Authors' Contributions
JDP, as principal investigator, supervised the project and manuscript, while JPV collected the data, conducted the analysis, and drafted the manuscript. SSB provided general feedback and suggestions on the direction and context of the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Convoy-Pal Acceptability Interview Guide.
[DOCX File, 26 KB - aging_v5i4e35592_app1.docx ]

Multimedia Appendix 2
Mean, standard deviation, and minimum and maximum values for the scales and subscales of the uMARS (Mobile Application Rating Scale: User Version).
[DOCX File, 17 KB - aging_v5i4e35592_app2.docx ]

References


Abbreviations

Convoy-Pal: Social Convoy Palliative Care
HF: heart failure
mHealth: mobile health
uMARS: Mobile Application Rating Scale: User Version
Original Paper

Mobile App Prototype in Older Adults for Postfracture Acute Pain Management: User-Centered Design Approach

Kevin Tran-Nguyen1, MSc; Caroline Berger1,2, BSc; Roxanne Bennett1, MSc; Michelle Wall1, MSc; Suzanne N Morin1*, MSc, MD; Fateme Rajabiyazdi3*, MSc, PhD

1Research Institute of the McGill University Health Centre, Montreal, QC, Canada
2University of Maryland, College of Information Studies, College Park, MD, United States
3Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, Canada

*these authors contributed equally

Corresponding Author:
Fateme Rajabiyazdi, MSc, PhD
Department of Systems and Computer Engineering
Carleton University
1125 Colonel By Dr
Ottawa, ON
Canada
Phone: 1 (613) 520 2600 ext 8229
Email: fateme.rajabiyazdi@carleton.ca

Abstract

Background: Postfracture acute pain is often inadequately managed in older adults. Mobile health (mHealth) technologies can offer opportunities for self-management of pain; however, insufficient apps exist for acute pain management after a fracture, and none are designed for an older adult population.

Objective: This study aims to design, develop, and evaluate an mHealth app prototype using a human-centered design approach to support older adults in the self-management of postfracture acute pain.

Methods: This study used a multidisciplinary and user-centered design approach. Overall, 7 stakeholders (ie, 1 clinician-researcher specialized in internal medicine, 2 user experience designers, 1 computer science researcher, 1 clinical research assistant researcher, and 2 pharmacists) from the project team, together with 355 external stakeholders, were involved throughout our user-centered development process that included surveys, requirement elicitation, participatory design workshops, mobile app design and development, mobile app content development, and usability testing. We completed this study in 3 phases. We analyzed data from prior surveys administered to 305 members of the Canadian Osteoporosis Patient Network and 34 health care professionals to identify requirements for designing a low-fidelity prototype. Next, we facilitated 4 participatory design workshops with 6 participants for feedback on content, presentation, and interaction with our proposed low-fidelity prototype. After analyzing the collected data using thematic analysis, we designed a medium-fidelity prototype. Finally, to evaluate our medium-fidelity prototype, we conducted usability tests with 10 participants. The results informed the design of our high-fidelity prototype. Throughout all the phases of this development study, we incorporated inputs from health professionals to ensure the accuracy and validity of the medical content in our prototypes.

Results: We identified 3 categories of functionalities necessary to include in the design of our initial low-fidelity prototype: the need for support resources, diary entries, and access to educational materials. We then conducted a thematic analysis of the data collected in the design workshops, which revealed 4 themes: feedback on the user interface design and usability, requests for additional functionalities, feedback on medical guides and educational materials, and suggestions for additional medical content. On the basis of these results, we designed a medium-fidelity prototype. All the participants in the usability evaluation tests found the medium-fidelity prototype useful and easy to use. On the basis of the feedback and difficulties experienced by participants, we adjusted our design in preparation for the high-fidelity prototype.

Conclusions: We designed, developed, and evaluated an mHealth app to support older adults in the self-management of pain after a fracture. The participants found our proposed prototype useful for managing acute pain and easy to interact with and navigate. Assessment of the clinical outcomes and long-term effects of our proposed mHealth app will be evaluated in the future.
Background
The rate of incidence of fractures is increasing globally, particularly in older adults [1]. Fractures are associated with acute pain, loss of autonomy, anxiety, recurrent falls, and, in a significant proportion, transition to chronic pain syndromes [2-4]. Furthermore, 74% of the patients visiting the emergency department, including those who present with at least moderate pain after a fracture or dislocation, are discharged with moderate to severe pain [5].

Multimodal approaches, including the use of medications, restorative therapies, and behavioral and complementary health approaches, are recommended for the management of acute pain conditions, including pain experienced following a skeletal fracture [6,7]. Despite a series of guidelines established by the US Institute of Medicine and the American Pain Society to manage pain [2,3], acute pain is underrecognized and undertreated in older adults both in and out of the hospital setting, leading to negative clinical outcomes [2]. In response to inadequate outpatient acute pain management, the US Department of Health and Human Services has emphasized the need for individualized self-management programs to support older patients in coping with and reducing their pain through pharmacological and nonpharmacological methods [6].

With the increased availability and use of mobile health (mHealth) tools on smartphones that are commercially available to users, there is a growing opportunity to develop mobile apps offering individualized pain management. The types of mHealth mobile apps that have been shown to have common clinical value offer at least one of the following: supporting clinical diagnosis, promoting behavior change and increasing patient adherence with treatment plans, supporting self-management of a condition, or delivering disease-related education [8]. However, systematic and scoping reviews have found that most mHealth apps were developed for chronic pain rather than for acute pain management [9-12]. A recent scoping review focused on mHealth in the context of surgery found that out of 13 studies, only 5 focused on addressing postsurgery acute pain [13]. These 5 studies aimed to reduce postoperative pain in patients by monitoring opioid use [14-16] and encouraging therapeutic adherence via smartphone functionalities such as alarms and accelerometers [17,18]. Among the mobile apps publicly available for iOS and Android devices, options catering specifically to acute pain are limited. A systematic review of commercially available pain management apps recommended 3 apps: Curable, Pathways, and Vivify; however, all of these apps were designed for chronic pain [19]. The systematic review found Achy Penguin to be the only available app in Canada that specifically manages acute pain, but it is designed for young children and does not fulfill the needs of older adults [19].

Previous studies on the use of mHealth to manage pain were promising in improving pain outcomes, but more research is required in this field, as many mHealth apps remain unvalidated by scientific means [12,20-22].

In addition, most of the pain management mHealth apps are designed without the involvement of health care providers [12] and older adults, resulting in apps that are ill-suited for an older audience [9,22]. Thus, there is a gap in the availability of innovative evidence-based mHealth tools and solutions to support older patients in the management of their acute pain once they leave the hospital.

Objectives
This study aimed to design, develop, and evaluate an mHealth app prototype to support older adults in the self-management of postfracture acute pain using a human-centered design (HCD) approach, which involves focusing on understanding the context of use, needs, and problems of the end users to develop the technological solution [23,24]. The main focus of this app is on medication management and adherence with support for other pain management needs such as educational materials and external available resources. The novelty of our work lies in the design and development process of the mHealth app, as we established evidence-based design requirements for our prototype and included older adults and health care professionals in the process.

Methods
Ethics Approval
Approval was obtained from the research ethics board of the Research Institute of the McGill University Health Centre (approval number 2021-7611), and all participants provided informed consent.

Overview
We reviewed the literature and conducted an informal competitive analysis to identify mobile apps available in the literature and on the market to manage postsurgery acute pain. We did not find any apps that were deemed suitable and clinically valid for older adults to manage their pain after a fracture. Thus, we aimed to develop a high-fidelity mHealth prototype app. We used an HCD approach to ensure that the end product was effective and efficient for the target users [25].

First, we identified the design requirements for a low-fidelity prototype using the results from prior surveys of members of the Canadian Osteoporosis Patient Network and clinicians. Next, we facilitated a series of participatory design workshops with older adults who had experienced a fracture and iteratively developed a medium-fidelity prototype. To find areas for improvement and gather evidence on the usability of our medium-fidelity prototype, we conducted usability tests, a method for hunting design and interaction problems in an
interface [24]. Finally, we designed a high-fidelity prototype based on the results of the usability tests (Figure 1). In addition, throughout the low- and medium-fidelity prototype phases of the study, we validated the content of our app through one-on-one discussions with pharmacists in our network.

Figure 1. The iterations of our mobile health app from (A) sketches, (B) low-fidelity prototype, (C) medium-fidelity interactive prototype, and (D) high-fidelity interactive prototype. The pages shown below display the logging pain functionality.

Phase 1—Initial Design Requirements

Overview

We used the information gathered through previous surveys administered by our research team with 305 members of the Canadian Osteoporosis Patient Network and 34 health care professionals (physicians, nurses, physiotherapists, and pharmacists). We used the results of the surveys to gather data from a large population to determine the content and functionalities needed when developing an mHealth app for acute pain self-management. The question asked in the survey to both groups was as follows: What do you believe to be the most important (1) content to include and (2) the functionalities to have on a mHealth app to empower adults aged >60 years to manage acute pain at home after discharge from the emergency department, following management of a skeletal fracture?

Data Analysis

The results of the surveys were analyzed to identify design requirements, high-level functionalities, and use cases by means of hierarchical task analysis specifically for the purpose of developing our prototype [26]. In addition, we identified a list of accessibility design guidelines for older adults from the literature [27-30]. Examples of accessibility guidelines that we considered in our design included using large font sizes, high color contrasts, large buttons, simple gestures, consistent layouts, and flattened menu structures. In addition to these accessibility guidelines, we opted for a hub-and-spoke navigation pattern in which users have to backtrack to the home page to access another part of the app, as prior studies have demonstrated that this pattern is easy to navigate for older adults [31,32].

We discussed and iterated sketches of the app within our multidisciplinary team, which included a clinician-researcher specialized in internal medicine, 2 user experience designers, a computer science researcher, and a clinical research assistant. Once we finalized the initial design of our app, we converted the sketches to a digital low-fidelity prototype using Axure (Axure Software Solutions Inc).

Phase 2—Participatory Design Workshops

Recruitment

In total, 6 older adults from Canada were recruited to take part in 4 participatory design workshops. Inclusion criteria for the workshop participants were as follows: they must (1) be aged ≥50 years; (2) have sustained at least one skeletal fracture after the age of 40 years; (3) be able to communicate in English; and (4) have access to the internet and own a desktop computer or laptop computer with a camera and microphone.

Procedure

Participatory design is a method that empowers users to become co-designers, inviting them to actively participate in the design process [33,34]. As such, we decided to conduct participatory design workshops so that our targeted end users could directly influence the design. The objective of these design workshops was 2-fold: to obtain feedback on the digital low-fidelity prototype and to uncover unanticipated requirements. We obtained informed consent and demographic information before commencing the participatory design workshops. We facilitated 4 workshops with the same group of participants from March 2021 to July 2021. We conducted the workshops remotely over Zoom (Zoom Video Communications) and audio and video recorded the workshops. Three members of the research team were present during the workshops. One moderated the session; one assisted the moderator in answering questions; and one observed and took notes, occasionally asking confirmatory questions. The participants were shown the prototype through Zoom’s screen-sharing functionality during the workshops. The aim of the first 3 workshops was to seek feedback and suggestions on different parts of the prototype. Upon analyzing the data gathered in the first 3 design workshops, we designed an interactive medium-fidelity prototype, ensuring that all issues raised during the first 3 workshops were addressed. Finally, we
demonstrated the medium-fidelity prototype as a whole in the fourth workshop for a final round of feedback.

**Data Analysis**

Two members of our research team were responsible for the thematic analysis following the steps proposed by Braun and Clarke [35]. The goal of this analysis was to summarize the feedback on our prototype and uncover additional user needs that were missed in phase 1 of the study. In the first step, the analysts read the transcripts from the first 3 design workshops and took notes. In the second step, they reviewed the transcripts and generated initial codes. The codes were then compared for agreement and subsequently applied to the fourth design workshop transcript. Our analysts used Quirkos (Quirkos Software) to facilitate the coding process. In the third step, they generated the initial themes, which were then reviewed in the fourth step. In the fifth step, another member of the team contributed inputs to further refine the themes. Finally, in sixth step, we as a group chose representative quotes for each theme and summarized our findings.

**Phase 3—Usability Testing**

**Recruitment**

We recruited a convenience sample of 10 older adults from the Greater Montreal Area (Quebec, Canada) to participate in usability tests. Inclusion criteria were as follows: the participants must be (1) aged ≥50 years, (2) have sustained at least one skeletal fracture after the age of 40 years, (3) be able to communicate in English, and (4) have access to the internet and own a desktop computer or laptop computer with a camera and microphone. These participants had not participated in the second phase of our study and did not have any prior knowledge of our app.

**Procedure**

To gain information about users’ behaviors while naturally using the product, we decided to conduct observational studies monitoring the use of our app by the users [36]. Following the guidelines in the literature [37], we used an iterative approach for our usability tests by dividing them into batches. Three participants did the initial test. We then addressed the most glaring usability issues and tested the app again with 4 participants. Finally, we tested the app with 3 additional participants. Because of the iterative nature of our prototype, we did not report quantitative data, such as error frequency. We collected informed consent and demographic information before the usability testing sessions. The same person moderated testing sessions remotely via Zoom, and we audio and video recorded the sessions. The moderator sent a secure link over the Zoom chat and instructed the participants to open the link to the medium-fidelity prototype in their browser and share their screen. Studies have demonstrated that using emulators is an acceptable method of mobile usability testing [38]; thus, we used a mobile device emulator on a desktop to test the mHealth app prototype. We chose to use an emulator for ease of use, as most participants were more familiar with using Zoom on their computers than on their mobile phones. The moderator asked the participants to use the different functionalities within the app while thinking back about their latest fracture experience. For example, “Think back to the time you broke one of your bone(s). Let’s say you’ve just taken two tablets of Tylenol for your pain. How would you keep track of your medication intake using the application?” They were also instructed to think aloud using a concurrent think-aloud approach [39]. Throughout the usability test, the moderator took notes on any difficulties the participants experienced. At the end of the usability test, the participants answered open-ended questions about their overall experience with the medium-fidelity prototype. The usability test script is presented in Multimedia Appendix 1.

**Data Analysis**

We transcribed the audio recordings of the usability test sessions to uncover any usability problems that may have been noted during the sessions. We analyzed and classified the errors that the participants made while performing each task and reviewed the answers to the open-ended questions.

**Results**

**Phase 1—Initial Design Requirements**

Surveys were conducted with 305 members of the Canadian Osteoporosis Patient Network (80% aged >60 years; 75% had a previous fracture) and 34 clinicians comprising physicians, nurses, physiotherapists, and pharmacists. We identified 3 categories of functionalities to be included in the app: support resources, diary, and educational materials (Multimedia Appendices 2 and 3 provide the full list of requirements). The support resources guide the users on how to use the app and when and where to seek medical assistance after fractures. The diary functionalities record the pain levels and medication intake of the users. The educational materials provide information on pain management, healing and recovery, mobility, and psychological well-being. We then produced a low-fidelity paper prototype that met all functional and accessibility requirements (Figure 2).

In preparation for the next phase of the study, we converted the paper prototype into a digital low-fidelity prototype. As the prototype was converted, we discussed the designs in the group and simplified or recategorized some functionalities to reduce the number of features for ease of use. This reorganization was also performed in an attempt to further flatten the menu structure of the app. For example, the “mood diary” and “reports” in the paper prototype (Figure 2) were merged with the “pain diary” in the first iteration of the complete digital low-fidelity prototype (Figure 3).
Figure 2. Sketches of the home page and three categories of functionality pages: educational materials, support resources, and diary.
Phase 2—Participatory Design Workshops

Overview

We conducted 4 participatory design workshops with 6 participants (4/6, 67% women; 2/6, 33% men; mean age 76.7, SD 9.5 years) from March 2021 to July 2021; all the participants had experienced a fracture after the age of 40 years. In the resulting codebook (Table 1), four overarching categories emerged from the thematic analysis of the workshops: (1) feedback on the user interface and usability, (2) request for additional app functionalities, (3) feedback on medical guides and educational materials, and (4) suggestions for additional medical educational materials. We do not provide quantitative information such as counting the number of times information (ie, words, terms, and sentences) was mentioned, as the goal of our study was to reveal unexpected findings in an inductive data collection and analysis process, and counting such information would not have theoretical value [40]. We used the feedback from the first 3 workshops to design a medium-fidelity prototype (Figure 4) in which we added and modified the content and features, refined the visual design, and added interactivity. In the fourth workshop, participants reviewed our medium-fidelity prototype, validated the changes, and provided feedback to further refine our medium-fidelity prototype.
Table 1. Codebook resulting from the thematic analysis.

<table>
<thead>
<tr>
<th>Code</th>
<th>Design workshop representative quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Feedback on app user interface and usability</td>
<td></td>
</tr>
<tr>
<td>1.1 Navigation between pages should be intuitive</td>
<td>“The problem when you’re a novice is getting back to where you started. I can’t tell you how many times I’ve had to log out and start all over again.” [Participant 03]</td>
</tr>
<tr>
<td>1.2 Pain scale color scheme should match the level of pain</td>
<td>“[The yellow to red] is a recognized sequence.” [Participant 05]</td>
</tr>
<tr>
<td>1.3 Screens should not be overcrowded</td>
<td>“Because I don’t want to make the screen any busier.” [Participant 02]</td>
</tr>
<tr>
<td>1.4 Users should be provided instructions on how to interact with the app</td>
<td>“You need to be told to click because we just don’t automatically click.” [Participant 03]</td>
</tr>
<tr>
<td>1.5 The app should limit the number of steps the user has to go through to complete a task</td>
<td>“It could be as simple as a button you push when you take your medication.” [Participant 01]</td>
</tr>
<tr>
<td>1.6 Urgent medical information should be highly visible and easily accessible</td>
<td>“If you click on it, it [could] take you over to more information on why a red flag symptom is important to act on immediately.” [Participant 01]</td>
</tr>
<tr>
<td>1.7 The app should have pleasing visual graphics</td>
<td>“Are you going to add any graphics [...] where you put all that whitespace?” [Participant 01]</td>
</tr>
<tr>
<td>2. Requests for additional app functionalities</td>
<td></td>
</tr>
<tr>
<td>2.1 Medication tracker</td>
<td>“I found it very helpful to make a list of when I took the medication because you think you’ll remember, but you don’t.” [Participant 02]</td>
</tr>
<tr>
<td>2.2 Logging pain reminder</td>
<td>“But could it also prompt you with, you know, the time for your next recommended time for your […] next evaluation of pain?” [Participant 01]</td>
</tr>
<tr>
<td>2.3 Warnings for medication overdosage</td>
<td>“Will [the application] tell you to make sure that you don’t take any more narcotics if you have extreme pain? Because that will be a very bad thing for you.” [Participant 04]</td>
</tr>
<tr>
<td>2.4 Data visualization of pain entries</td>
<td>“I’m thinking of the Environment Canada weather app. What we could do is have the pain scale on the vertical axis and the times across the bottom.” [Participant 01]</td>
</tr>
<tr>
<td>2.5 Categorizing contacts in the address book</td>
<td>“Your contact afterwards may not be your family doctor. So [you may need more than] only one health professional contact that you can put in the app.” [Participant 02]</td>
</tr>
<tr>
<td>3. Feedback on app medical guides and educational materials</td>
<td></td>
</tr>
<tr>
<td>3.1 App provides helpful information on red flag symptoms</td>
<td>“I like the idea of red flag symptoms because those are the things that you should address immediately.” [Participant 01]</td>
</tr>
<tr>
<td>3.2 App acts as a reinforcement tool for the information given at the hospital</td>
<td>“So this, to me, is a reinforcement of stuff you’ve already been told.” [Participant 02]</td>
</tr>
<tr>
<td>3.3 App provides helpful information on pain management strategies</td>
<td>“He was given no advice on pain [management], so I think this is excellent.” [Participant 06]</td>
</tr>
<tr>
<td>3.4 App allows for better communication between the user and their health care provider</td>
<td>“It’s difficult for them to share what it has been like for the past 6 weeks or whatever, and the tool will help them communicate that to their doctor.” [Participant 02]</td>
</tr>
<tr>
<td>3.5 The app should use layman’s terms</td>
<td>“Tightness in the throat might be better explained as difficulty in swallowing.” [Participant 05]</td>
</tr>
<tr>
<td>3.6 The app should use inclusive language</td>
<td>“The ‘back on your feet’—it’s just sensitive to people, let’s say they’re in a wheelchair, right?” [Participant 02]</td>
</tr>
<tr>
<td>4. Suggestions for additional medical educational materials</td>
<td></td>
</tr>
<tr>
<td>4.1 The app should provide information about do’s and do not’s after injury</td>
<td>“So, there’s really some do’s and don’t’s that you should know when you leave the hospital.” [Participant 02]</td>
</tr>
<tr>
<td>4.2 The app should guide users when and where to seek medical assistance</td>
<td>“Is there going to be a point where [the application] says, at what stage should you contact your doctor to discuss something else to help better alleviate your pain?” [Participant 02]</td>
</tr>
<tr>
<td>4.3 The app should have information about early mobility</td>
<td>“We’re concerned about how to manage not only the pain but the movement.” [Participant 03]</td>
</tr>
<tr>
<td>4.4 The app should contain links to external websites for tailored information about various injuries</td>
<td>“Is there a website or something that you could direct people to so that they could get […] information […] for immediate aftercare of a spinal fracture?” [Participant 03]</td>
</tr>
<tr>
<td>4.5 The app should provide personalized medical content</td>
<td>“With a vertebral fracture, you’re handed your pain medication, and you were out in the cold. So some kind of very specific suggestions concerning coping with a vertebral fracture [would be nice] to have in the application.” [Participant 03]</td>
</tr>
</tbody>
</table>
Feedback on User Interface and Usability

Participants commented on the appearance and usability of the app. The participants liked the color scheme used for the pain scales, which had a spectrum from yellow (toward less pain) to red (toward more pain) that was intuitive. They also liked having an option to easily return to the home page via a large icon on the header. The participants wanted urgent medical information, such as warnings for life-threatening symptoms, to be easily discernible from the rest of the content to create a sense of urgency. Participants wanted explicit step-by-step instructions on how to interact with the app, such as “Please tap here to continue” and an explanation of the rationale behind each functionality. They pointed out usability issues, such as crowded screens, unintuitive navigation, and poor content discoverability.

We resolved any usability issues pointed out by the participants as our priority because we consider these to be impediments to using the app effectively. These changes were made after each workshop and included (1) decreasing the complexity of the app by removing unnecessary screens; (2) making important information more salient by bolding it with high-contrast colors (eg, red); (3) providing tutorials, in the form of pop-ups, with instructions and explanations on functionality use and benefits; and (4) adding more icons and illustrations to aid navigation and make the app aesthetically pleasing.

Requests for Additional App Functionalities

The participants proposed additional functionalities to be included in the app that would be helpful for pain management. Namely, the participants felt that reminders to log their pain would motivate them to use the app; they wanted to visualize their pain levels on a chart as time progresses, as this would facilitate at-a-glance monitoring; participants indicated having difficulty keeping track of their medication intake owing to high levels of pain and the brain fog caused by pain medication; and they requested functionalities to keep track of their intake and to warn them if they attempted to take medications at a higher frequency than prescribed. Although participants appreciated the contact book of health care providers in the app, they indicated that it would be useful to be able to categorize contacts; this function will help them and their caretakers to quickly identify who to call.

Figure 4. Digital medium-fidelity prototype, showing the home page, pain diary page, symptom checker page, educational materials page, and medication log page.
Consistent with the demands of the participants, we designed and implemented these functionalities after the first 3 workshops. These included (1) reminders to log pain entries; (2) a medication diary, which included warnings if the user attempted to record medication intake too often; and (3) the ability to categorize contacts as “Family Doctor,” “Pharmacist,” “Occupational or Physical Therapist,” “Homecare,” “Clinic,” “Hospital,” and “Other” when saving a contact in the app. We then presented and validated these new functionalities in the fourth workshop, which were all well received.

**Feedback on Medical Guides and Educational Materials**

Participants responded favorably to medical content on acute pain management; they felt they were not given adequate information after hospital discharge and that the mHealth app would provide useful information. Participants also saw the value of the “pain diary,” as it would allow them to monitor their pain levels and communicate their pain journey more effectively with their health care providers. Participants critiqued some of the wordings of content in the app. In response to their comments, we simplified medical jargon into layman’s language. We also changed the wording to offer a more inclusive language.

**Suggestions for Additional Medical Content**

Although the participants perceived the medical content on acute pain management as satisfactory, they indicated that the app lacked recovery and early mobility information. Furthermore, 2 participants stressed the importance of including information about early mobility when recovering from injuries such as hip or vertebral fractures. Participants also wanted practical advice regarding when to seek medical assistance, such as whether they should contact their health care provider if their severe pain does not abate.

Suggestions for additional medical content were reviewed by the clinician-researcher in our team and then added to the app. These included (1) specific practical advice on “dos” and “don’ts” for different types of injuries to supplement general advice, (2) an alert encouraging users to seek medical advice when their pain is uncontrolled (3 consecutive pain scores rated at $\geq 7$), (3) resources leading to external links (such as Osteoporosis Canada) for information outside the scope of this app, and (4) information on early mobility.

**Phase 3—Usability Testing**

**Overview**

We evaluated the usability of our medium-fidelity prototype with 10 participants (7/10, 70% women; 3/10, 30% men; mean age 68.6, SD 4.12 years) in October 2021. All the participants had experienced at least one fracture after the age of 40 years. We analyzed the results of the usability test sessions to design the high-fidelity prototype (Figure 5) and finalized the content, features, and visual design required in the preparation for professional app development.
Overall Impressions of the App
Participants’ overall impressions of the app were positive, with many indicating that they would have used it if it were available at the time of their injury.

Ease of Use
All the participants found the app easy to use, including even those who stated that they usually found new technology difficult to use:

Even for someone like me, I found it easy. But I would have had my son or daughter to help me if I had any questions. [Participant 15]

Personalization
Each person’s experience with fractures was different. Participants who had minimal pain were interested in educating themselves on their injuries and steps toward recovery rather than addressing pain:

I didn’t need medication. I had a little pain. [...] So, for me, general information would have been good

because there were many things I didn’t know. [Participant 09]

Participants who experienced more pain after their fracture were interested in app functionalities directly related to managing acute pain, such as recording the time and dosage of pain medication intake:

I think the very first thing I would be interested in is the medication log. Because seeing how much I need to take day to day, I think, is crucial. [Participant 15]

Communication Facilitator
Participants perceived value in using the app as a communication tool when visiting their health care providers. Participants thought the pain diary could act as a memory aid to help them recall their overall health and issues they face:

I think it’s great because you go to the doctor and you sort of forget to say things […] or you forget to ask some questions. [Participant 14]
Types of Difficulties During Task Performance

Overview
Although the overall participants’ experience with the prototype was positive, the usability tests revealed problems with the user interface and areas for improvement. We addressed the problems encountered by the participants during the usability tests in the final version of our prototype. Multimedia Appendix 4 provides a full list of the identified and resolved difficulties.

Identifying Interactable Elements
A few participants failed to notice a button or an input field or mistook noninteractable elements for buttons, such as textboxes with a border.

Navigation
Lack of navigation affordance caused issues among some participants. The participants failed to realize that they could scroll on pages to access information beyond the screen frame. The hub-and-spoke navigation pattern worked well, as we did not notice participants struggle with returning to the home page or accessing different parts of the app. Most of the navigation problems were related to the nomenclature of the items on the main menu and some buttons, as participants noted that the name of the item or the button was not indicative of the feature or function.

Data Entry Interactions
In some instances, participants either failed to enter the expected information or completely missed inputting information. Incorrect field completion and omission of input may have been because of the high cognitive effort required to carefully read questions or confusion caused by the question’s phrasing.

Lack of Error Recovery
Some participants inadvertently skipped the onboarding process because they were unsure if they were required to continue and pressed the “skip” button.

Feedback Provided During Task Performance
In addition to the identified usability problems, the participants provided feedback on aspects of the app that were confusing or cumbersome.

Interpretation of Information
As the participants navigated to access different functionalities of our app, tutorials appeared as pop-ups on how to use the functionality. A few participants skipped some of the tutorials and noted that they were too long. In addition, some participants indicated that they would have liked the tutorials to explain the reasoning and benefits of using a certain feature. For example, they wanted to know how long they should be using the pain diary and how it would help them manage their pain.

High Loads of Cognitive Effort Required
Some participants indicated that they would not use the Brief Pain Inventory—short form part of the “pain diary” regularly, as they thought it took too long to complete. The Brief Pain Inventory—short form is a 10-item, validated and self-administered questionnaire to assess pain and its impact on daily functioning [41]. In addition, most participants had trouble remembering which medications and dosages they were prescribed when the app asked them to input their prescriptions. Some questions in the app were poorly formulated; for example, participants expressed confusion over a question in the profile setup, which asked, “Were you prescribed or recommended to take acetaminophen for your pain?” As the possible answers were either “Yes” or “No,” participants found it hard to answer the double-barreled question as they thought they were being asked 2 questions in 1.

Discussion

Principal Findings
This study demonstrated how we designed, developed, and evaluated an mHealth app prototype to empower older adults to self-manage their acute pain after a fracture using a 4-step HCD approach: (1) definition of context of use, (2) identification of user requirements, (3) production of design solutions, and (4) evaluation of design solutions to design our mHealth app [42]. Our multidisciplinary team is one of the strengths of our study, as experts from various fields provided input in the design of the app. Inputs from a physician and from pharmacists were considered, as they ensured the accuracy of our medical content. Another strength of our study is the involvement of older adults, which allowed us to gain a deeper understanding of their needs and frustrations when using mHealth technologies to manage their acute pain.

It is widely believed that older adults are often fearful and unwilling to try new technologies [43,44]; our findings challenge this stereotype, as the participatory design workshops and usability tests revealed a need and enthusiasm for an acute pain management mHealth app. Many participants from our design workshops and usability tests expressed their willingness to use mHealth technology to assist in their pain and injury management journey.

In line with other studies on mHealth tools, we also found that people are looking for clear, concise, and personalized health content [45-47]. Methods of managing pain and recovery may differ based on the type of injury, and a one-size-fits-all approach will rarely meet the users’ needs; health content that is relatable to one’s particular case is often perceived as more beneficial [48]. As stated in the literature [49,50], we found that it is important to provide clear instructions on how to use the app for older adults. Older adults often view the use of technology as a series of steps or procedures and are less inclined toward trial-and-error learning styles owing to the fear of “breaking something” [51]. In addition, our participants highlighted the importance of understanding the benefits of using the app; previous studies also echoed the importance of communicating the benefits of using the proposed technology to older adults [52].

Through design workshops, we found that participants who had sustained skeletal fractures considered mobility as important as pain management. We had originally emphasized pain management in the initial phase of the study but quickly pivoted to include more educational materials on mobility in the second phase. The sentiment of mobility being an important topic was
echoed by the participants in the usability tests in the third phase, who were pleased to see a large number of educational materials related to mobility.

We established evidence-based design requirements from needs assessment studies, accessibility design guidelines for older adults from the literature, and participants’ feedback on our design iterations. These efforts contributed to our final mHealth prototype design that all the participants perceived as useful and easy to use. Although the overall results of our usability tests were positive, we uncovered some design problems. Similar to previous studies, scrolling [53,54], identifying buttons to trigger an action [54-56], and interacting with nonactionable targets [32] were the most common problems encountered in our evaluations. Therefore, we recommend that designers be mindful of these potential difficulties when designing mHealth apps for older adults.

Limitations
Our study is not without limitations. First, our studies were conducted remotely owing to the COVID-19 pandemic restrictions, which may have affected the participants’ interactions with the prototype. We used only an emulator on a desktop computer to test our medium-fidelity mHealth prototype; thus, it is possible that some usability problems were not detected. In the future, the prototype should be tested on a smartphone to approximate real use conditions. Second, convenience sampling may have biased our results, as we recruited participants who had participated in previous studies on bone health or who had been evaluated in orthopedic clinics at our center. Design workshop participants were well aware of osteoporosis and its negative impacts and may have had prior knowledge related to injury management. Nevertheless, we believe that they represent the population of patients with skeletal fragility who might use such a tool for pain management following a fracture. In the future, the prototype should be tested with participants with limited or no knowledge of managing an injury. Third, the protocol for the usability tests required participants to have access to Zoom. Consequently, these participants likely had higher technological literacy than those who would never have used videoconferencing tools. In the future, this app should be evaluated by participants with low technological literacy, and the design should be modified accordingly.

Conclusions
Our prototype results from the needs assessment surveys and the insights provided in co-design workshops and usability tests, with content developed in partnership with practicing health care professionals. Our prototype is promising, as the usability test results indicate that the prototype was easy to use for the older adults who participated in this study and contained useful materials. Researchers aiming to develop mHealth technologies would benefit from an HCD approach, as this method promotes the establishment of evidence-based requirements and eliminates potential frustrations early in the design process through continuous evaluation of iterations. In future studies, we plan to professionally develop this prototype on mobile devices and evaluate the impact of the app’s use on patient health outcomes through clinical trials and longitudinal studies.

Acknowledgments
This study was supported by an operating grant from the Canadian Institute of Health Research (201410 PHE) and the Canadian Natural Sciences and Engineering Research Council (RGPIN-2021-04222). The funding sources had no role in the design of the study, analysis of the results, writing of the manuscript, or decision to submit the manuscript for publication. SNM is a scholar of the Fonds de Recherche du Québec en Santé.

The authors would like to thank the patient partners for their time in attending the design workshops as well as the pharmacists who validated the contents of our app. Finally, the authors would like to thank Khiran Arumugam and Kathy Khorramak for assisting the team in testing our prototypes.

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

Authors’ Contributions
KTN and FR had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Usability test interview plan.
[DOCX File, 17 KB - aging_v5j4e37772_app1.docx ]

Multimedia Appendix 2
Initial design requirements identified from prior surveys with members of the Canadian Osteoporosis Patient Network and health care providers.

References


2. Institute of Medicine (US) Committee on Advancing Pain Research, Care, and Education. Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research. Washington, DC, USA: National Academies Press (US); 2011.


Abbreviations

HCD: human-centered design
mHealth: mobile health

©Kevin Tran-Nguyen, Caroline Berger, Roxanne Bennett, Michelle Wall, Suzanne N Morin, Fateme Rajabiyazdi. Originally published in JMIR Aging (https://aging.jmir.org), 17.10.2022. 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.
Lessons and Reflections From an Extended Co-design Process Developing an mHealth App With and for Older Adults: Multiphase, Mixed Methods Study

Catherine Tong¹, DPhil; Alison Kernoghan¹, MSc; Kassandra Lemmon¹, BSc; Paige Fernandes¹, MSc; Jacobi Elliott¹,², DPhil; Veronica Sacco¹, MSc; Sheila Bodemer¹; Paul Stolee¹, DPhil

¹School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada
²Lawson Health Research Institute, London, ON, Canada

Abstract

Background: There are many mobile health (mHealth) apps for older adult patients, but research has found that broadly speaking, mHealth still fails to meet the specific needs of older adult users. Others have highlighted the need to embed users in the mHealth design process in a fulsome and meaningful way. Co-design has been widely used in the development of mHealth apps and involves stakeholders in each phase of the design and development process. The involvement of older adults in the co-design processes is variable. To date, co-design approaches have tended toward embedding the stakeholders in early phases (eg, predesign and generative) but not throughout.

Objective: The aim of this study was to reflect on the processes and lessons learned from engaging in an extended co-design process to develop an mHealth app for older adults, with older users contributing at each phase. This study aimed to design an mHealth tool to assist older adults in coordinating their care with health care professionals and caregivers.

Methods: Our work to conceptualize, develop, and test the mHealth app consisted of 4 phases: phase 1, consulting stakeholders; phase 2, app development and co-designing with older adults; phase 3, field-testing with a smaller sample of older adult volunteer testers; and phase 4, reflecting, internally, on lessons learned from this process. In each phase, we drew on qualitative methods, including in-depth interviews and focus groups, all of which were analyzed in NVivo 11, using team-based thematic analysis.

Results: In phase 1, we identified key features that older adults and primary care providers wanted in an app, and each user group identified different priority features (older adults principally sought support to use the mHealth app, whereas primary care providers prioritized recording illnesses, immunizations, and appointments). Phases 2 and 3 revealed significant mismatches between what the older adult users wanted and what our developers were able and willing to deliver. We were unable to craft the app that our consultations recommended, which the older adult field testers asked for. In phase 4, we reflected on our abilities to embed the voices and perspectives of older adults throughout the project when working with a developer not familiar with or committed to the core principles of co-design. We draw on this challenging experience to highlight several recommendations for those embarking on a co-design process that includes developers and IT vendors, researchers, and older adult users.

Conclusions: Although our final mHealth app did not reflect all the needs and wishes of our older adult testers, our consultation process identified key features and contextual information essential for those developing apps to support older adults in managing their health and health care.

(JMIR Aging 2022;5(4):e39189) doi:10.2196/39189

https://aging.jmir.org/2022/4/e39189

JMIR Aging 2022 | vol. 5 | iss. 4 | e39189 | p.183

(page number not for citation purposes)
KEYWORDS
mobile health; mHealth; older adults; health care providers; co-design; user experience or UX design; qualitative; apps; elderly; health care; care provider

Introduction

Background

There is a growing interest in the role of technology in supporting older adults’ health and well-being. Emerging technologies may help older adults monitor their health by facilitating coordination of care, communication with members of their care team (including health care professionals and family caregivers), and self-management [1-3]. Mobile health (mHealth) technologies can improve patient experience [4], and enhance the delivery of health care by improving communication and collaboration and supporting health care professionals [3]. As Cameron et al [5] argue, “mobility is central to the notion of participatory healthcare,” allowing patients to engage in health care unconstrained (or less constrained) by time and space. While there are many definitions of mHealth [5], which overlap with eHealth, here we draw on the World Health Organizations description and distinction: mHealth is “the use of mobile wireless technologies for public health, or mHealth, is an integral part of eHealth, which refers to the cost-effective and secure use of information and communication technologies in support of health and health-related fields” [6].

The boundaries of mHealth have expanded rapidly with technological advancements and an increasing trend of accessing the internet through mobile and handheld tablet devices [7,8]. Despite the broad willingness of individuals to use mHealth technology to manage their health [9], research and product design in this field have been predominantly directed at one cohort of users: younger people [10,11]. Recent findings are challenging traditional stereotypes that suggest that older adults are afraid of, or unwilling to use, technology; older adults’ use of technologies, such as computers, mobile phones, tablets, and smartphones, has been steadily increasing in the past decade [12-14]. In addition, the COVID-19 pandemic has further accelerated older adults’ adoption of new technologies [15,16]. Older adults increasingly recognize that technology can play a role in supporting self-management practices through health monitoring and access to information [17-20].

There is emerging evidence specific to older patients, that mHealth tools could be adopted to support pain management [21], increase mobility (eg, [22]), mitigate fall risk [23], support healthy habits (eg, [24]), and manage a range of chronic conditions (eg, [25]). Many mHealth apps have been designed to support chronic disease management [26], including in older adult patients [3]. The development of mHealth apps has tended to echo the health care system that deals with specific diseases, conditions, and health care goals separately. What is lacking is a comprehensive mHealth app that would support an integrated approach to managing an older patient’s individual health care goals, needs, appointments, medication reminders, and health care communications.

Although some have questioned the clinical value of mHealth, because of a lack of evidence demonstrating broad impacts on patients [27], the sector has “exploded” [5]. There are many mHealth apps for older adult patients, but Wildenbos et al [28] scoping review found that broadly speaking, mHealth still fails to meet the specific needs of older adult users. According to Wildenbos et al [28], the development of mHealth apps for older adult users must consider cognition, motivation, physical ability, and perception, and be specifically mindful of physical and perceptual barriers. These considerations have also been echoed by Li et al [29] in their reflections on mHealth apps for older users.

Co-design and User Experience With Older Adults

The comments from Wildenbos et al [28] and Li et al [29] highlight the need to embed users in the mHealth design process in a fulsome and meaningful way. The principle of designing with the user, for the user, is reflected in both a “co-design” approach and user experience (UX) approach. Co-design has been widely used in the development of mHealth apps (see [30]) and involves stakeholders in each phase of the design and development process. The involvement of older adults in the co-design processes is variable [31]. The systematic review by Noorbergen [30] found that co-design approaches tend to embed stakeholders in early phases (eg, predesign and generative) but not throughout. Conceptually similar to co-design, UX or UX design emphasizes incorporating the perceptions of users resulting from their own experiences using a service or product through processes such as usability testing [32,33]. Within the UX field, principles of human-centered design can support the development of products that address the needs and capabilities of users [34]. Focusing on the needs and capabilities of users also reflects what Wildenbos et al [28] and Li et al [29] specifically stated about designing for older users. More recently, researchers have offered specific UX design approaches with and for older adults, notably Russell Kirksey [33,35]. Relationships and trust building have been identified as important elements of these approaches [31]. This growing body of literature suggests that without embedding the user throughout, without focusing on the needs and capabilities of users, or without relationship and trust building between users and developers, co-design may not produce the intended outcomes.

The aim of this study was to reflect on the processes and lessons learned from engaging in an extended co-design process to develop an mHealth app for older adults. The aim was to design an mHealth tool to assist older adults in coordinating their care with health care professionals and caregivers, with an emphasis on primary care. Our research question and objective for this manuscript are as follows:

- Research question: What do older adults, primary care providers, and other stakeholders wish to see in an mHealth app that supports older patients in managing their health and health care?
- Manuscript objective: What did we learn from an extended co-design process involving older adults and developers?
How may our reflections and lessons learned inform our future work and those of others?

This project was nested within a broader, multicomponent intervention designed to transform primary care for older patients and their caregivers in 3 Canadian provinces [36]. The process for informing, developing, and testing the app is presented below.

**Methods**

**Overview**

Our work to conceptualize, develop, and test the mHealth app consisted of 4 phases: phase 1: consulting stakeholders (Jan-Nov, 2017); phase 2: app development and co-designing with older adults (Jan-Sept, 2018); phase 3: field-testing with a smaller sample of older adult volunteer testers (Jan-April, 2019); and phase 4, reflecting, internally, on lessons learned from this process (Sept-Dec, 2020). The methods used to track, report, and understand each phase are outlined below. We are a team of mixed methods researchers, gerontologists, and professionals in the geriatric health care sector who came together to implement a multifaceted intervention to improve primary care for older patients. This was the first app development initiative undertaken by our team.

**Phase 1 Methods: Consulting Stakeholders, After a Scoping Review**

Before phase 1, we conducted a scoping review of the types of mHealth tools that exist to support care coordination for older adults living in the community, as well as their existing and desired features and implementation issues (results reported elsewhere; [37]). Findings from the scoping review informed a consultation process with stakeholders, as per the scoping review methodology by Levac et al [36]. Through the consultation step and alignment with the principles of a co-design approach [30], we aimed to better understand the mHealth preferences of key stakeholders, including older adults, family caregivers, and primary care providers.

A total of 26 participants were recruited from both urban and rural locations in Southern Ontario, Canada. Data collection included individual interviews with 5 primary care providers, 1 caregiver, and 1 technology expert, and 4 focus group interviews (4-6 participants each) with older adults and caregivers. Older adults were defined as persons aged ≥65 years who were living in the community accessing primary care services and who were able to speak English and provide their own consent. Family caregivers were persons of any age who took the role of caring for older adults living in the community. Providers included persons of any age who played the role of primary care providers, such as family physicians or nurse practitioners. As this work was embedded in a larger project aimed at transforming primary care for older patients, we leveraged existing partnerships with 5 clinics and relevant organizations (eg, patient advisory groups) to recruit, using both a recruitment script (delivered by gatekeepers in each group) and recruitment posters. Participants were interviewed (either individually or via a focus group) and asked to complete a priority-setting questionnaire, where they ranked mHealth app features, which had been derived from the scoping review, in order of importance. The interview guides are presented in Multimedia Appendix 1, and the questionnaires in Multimedia Appendix 2.

Semistructured interviews and focus groups with participants were used to obtain a richer understanding of mHealth’s preferences. These individual and focus group interviews lasted approximately 45-60 minutes and were digitally recorded and transcribed. Interview data were analyzed by 1 graduate student (PF) using the thematic analysis approach by Braun and Clarke [38], supported by NVivo (Version 11, QSR International). Data from the questionnaire were entered into an Excel spreadsheet, averaged, and ranked based on the mean ratings. SDs were calculated based on sample variance and reported as indicators of consensus. The highest rated features were considered as priorities to be included in an app.

**Phase 2 Methods: App Development and Co-designing With Older Adults**

After completing the background scoping review and consulting with stakeholders to understand what should go into an app to support older patients and their caregivers, particularly in primary care, we began the app development and co-design process. Our research team led co-design sessions with older adults, and we contracted local developers to create the coding required for the app. App development and consultations with older adults occurred simultaneously, with the intent of using older adult feedback at each stage of the development process. Older adults were recruited from the Seniors Helping as Research Partners (SHARP) group, with which our team had worked for more than 7 years, and who were also involved in the interviews described in phase 1. During the app development process, we conducted 3 in-person co-design sessions with approximately 6 older adults per session, led by team members trained in facilitation and qualitative data collection. We communicated the findings of the older adult co-design sessions via team meetings with developers and more than 100 emails. To report and reflect on this process, we have drawn on the development contracts and letters of support, field notes from the co-design sessions, and minutes and recordings from 15 meetings with the developers, whom we have anonymized here.

**Phase 3 Methods: Field-testing With Older Adults**

Participants who had been involved in the earlier stages of the co-design process were contacted by researchers via email to invite their participation in testing the app. A total of 6 participants (5 older adults and 1 caregiver) from a midsized city in Southern Ontario agreed to field-test the app for a 6-month period. Interested participants completed individual, in-person training tutorials (20-30 minutes in length) with a research associate. During one-on-one training sessions, research associates assisted participants with logging into the app using an iPad or Samsung tablet provided by the research team to ensure that testing was completed in both Apple (iPad) and Android (Samsung) formats. The researchers guided participants in a step-by-step tutorial on how to use each of the features available within the app. An Apple- or Android-specific user guide for the app was provided to each participant to provide basic information on how to use the tablet device and app. A

https://aging.jmir.org/2022/4/e39189

JMIR Aging 2022 | vol. 5 | iss. 4 | e39189 | p.185

(page number not for citation purposes)
research assistant connected individually with each field tester at the 2-week mark, and then monthly. Communication with the field testers depended on personal preference and included telephone calls, in-person meetings, and emails with the research team. Field testers were asked to take notes on specific features in the app; these note-taking templates are available in Multimedia Appendix 1. Scheduled conversations with field testers were digitally recorded and guided by the interview questions listed in Multimedia Appendix 1.

**Phase 4 Methods: Reflecting on Lessons Learned**

To explore our team’s reflections and lessons learned from the process, a small focus group session (n=3) was held with members of the team closely involved with the co-design project from its inception. We recorded and transcribed the focus group and applied thematic analysis [38] to identify the lessons learned.

**Ethics Approval**

We received ethics clearance from the University of Waterloo’s Office of Research Ethics (ORE # 44428) for phases 1, 2, and 3; phase 4 only included members of the research team and all coauthors who agreed to share and record their reflections for this manuscript.

**Results**

In the following sections, we highlight the findings for each phase. In phase 1, we identify the mHealth features that older adults and primary care providers value and contextual considerations from interviews and focus groups; in phase 2, we share observations from the app development and co-design process; in phase 3, we briefly describe field-testing; and in phase 4, we describe the lessons learned from this co-design process, drawing on a reflective focus group with our research team.

**Phase 1 Findings: Consulting Stakeholders**

In the questionnaires, older adults and caregivers reported that they were most interested in training and supporting the use of the tool, keeping a contact list with their care team, reminders to bring items to appointments, and the ability to track their illnesses. Meanwhile, primary care providers were interested in older adults bringing their medications, appointment details, contact list of the older patient’s care team, goal setting, tracking exercise, alerts if their health data were out of range, and reminders to track health information. Table 1 displays the top 10 mHealth feature priorities of older adults, caregivers, and primary care providers identified in the questionnaires.

Textbox 1 outlines the 8 major themes identified in our analysis of the interviews with stakeholders. While the surveys identified desirable features and support for an mHealth app, the interviews revealed and expanded on important contextual factors that could support or impede the use of such an app. These contextual factors were connected to the broader (disjointed) health care system where we work, access to technology (eg, limitations in more rural areas), and more individual-level considerations, including individual patient and caregiver differences and preferences, access and comfort with technology, and health literacy.

For example:

*Dr X has something, another doctor has something, your specialist has something, another specialist got something, another bone specialist has got something, the (police) has something, the fire department’s got something from my wife. Everybody got something, but what do you do with it all?* [caregiver 2]

*...and there are some not so good EMRs and you couldn’t interface with anything.* [Health care provider 1]

Participants also discussed the utility and potential of mHealth apps and generally felt very positive about mHealth; however, important considerations may limit uptake or reproduce existing health inequalities (eg, for people with limited resources, low health literacy, very poor health). Participants also emphasized that an mHealth app would likely be more successful if it was offered and aimed at caregivers supporting frail older adults:

*A lot of my patients have great caregivers, daughters, sons who come to appointments with them. They would be more likely to adopt the app like that and keep it up to date and they have their own busy life.* [Health care provider 4]

Consultation processes with key stakeholders confirmed that older adults and primary care providers have a strong interest in mHealth tools and pointed to features that should be integrated into an mHealth tool to support care coordination. This background work laid the foundation for the next phases of the project, including partnering with an app developer to create an mHealth tool and testing older adults and family caregivers.
<table>
<thead>
<tr>
<th>Rank</th>
<th>Health care provider</th>
<th>Value, mean (SD)</th>
<th>Older adult or caregiver</th>
<th>Value, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>My immunization records</td>
<td>5.00 (0.00)</td>
<td>I will be able to call a telephone support line if I need help using the app or setting it up</td>
<td>4.44 (1.34)</td>
</tr>
<tr>
<td>2</td>
<td>My illnesses</td>
<td>5.00 (0.00)</td>
<td>I will be given a face-to-face training session on how to use the app</td>
<td>4.39 (1.42)</td>
</tr>
<tr>
<td>3</td>
<td>Appointment name (eg, Cardiologist appointment, Dr ____)</td>
<td>5.00 (0.00)</td>
<td>I will be given a user manual with written instructions of how to use the app</td>
<td>4.28 (1.41)</td>
</tr>
<tr>
<td>4</td>
<td>I will be able to keep a contact list and information of all those involved in my care team (eg, physician, Nurse, Specialists, etc.)</td>
<td>5.00 (0.00)</td>
<td>Appointment name (eg, Cardiologist appointment, Dr ________)</td>
<td>4.22 (1.52)</td>
</tr>
<tr>
<td>5</td>
<td>My medication</td>
<td>4.8 (0.45)</td>
<td>There will be a tutorial within the app to explain to me how to set-up and use it</td>
<td>4.22 (1.44)</td>
</tr>
<tr>
<td>6</td>
<td>I will have the ability to give access to others (health care providers or caregivers)</td>
<td>4.8 (0.45)</td>
<td>Prepare for appointments—bring medications</td>
<td>4.16 (1.50)</td>
</tr>
<tr>
<td>7</td>
<td>Prepare for appointments—bring medications</td>
<td>4.8 (0.45)</td>
<td>Having the option of a paper-based or hard-copy version rather than web-based version</td>
<td>4.12 (1.27)</td>
</tr>
<tr>
<td>8</td>
<td>Appointment details—date and time</td>
<td>4.8 (0.45)</td>
<td>Prepare for appointments—bring health documentation</td>
<td>4.11 (1.56)</td>
</tr>
<tr>
<td>9</td>
<td>I will have the ability to track my symptoms</td>
<td>4.6 (0.55)</td>
<td>I will be able to keep a contact list and information of all those involved in my care team (eg, physician, Nurse, Specialists, etc.)</td>
<td>4.11 (1.66)</td>
</tr>
<tr>
<td>10</td>
<td>I will be able to create personal health goals</td>
<td>4.6 (0.55)</td>
<td>Appointment details—location</td>
<td>4.06 (1.70)</td>
</tr>
</tbody>
</table>

*These priorities were identified based on the mean averages from each question in the questionnaire (5-point scale, with 5 being very interested).
Textbox 1. Summary of themes and subthemes from qualitative consultation with key stakeholders.

### System level gaps impact care coordination and self-management
- Problems with information transfer between health care professionals or settings
- Lack of standardization in care coordination practices
- Lag periods between appointments
- Short appointment times with health care professionals
- Challenges with navigating the system

### Microlevel issues impact or prevent self-management
- No standard tracking method
- Patients’ needs vary from simple to complicated conditions
- No equipment at home to monitor own health
- Lack of understanding of health conditions
- Provider does not provide all information to the patient or caregiver
- Caregiver feels burdened managing information

### Older adults currently self-manage their health in various ways
- Tools patients use to keep track of their health information: spouse or caregiver; memory; diary or notebook; pill boxes; paper copies of documents

### Positive experiences empower older adult patients to self-manage health
- Importance of self-advocacy to get information
- Understanding health status
- Building trust or relationships with patients and health care

### Technology can support self-management practice in various ways
- Monitoring via devices
- Phone reminders
- Memo or notepad on phone
- Web-based laboratory results
- Phone calendar

### Apps or technology can support current practices for older adults and caregivers
- Participants’ vision of using apps
- Suggested app features
- Suggested design esthetics

### Technology can be a barrier to adopting or accepting self-management practices
- Limited access to technology
- Privacy concerns
- Financial barriers
- Negative attitude toward technology
- Age as a barrier
- Cognitive impairment
- Technology illiteracy
- Cultural differences
- Transition from paper to technology
- Negative attitudes to tracking health
Considerations for implementing technologies for patients and health care professionals

- Training education for the following: providers on the technology and how it is used; helping end users use the technology
- Developing an implementation strategy for patients and health care professionals
- Need for discussions to be had on what information patients need to track

Phase 2 Findings: App Development and Co-designing

With Older Adults

Our team is situated in the largest technology hub in Canada, and we relied on guidance from colleagues familiar with the local developer context to select an appropriate vendor when developing our research proposal for funding. We engaged in conversations with vendors in 2017 and worked with university staff (eg, our university privacy officer) before selecting a developer to understand issues related to privacy, given that the app would store personal health information for users. Given the nature of our research funding, we had to have vendors selected, with “letters of support” at the point of applying for funding, long before any funds were available, or details were finalized. We selected our vendor, given their prior experience in developing mHealth apps, who began work in early 2018.

Drawing on our consultation work, the app was intended to help store and organize older adults’ health care information, including the professionals and clinics involved in their care, appointments, and medication lists. The app also reminded users of upcoming appointments and tasks, in addition to allowing older adults to track and monitor health information, such as weight, blood pressure, or physical activity. Older adults from SHARP group reviewed paper mockups of the health app and provided preliminary feedback. SHARP members predominantly highlighted issues around accessibility, including small font size, use of colors, minimizing language complexity, and simplifying navigation. App developers implemented this initial feedback to create a web-based app prototype. Older adults from the SHARP group were asked to review the prototype and provide comments on the platform. Researchers and app developers incorporated some of this feedback into the app design before its launch on the App Store and Google Play Store, which are available in English and French.

The next step involved testing the app with group members of SHARP. Older adults in this group tested the app for approximately 2 weeks, on their own devices or a loaned device from our team, before providing additional comments to the research team. While users appreciated the general appearance of the app and app icon, some older adults found it cumbersome to navigate through the app and the accompanying manual. Older adults suggested providing additional training on how to use all the features on the app, along with providing styluses. Researchers and developers have attempted to incorporate this feedback into an update for the app before starting field-testing; however, many suggested could not be implemented, as our vendor perceived these additional changes to be outside the scope of our agreements. For example, older adults wanted the option to view their calendar of appointments both weekly and monthly, but developers deemed this “out of scope.” Testers would also have appreciated the option to consolidate the calendar in the app with the existing calendar on their devices; likewise, this was deemed out of scope.

Expectations and Challenges in Working With the Vendor

We contracted the vendor for 11 months. In the first 2 months, we regularly met with the vendor and concurrently engaged in focus groups with older adults (outlined above) to review the app development process. While we, as health researchers, called this “co-design” or “co-creation” [39,40], many in the development sector would have labeled this UX.

By month 4, we realized a mismatch between our expectations and what the vendor was willing or able to deliver. We provided feedback from the co-design process with older adults in months 2 and 3, but very little user feedback was implemented. Although our initial agreement noted that the vendor would “rely heavily on the user’s input,” and the contract stated the developer planned to “engage with the project team in an iterative fashion,” we viewed this engagement as quite limited. While some requests were justifiably out of scope (eg, linking the app to the existing electronic medical record [EMR] systems of the users’ physicians), many requests appeared (in our minds) relatively straightforward and were covered by our initial agreements. There were also occasions in what the older adults wanted (eg, specific ways to navigate the app or visual preferences) were not aligned with what the developers deemed “best practices” in their field (eg, which some users requested text to be in all caps for readability and vision issues, this is generally avoided in app development). This example, however, raises the question: When designing an app for older adults, should their preferences or “best practices” take precedence?

The final product was not reflective of user input and feedback.

Phase 3 Findings: Field-testing With Older Adults

After launching the app on Apple and Google Stores, we distributed the devices with the app (and styluses to support usability) to 6 volunteer field testers with the intention of eliciting their feedback and tracking their use of the app for 6 months (eg, to determine which features were being used most often, to understand if the app was taken to medical appointments or used in conjunction with caregivers, etc.). Unfortunately, 2 participants returned their devices before study commencement: one because they found the app too complex and the other because of visibility issues (they reported that they would have preferred all text in the app to be capitalized for better readability or a setting that allowed the user to change the font to capitalized, depending on one’s preference). The initial results showed that the app supported the management of some aspects of participants’ health or health care. For example, one participant found the Reminder function in the “My Calendar” element to be especially helpful in
organizing their health care (eg, appointment and medication reminders) and another found the Care Team feature to be particularly beneficial for consolidating the contact information of their health care professionals. However, the bulk of the feedback reflected that most inputs from older adults in phase 2 (development and co-design) were not addressed. There were consistent usability and accessibility issues (eg, small font size, readability, and overall complexity of the app) and features that were available but not functional in a way that the participants envisioned (eg, there was no mechanism to link the calendar in the app to the existing calendars on the devices, so participants had to maintain 2 calendars, one for their health in the app, and one on their device for regular scheduling of personal and professional events). Field-testing was halted after 3 months, given consistently poor feedback from the remaining testers. Although we launched this publicly, the app was not rolled out as part of our wider intervention in primary care, given the usability concerns raised above.

**Phase 4 Findings: Reflecting on Lessons Learned**

**Overview**

Suspending field-testing and not rolling out the app for use on a wider scale was not the intended outcome, and our team concluded this work with a reflective focus group to try and better understand the lessons learned. The lessons are summarized in Table 2 and in the narrative below.

<table>
<thead>
<tr>
<th>Lessons learned</th>
<th>Description</th>
<th>Supporting quotations, from the research team</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Selecting a strategic partnership with aligned goals</td>
<td>Ensuring that partnerships between research teams and tech developers are grounded in an understanding of each other’s goals and priorities and that the project will be mutually beneficial for both groups.</td>
<td>“We need to do our due diligence to go out and interview different teams...and select one that shares the same values and wants to work the same way that we have in mind.” [P 1]</td>
</tr>
<tr>
<td>2. Including a person on the research team with content expertise in tech development</td>
<td>Researchers commented on not speaking the same “language” as tech developers, and this posed some challenges around what was or was not possible as the project evolved. Incorporating a consultant with relevant experience in the technology and app development space was proposed as a suggestion.</td>
<td>“Have someone even in the know, review what the other company is saying...So maybe have someone look over their letter of support or what their proposal would be.” [P 3]</td>
</tr>
<tr>
<td>3. Facilitating direct relationships between users (ie, older adults) and tech developers from the beginning of the project</td>
<td>The research team can play a significant role in coordinating relationship-building between older adult users and app developers, commencing at the start of the project. As research teams may already have preexisting connections with older adults in the community, they can leverage these relationships to bridge the gap between users and tech developers to co-develop impactful products.</td>
<td>“And then during that meeting at the end of the day, functionally, what does an older adult want, and they can talk about that a lot. And then the tech people can sit there and sort of analyze what’s possible or not...if we brought these groups together from the beginning, and had more planning ahead of time, we might have started a little bit differently.” [P 2]</td>
</tr>
</tbody>
</table>

**Selecting a Strategic Partnership With Aligned Goals**

One of the most significant findings resulting from this project is ensuring that partnerships between research teams and technology developers are grounded in an understanding of each other’s goals and priorities and that the project will be mutually beneficial for both groups. Our research team felt this was not the case for our project, evidenced by the following member’s statement: “We’re going into a business agreement with another organization who doesn’t have the same end goals as we do with this product” [P 2].

Selecting these strategic partnerships can be a challenge for research teams to write grant proposals, requiring them to submit partnership letters and draft a budget under time constraints. However, building in time to meet with different vendors and determining which company’s values and approaches best align with those of the research team is an imperative step in supporting a productive partnership. Another difficulty we encountered was the different understandings of what is involved in a partnership. On the basis of our team’s personal experiences, the developer we worked with was rooted in a “business model” where strictly following the contract was prioritized. However, our understanding of a partnership involves more iterative and flexible processes, such as opportunities for ongoing feedback and making necessary adjustments to the app. When developing contracts with a vendor, researchers should ensure that their team wishes to engage in iterative processes, such as multiple rounds of feedback, which must be explicitly built into the contract from the outset.

**Including a Person on the Research Team With Content Expertise in Tech Development**

Researchers commented on not speaking the same “language” as tech developers, which created some tension around what was or was not possible as the project continued. Because of the iterative nature of the co-design processes, participants involved in testing the app provided feedback at multiple stages. Throughout this process, certain suggestions could not be actioned because the app developers felt that they extended beyond the project’s scope or were not aligned with “best practices” in the developer field. The research team agreed that some recommendations were beyond the project’s scope, such as “older adults really wanting a system that integrates in with practices” in the developer field. The research team agreed that some recommendations were beyond the project’s scope, such as “older adults really wanting a system that integrates in with their primary care EMRs...so that they could have conversations with physicians through the app” [P 1]. It would have also been helpful, particularly at the outset, to have a team member who understood both the length of time it takes to co-design an app and the longer-term commitments required to maintain the app. Finally, it is incredibly challenging to integrate an mHealth app into a health care system that itself is not integrated (eg, if a primary care provider and specialist do not use the same EMR.

Table 2. Lessons learned in the development of an mHealth app with and for older adults.
system, how would an app for older adults integrate with these disparate systems? However, our team felt that other concerns raised by older adults, especially pertaining to the app’s visual design and operability of features, were more feasible for technology developers to address. One member of the research team suggested incorporating a collaborator with relevant experience in the technology and app development space, which could help the team navigate any areas of contention as the project evolves.

**Facilitating Direct Relationships Between Users (ie, Older Adults) and Tech Developers From the Beginning of the Project**

The research team can play a significant role in coordinating relationship-building between older adult users and app developers, commencing at the start of the project. As research teams may already have pre-existing connections with older adults in the community, they can leverage these relationships to bridge the gap between users and technology developers to co-develop impactful products. In the context of our project, researchers took on the role of being the “middlemen” between older adult users and tech developers. We elicited feedback from older adults and relayed information back to tech developers; however, these 2 groups were never brought together, which posed significant limitations. This approach was inefficient because researchers were communicating separately with both groups and tech developers did not receive input directly from users to better appreciate each other’s perspectives. One of our team members highlighted this issue by saying, “If we brought these groups together from the beginning, and had more planning ahead of time, we might have started a little bit differently” [P 2].

**Discussion**

**Principal Findings**

Through our multiphase, mixed methods co-design project, we learned that (1) older adults and primary care providers are keen on an integrated app that helps older adult users manage their health and health care, (2) older adults prioritized multiple modes of support to use the app effectively, whereas primary care providers emphasized the importance of features that helped track (eg, track appointments, the care team, vaccines, etc), (3) co-designing with older adults and developers was fruitful, in terms of learning, but ultimately challenging. We also included a post hoc analysis of the oft-cited “pitfalls” [30] of mHealth development, with the intent to be reflective and inform the future work of our team and others. Both steps offer novel insights into mHealth development for older users. Implications for app development and practice, and implications for future UX and co-design research, are further detailed below.

**Implications for mHealth App Development and Practice**

Our consultation phase identified the features, functions, and important considerations (eg, privacy, accessibility, and affordability) that older adults, caregivers, and primary care providers wish to see in mHealth apps designed to support the health and health care of older patients. Our findings specific to mHealth features and considerations are consistent with other research (eg, [3]) and highlight the importance of perceived usefulness and perceived ease of use [41]. Notably, primary care providers ranked immunization record forms and a list of illnesses as their top 2 features, whereas older adults identified IT user support and training to use the app as their 2 most highly ranked features. We sought to address these priorities and create an app that reflected the preferences of both primary care providers and intended users (ie, older adults). In our case, these preferences were not contradictory: the providers wanted specific features (eg, something to track appointments and vaccines), whereas the older adults wanted support tools to accompany the app (eg, manuals or IT support) so that they could use the features.

When working with the developers, we also saw differences between what the “experts” or professionals were recommending and what older users asked for (eg, asking for design elements that contradict best practices and design standards). In our experience, both users and developers have strong reasons for their preferences, and (in hindsight) we should have built more time into the initial contract for developers to work directly with older adults to come to a shared understanding of conflicting elements.

Some of the features that both primary care providers and older adults wanted were not possible in our health care system. It is fundamentally impossible to integrate an mHealth app into a broader, multifaceted health care system that is not itself integrated [42-44]. In a context in which primary care clinics, hospitals, specialists, and home and community care are potentially all using different EMR systems, or perhaps not using an EMR at all [45], we were not able to respond to the requests for features that connected patients to their records, health histories, appointments, etc, across numerous, disparate systems. While we were aware of the disconnected nature of our system, perhaps we did not fully appreciate the argument by Kircskey [33] that “to find any measure of success, a fully functional app for older users should be integrated into the entire health-care system.” Our participants echoed Kircskey in our consultation phase: the execution of an integrated app is limited in contexts in which the health care system itself has yet to integrate. For those working in a disconnected health care system, and with co-designing technologies, it may be helpful to provide the co-design participants with some context regarding what is (and what is likely not) possible in their particular setting and context.

**Comparison With Prior Work and Future Directions:**

**UX Design**

Unlike most UX approaches in the review by Noorbergen [30], we focused on including users in each step of the co-design process. Consistent with the UX literature [32,33,35,44], we aimed to address the perceptions, needs, and capabilities of our users; however, our IT vendor was not always able to address their needs and requests in the development process. Although UX researchers [31] have emphasized the importance of trust and relationships, in our experience, this was not lacking with the field testers but with the developers. Devoting time to relationships and trust building, including time for socializing,
is recommended as a facilitator of the co-design process [31], but this is complicated when working with developers who must charge for the time they devote to the project.

As more health researchers attempt to leverage technology’s potential to improve the well-being of older adults [1], it is possible (in our experience) that researchers will confront unforeseen challenges in the tech development sector [30]. One aspect overlooked by our team is the importance of ensuring that researchers and app developers are committed to adopting a UX lens when working with user stakeholders to co-create mHealth technologies for older adults [33,35]. Our study’s design broadly aligns with the aspects of human-centered design outlined by Harte et al [34], such as engaging in iterative processes and involving users throughout the different phases of the app’s design and development. Our work further adds to this conversation by showcasing challenges surrounding partnerships between researchers, app developers, and stakeholder users that research teams should consider [46] when co-designing an mHealth tool for older adults, even when following the appropriate guiding principles and standards of UX.

UX design has notably developed and evolved in the last few years, including the application of UX to the development of mHealth apps [33,47-49], and it is likely that teams engaged in this work will find it easier to find developers versed in the UX principles. There is also a broad body of literature on this topic. Much has been written about the barriers to mHealth usage by older adults (eg. [21,28,29]), and our findings from the consultation phase echo the many recommendations and considerations specific to visibility (eg, clear text, contrast), accessibility (eg, ability to zoom in, translate, change font size), and the importance of ease of use (eg, simple navigation, explicitly noting links). We have also seen the emergence of literature on how to design mHealth apps with and for older adults (eg. [29,33,35]), as well as insights and guidance on designing for older persons living with dementia [50], and individuals with a range of physical limitations [51]. This newer body of literature reflects our approach, which embeds older adults in every developmental phase. The limitations of our final product were not necessarily because of a faulty approach, per se, but rather a development partnership that did not prioritize or reflect what we were hearing in our engagement with older adults. UXs and feedback must be meaningfully adopted at each stage of the development process [29,30,33]; this can help ensure that all perspectives are considered and to avoid unrealistic expectations [31]. We would recommend that researchers, clinicians, and developers entering into the mHealth app development process take some time to ensure everyone is on the same page, not just technically but also in their approaches to UX or co-design, long before any contracts are developed.

Agism Against Older Adults and Gerotechnology

The COVID-19 pandemic illuminated and reproduced agist (against older adults) attitudes that have permeated our society for decades; if not longer [52,53]. We live in an ageist society that prioritizes the experiences, preferences, wants, and needs of younger people [54]. Although we would never suggest that the developers that we worked with were explicitly agist, they are (just as we are) part of a society that prioritizes the needs of younger people and are also part of an industry that has historically done a better job designing for younger people [10,11]. We are encouraged to see guidelines and recommendations for designing apps for older users (eg. [29]) and the rising prominence of gerontechnology; a gerontological discipline dedicated to the design and adoption of new technologies for older people [55]. Although researchers have been engaged in gerontechnology for more than 20 years [56], our work suggests that more work is required. Our findings demonstrate a clear desire for an app (or apps) that supports patients in managing information about their appointments, care team, chronic conditions, prescriptions, vaccinations, etc. We are sharing these results in a time when the population is aging, older adults are living longer (but often with a higher number of chronic conditions and prescriptions to manage) [57], when (on account of the COVID-19 pandemic) keeping track of vaccinations has become even more important and older users have increasingly integrated technology and smartphones in their daily lives [15,16]. In an ageist and aging society, there is both a moral imperative and strong business case to be made for designing for the older user.

Strengths and Limitations

The review by Stowell et al [58] has shown that UX design has often overlooked the experiences and input of racialized users, and our sample of older adults was also predominantly White and of relatively high socioeconomic status (a description of the SHARP group can be found in [59]). Given the nature of our findings, this study also does not include the perspectives of the developers we worked with; it is likely that their interpretation of our process and the product delivered would be different. Teams engaging in a UX development process with an external partner may wish to proactively build a “postlaunch debrief” session into their initial contracts, knowing that these processes can be complex and merit a post hoc analysis of what did and did not work for both parties. Finally, although our final mHealth product has real limitations, we would also argue that peer-reviewed literature tends to focus on success stories, and there is great value in reporting on missteps and deviations to inform the work of others [60]. We situate some of our findings (namely phases 2 and 4) within broader calls for publishing “negative” results (eg. [61]), not just research success stories. We believe in the power and importance of designing with and for older adult users and have continued to do so across several high- and low-tech projects.

Conclusions

Although our final mHealth app did not reflect all the needs and wishes of our older adult testers, our consultation process identified key features and contextual information essential for those developing apps to support older adults in managing their health and health care. Furthermore, our reflective process identified important factors to consider when health researchers and gerontologists enter the app development sector. In the words of Karl Popper [62], “every refutation should be regarded as a great success,” and we hope that the reflections, and refutations, shared here will inform and support the future work.
of others seeking to support the health of older adults using mHealth apps.

Acknowledgments
The authors extend their thanks to the primary care providers, participating clinics, family caregivers, and older adults, for taking the time to share their experiences and perspectives. This project was funded by the Canadian Frailty Network (CFN; grant # TG2015-24), which is funded by the Government of Canada’s Networks of Centres of Excellence program. Ms Fernandes’ work was supported by the Canadian Frailty Network and University of Waterloo’s School of Public Health Sciences. Ms Lemmon’s work was supported by University of Waterloo’s Hallman Fellowship.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Interview guide and field-testing notes template.
[DOCX File, 26 KB - aging_v5i4e39189_app1.docx]

Multimedia Appendix 2
Questionnaires.
[DOCX File, 48 KB - aging_v5i4e39189_app2.docx]

References


Hsieh KL, Fanning JT, Rogers WA. Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: 10.2196/jmir.13817] [Medline: 31970289]

Vo V, Auroy L, Sarradon-Eck A. Patients’ perceptions of mHealth apps: meta-ethnicographic review of qualitative studies. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e13817 [FREE Full text] [doi: 10.2196/jmir.13817] [Medline: 31970289]


Fernandes P. Understanding how mHealth Tools Can Support Care Coordination between Older Adults, and Healthcare Providers: A Scoping Study. IWSSpace 2018 [FREE Full text]


Jessen S, Mirkovic J, Ruland CM. Creating gameful design in mHealth: a participatory co-design approach. JMIR Mhealth Uhealth 2018 Dec 14;6(12):e11579 [FREE Full text] [doi: 10.2196/11579] [Medline: 30552080]


Designing Mindfulness Conversational Agents for People With Early-Stage Dementia and Their Caregivers: Thematic Analysis of Expert and User Perspectives

Cassandra E L Seah¹, BA; Zheyuan Zhang¹, BA; Sijin Sun¹, BSc; Esther Wiskerke², MSc; Sarah Daniels³, MSc; Talya Porat¹, PhD; Rafael A Calvo¹, PhD

¹Dyson School of Design Engineering, Imperial College London, London, United Kingdom
²Ladywell Dementia Team, Lewisham Council, London, United Kingdom
³Care Research & Technology Centre, UK Dementia Research Institute, London, United Kingdom

Corresponding Author:
Cassandra E L Seah, BA
Dyson School of Design Engineering
Imperial College London
Imperial College Rd, South Kensington
London, SW7 9EG
United Kingdom
Phone: 44 7742540443
Email: c.seah@imperial.ac.uk

Abstract

Background: The number of people with dementia is expected to grow worldwide. Among the ways to support both persons with early-stage dementia and their caregivers (dyads), researchers are studying mindfulness interventions. However, few studies have explored technology-enhanced mindfulness interventions for dyads and the needs of persons with dementia and their caregivers.

Objective: The main aim of this study was to elicit essential needs from people with dementia, their caregivers, dementia experts, and mindfulness experts to identify themes that can be used in the design of mindfulness conversational agents for dyads.

Methods: Semistructured interviews were conducted with 5 dementia experts, 5 mindfulness experts, 5 people with early-stage dementia, and 5 dementia caregivers. Interviews were transcribed and coded on NVivo (QSR International) before themes were identified through a bottom-up inductive approach.

Results: The results revealed that dyadic mindfulness is preferred and that implementation formats such as conversational agents have potential. A total of 5 common themes were also identified from expert and user feedback, which should be used to design mindfulness conversational agents for persons with dementia and their caregivers. The 5 themes included enhancing accessibility, cultivating positivity, providing simplified tangible and thought-based activities, encouraging a mindful mindset shift, and enhancing relationships.

Conclusions: In essence, this research concluded with 5 themes that mindfulness conversational agents could be designed based on to meet the needs of persons with dementia and their caregivers.

(JMIR Aging 2022;5(4):e40360) doi:10.2196/40360

KEYWORDS

mindfulness; dyadic; dementia; caregivers; user needs; intervention; mindfulness; user; feedback; design; accessibility; relationships; mindset; essential

Introduction

Background

Dementia has been predicted to affect an estimated 78 million people by 2030 [1]; therefore, new interventions are needed to support caregivers and persons with dementia (dyads). Mindfulness-based interventions (MBIs) are an example receiving increasing attention. MBIs for participants with dementia have been shown to improve quality of life [2,3] and decrease depressive symptoms [2]. MBIs have also played a...
role in improving self-reported stress, depression, anxiety, burden, quality of life, cognition, and mood in caregivers of persons with dementia [4-14].

Few studies of MBIs use a dyadic approach in which both the person with dementia and their caregiver engage together in a mindfulness activity [15-19]. Dyadic MBI approaches have shown benefits, such as improved well-being, quality of life, depressive symptoms, relaxation, awareness, acceptance, and resilience [15,16,18]. Although studies reveal positive benefits for dyadic MBIs, there is a lack of focus on how the interventions are designed for the stakeholders. Studies often adapted mindfulness interventions without stating the specific adaptations or the reasons why they were made [3].

Most dyadic MBI studies were performed in physical group settings [5,15-17], with one study conducting home sessions using recordings [20] and another using guided telephone calls [4]. Introducing more advanced technologies such as conversational agents to dyadic mindfulness for persons with dementia and their caregivers is a novel approach and not previously studied. Conversational agents refer to artificial intelligence or computer programs that use natural language processing to converse with people [21]. The use of conversational agents could improve accessibility, as they allow users to access them in the comfort of their own homes instead of traveling to a physical location. They also allow users to access them at any time, as opposed to having to wait for a guide through physical groups or telephone calls. Conversational agents may also support individuals with limited digital competencies, which may be beneficial not only for older adults but also for people with cognitive challenges [22], offering lower barriers to use by using voice as a communication medium, as opposed to using a graphical interface device. Users simply need to speak to conversational agents instead of learning how to navigate through digital interfaces, thereby enhancing accessibility. Conversational agents may also provide more personalization and guidance as compared with recordings, by guiding dyads step by step in a personalized manner, as opposed to a static video or voice recording. They may also provide more scalability as compared with physical groups or guided telephone calls, as they do not require a mindfulness expert to be present each time, allowing for automation as opposed to a manual approach. Conversational agents could benefit current mindfulness interventions for dyads by improving accessibility, scalability, guidance, and personalization, thus offering opportunities for mindfulness to be seamlessly integrated into the dyads’ lives. To understand this, research on conversational agents for persons with dementia, which was relatively scant [6], was explored, with some interesting use cases such as dementia detection [7]. Unfortunately, conversational agents for dementia are inadequately discussed in the scientific community, with studies lacking rigor [6]. However, a recent study revealed that persons with dementia can use embodied conversational agents independently in their home environment [8]. This showed good engagement with the system, revealing the potential for conversational agents to meet the basic and higher-level needs of people living with dementia [8]. To the best of our knowledge, this is the first study of dyadic MBIs supported by technologies such as conversational agents.

Dyadic mindfulness conversational agents may be a more useful way to engage persons with dementia and their caregivers. Although dyadic mindfulness has shown benefits for persons with dementia and their caregivers, physical group settings and guided telephone calls were used, which were not scalable, as they required a trainer for each session, and were not accessible, as users could not use them whenever they wanted. Physical group settings were also not accessible as they required users to travel to a physical location. Recorded formats provide for accessibility and scalability needs, given that users know how to navigate the digital space, but are not personalized and guided such as physical group sessions or telephone calls. A way to bridge this gap is to use conversational agents for dyadic mindfulness, providing accessibility and scalability while also ensuring guidance and personalization for users. However, for future studies to design and test the effectiveness of dyadic mindfulness conversational agents, the needs and preferences of persons with dementia and their caregivers need to be understood first.

To understand the design recommendations, we first examined the existing literature. The recommendations below only address designing dyadic MBIs or conversational agents for persons with dementia. Few recommendations could be found, as studies primarily focused on the effectiveness of dyadic MBIs or conversational agents rather than on the design of the interventions. Recommendations for designing dyadic mindfulness programs for persons with dementia and their caregivers included the following: first, allowing separate interventions for each individual on different occasions [3]. Care should be taken when providing interventions with a dyadic approach because caregivers may be reluctant to fully focus on themselves, as they are concerned about their partners. They may also not feel comfortable discussing their concerns. Second, mindfulness guides must be able to manage participants’ negative emotions [20]. Mindfulness interventions may provide a mental space for negative emotions to arise, and trainers need to be able to guide participants through the process. Third, the needs of caregivers should be met [3]. It was important to ensure that the needs of caregivers were taken care of, as they may be more worried about the well-being of the persons with dementia than about their own well-being. On the other hand, recommendations for designing conversational agents for persons with dementia included improving the quality of speech recognition [8]. Automatic speech recognition quality and synthesis were technical problems that required improvement because they negatively impacted the adaptability and usability of conversational agents. Recommendations from the existing literature only addressed parts of the proposed intervention—dyadic mindfulness or conversational agents for persons with dementia included improving the quality of speech recognition [8].
Research Question

What is important when designing mindfulness conversational agents for persons with dementia and their caregivers?

Methods

Ethical Considerations

Ethics approval was granted for this study by Imperial College Research Ethics Committee (21IC6573). Informed consent was obtained from all the participants before they participated in the study.

Study Design

Semistructured interviews were conducted via Microsoft Teams. The first author interviewed 5 dementia experts, 5 mindfulness experts, 5 persons with early-stage dementia, and 5 dementia caregivers individually based on a predefined topic guide and through a 1-hour interview. Dementia and mindfulness experts were recruited between March and June 2021. Persons with dementia and their caregivers were recruited between October and November 2021. Multimedia Appendix 1 outlines the structure that was adopted for the interviews with experts and users.

Participants

Individuals with expertise in mindfulness training or dementia were recruited from research networks of coinvestigators using an opportunity sampling method. Persons with dementia and caregivers were recruited through dementia groups on the social media platform Facebook (Meta Platforms Inc), where a volunteer sampling approach was used. Persons with early-stage dementia and caregivers individually responded to advertisements on various dementia groups. As it was particularly challenging to recruit individuals who were either caregivers to persons with dementia or persons with early-stage dementia, the caregivers and persons with dementia who participated in this study were given GBP £20 (US $23.60) and GBP £40 (US $47.10), respectively, as an appreciation for their time. Interviewees were informed that they could withdraw at any time without giving a reason. All the participants had to be aged ≥18 years, able to use Microsoft Teams, and communicate in English. In addition, the following criteria had to be met:

- Dementia experts had to be service providers in health care facilities working with and planning programs for persons with dementia.
- Mindfulness experts had to be teachers and practitioners of mindfulness methodologies or therapies, such as Mindfulness-Based Cognitive Therapy or Mindfulness-Based Stress Reduction or mindfulness activities such as meditation without a clinical focus or equivalent.
- Persons with dementia were in the early stage of the illness and able to consent for themselves.
- Caregivers had to be primary caregivers of an individual with a dementia diagnosis (early stage of any type) and must have provided care for at least 3 months before recruitment.

Data Collection and Analysis

The sessions were recorded using Microsoft Teams before they were transcribed. The transcripts were analyzed using NVivo (QSR International; version 12). A thematic analysis approach was used to identify, analyze, and report patterns within the data [9]. Thematic analysis is a flexible approach that organizes and describes data sets that are rich in detail, allowing for the interpretation of various aspects of the research topic [9]. This approach was used because it enabled the identification of important themes across expert and user feedback.

For expert feedback, an inductive bottom-up approach was used by 2 researchers independently. Data were coded as the researchers read through the data. Codes were created and applied, identifying emerging topics as the data were analyzed. Pattern coding was then performed to condense the data into fewer analytical concepts. Themes were subsequently identified from the pattern codes and compared by the 2 researchers across different experts’ data. Themes were chosen based on the following: (1) whether they were identified by both researchers and both types of experts, and (2) if they were not identified by both researchers, they should have had a significant number of codes from expert feedback that both researchers subsequently explored and agreed to include. Cohen κ coefficient [10], was calculated to measure the level of agreement between the 2 researchers. Calculations using the formula reflected a value of 0.69, showing moderate agreement between researchers for expert feedback [11].

For user feedback, attribute codes were first developed based on the main topics from the interview questions, organizing data by (1) living situation, (2) socioeconomic factors, (3) health factors and care needs, (4) coping methods, (5) hobbies and daily life, (6) main challenges, (7) use of technology, (8) use of mindfulness, (9) factors to consider for design, and (10) individual activity preferences. After this, 2 researchers, CELS and ZZ, separately used an inductive method—an open coding approach—before pattern coding was performed to condense the codes. Themes were subsequently identified from the pattern codes. The 2 researchers then came together to compare the themes, where both researchers reached a consensus about which themes to include after discussion and justification. Themes were chosen based on the following: (1) whether they were identified by both researchers and both types of users, and (2) if they were not identified by both researchers, they should have had a significant number of codes from user feedback that both researchers subsequently explored and agreed to include. Calculations using the Cohen κ coefficient reflected a value of 0.77, showing moderate agreement between researchers for user feedback [11].

Representative data consisting of expert and user quotes were included to support the themes identified. A cross-comparison analysis across themes from expert and user perspectives was performed to identify similarities. The analysis was reviewed by 4 other researchers. Data and all appropriate documents will be stored for a minimum of 10 years after the study is completed. All video recordings were transcribed in a timely fashion and removed from Microsoft Teams.
Results

Participants
A total of 5 dementia experts and 5 mindfulness experts who fulfilled the inclusion criteria completed the interviews, and all 10 interviews were analyzed. Most of the dementia experts (3/5, 60%) interviewed were facilitators at day activity centers, and all of them (5/5, 100%) were female. All the mindfulness experts (5/5, 100%) were mindfulness teachers, and most mindfulness experts (3/5, 60%) were male.

In all, 5 caregivers and 7 persons with early-stage dementia responded to the advertisements posted on Facebook dementia groups; 2 persons with dementia were excluded because they had difficulty understanding questions or experienced challenges using Microsoft Teams. The 10 users completed the interviews, and all 10 interviews were analyzed. A total of 60% (3/5) of caregivers cared for parents with dementia, whereas 40% (2/5) of caregivers were spouses of persons with dementia. Most caregivers (4/5, 80%) were female. Of the 5 people with early-stage dementia who were interviewed, 4 (80%) were supported by their spouses. They were able to manage daily tasks with some difficulty but had assistance from others. However, 20% (1/5) of persons with dementia lived independently in a retirement village, with assistance from paid caregivers who visited occasionally. Most persons with dementia (3/5, 60%) were female. Table 1 presents more information on the participants.

Table 1. Participant details (N=20).

<table>
<thead>
<tr>
<th>Type of participant</th>
<th>Value, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dementia expert (n=5)</td>
<td></td>
</tr>
<tr>
<td>Nursing home allied health professional</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Hospital staff</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Day activity center facilitator</td>
<td>3 (60)</td>
</tr>
<tr>
<td>Mindfulness expert (n=5)</td>
<td></td>
</tr>
<tr>
<td>Mindfulness meditation teacher</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Mindfulness movement teacher</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Mindfulness teacher</td>
<td>3 (60)</td>
</tr>
<tr>
<td>Persons with dementia (n=5)</td>
<td></td>
</tr>
<tr>
<td>Person with early-stage vascular dementia</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Person with early-stage dementia</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Person with early-stage frontotemporal dementia</td>
<td>3 (60)</td>
</tr>
<tr>
<td>Caregivers (n=5)</td>
<td></td>
</tr>
<tr>
<td>Carer for person with early-stage vascular dementia</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Carer for person with early-stage frontotemporal dementia</td>
<td>1 (20)</td>
</tr>
<tr>
<td>Carer with experience caring for person with early-stage dementia</td>
<td>3 (60)</td>
</tr>
</tbody>
</table>

Preferences for Conversational Agents That Support Dyadic Mindfulness
Most participants, all (5/5, 100%) caregivers and 60% (3/5) of persons with dementia, had past experiences using various forms of mindfulness practices. Although not all users had experiences with mindfulness, all 10 (5 caregivers and 5 persons with dementia) participants mentioned that they would use mindfulness practices after trying a mindful exercise. This indicated that mindfulness was appealing to the participants. With regard to practicing mindfulness as a pair, 80% (4/5) of the caregivers and 60% (3/5) of the persons with dementia expressed interest. However, 20% (1/5) of caregivers and 40% (2/5) of persons with dementia preferred to engage in mindfulness individually. According to most users, dyadic mindfulness was preferred.

Most participants, 60% (3/5) of caregivers and 80% (4/5) of persons with dementia, were familiar with using conversational agents. They used them to contact people, monitor the home, plan things using the calendar, set reminders and alarms, have conversations, tell the weather, and play music. They had past experiences with various types of conversational agents—Alexa, Siri, and Google Assistant. Caregivers and persons with dementia were asked to speculate on format preferences (conversational agents or videoconferencing software) to practice dyadic mindfulness; 60% (3/5) of caregivers and 60% (3/5) of persons with dementia preferred using conversational agents. Conversational agents were preferred, as they were something that the users could do on demand instead of having to arrange and schedule a zoom session. The users mentioned that conversational agents can be accessed whenever they wanted and that it did not matter whether it was at night. However, 40% (2/5) of caregivers and 40% (2/5) of persons with dementia preferred using videoconferencing owing to a more personal touch. Overall, mindfulness interventions designed for dyads could incorporate the use of conversational agents, given the majority’s familiarity with the technology and their likelihood of using it. However, the interviewees were
tech-savvy and may not be representative of the target population. It may be helpful for future studies to include participants who are not tech-savvy to understand their perspectives as well.

**Designing for Dyadic Mindfulness Conversational Agents—Common Expert and User Themes**

**Overview**

A total of 9 expert themes and 7 user themes were identified from the inductive thematic analysis of expert and user interviews to identify what was necessary in designing mindfulness conversational agents for persons with dementia and caregivers. The expert and user themes identified are listed accordingly in Textbox 1, where 5 common themes were identified.

**Textbox 1.** Comparison of expert and user themes.

<table>
<thead>
<tr>
<th><strong>Expert themes</strong></th>
<th><strong>User themes</strong></th>
<th><strong>Common themes</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Embracing ability</td>
<td>Alleviate stress</td>
<td>Enhancing accessibility</td>
</tr>
<tr>
<td>Evoking the 5 senses</td>
<td>Alleviate worries</td>
<td>Encouraging mindful mindset shift</td>
</tr>
<tr>
<td>Fostering engagement</td>
<td></td>
<td>Enhancing relationships</td>
</tr>
<tr>
<td>Creating habits</td>
<td></td>
<td>Cultivating positivity</td>
</tr>
</tbody>
</table>

**Enhancing Accessibility**

First, 60% (3/5) of dementia experts, 60% (3/5) of mindfulness experts, all (5/5, 100%) caregivers, and 80% (4/5) of persons with dementia emphasized the importance of keeping solutions designed for persons with dementia and their caregivers as accessible as possible, with activities explained in a simple and straightforward manner. It was necessary to avoid complex sentences, as they may cause difficulties in following instructions according to person with dementia D5. Second, providing guidance may also help facilitate accessibility, according to 40% (2/5) of dementia experts, 80% (4/5) of mindfulness experts, and 20% (1/5) of persons with dementia. Having expert guidance could help participants to have clarity and offer more support for people who need it, especially for those who may be progressively getting worse. Third, it would also be essential to consider how the activities designed for the dyad progress as the condition of the person with dementia declines, ensuring that users would still be able to use solutions with ease, as mentioned by 20% (1/5) of dementia experts, 80% (4/5) of caregivers, and all (5/5, 100%) persons with dementia. This could be accomplished by breaking down the tasks or simplifying programs for persons at different stages of dementia.

And it has to be as simple as possible, even for early stage, you know, as simple as possible. [Dementia expert D4]

So your version of simple and what actually works for me can be very, very different. [Person with dementia D1]

**Encouraging Mindful Mindset Shift**

Participants in the 4 groups believed that a mindful mindset shift may be able to benefit the dyad through present-moment awareness and having no attachments or aversions to experiences. Present-moment activities were used and preferred by 80% (4/5) of dementia experts, all (5/5, 100%) mindfulness experts, all (5/5, 100%) caregivers, and all (5/5, 100%) persons with dementia to enhance present-moment awareness. Worries that come with a progressing illness could be mitigated through the effects of present-moment activities, where the hustle and bustle stop for a moment and there is a sense of calm, changing the relationship with negative experiences to one of ease rather than one of struggle. In addition, 80% (4/5) of mindfulness experts and 40% (2/5) of persons with dementia identified the need to have no attachment or aversion to events, learn to have acceptance, and be comfortable with discomfort. Where pleasant or unpleasant feelings arise, thoughts should be allowed to come and go while being mindful—not reacting to or judging them.
This means that the participants would be comfortable with discomfort because a “different way of being is actually okay.” Moreover, 40% (2/5) of persons with dementia exemplified this by accepting and acknowledging uncomfortable experiences and waiting patiently for negative emotions to pass. Interventions for dyads should encourage a mindset shift, allowing dyads to stay in the present moment and learn to have no attachment or aversion to experiences.

**Being in the moment is just so important, isn’t it, for everybody, really, but I think specifically for someone who lives with an illness (dementia) and something which is obviously progressing...** [Dementia expert D1]

If I’m upset, I just acknowledge that I’m upset and I have a reason to be upset and wait till it passes, really. [Person with dementia D5]

### Enhancing Relationships

Social connectedness was widely identified as a factor that should be incorporated into dyad activities by 80% (4/5) of dementia experts, all (5/5, 100%) mindfulness experts, all (5/5, 100%) caregivers, and 80% (4/5) of persons with dementia. In particular, social connectedness between the dyads was important according to mindfulness experts M2 and M5. This was further emphasized by 60% (3/5) of caregivers and 40% (2/5) of persons with dementia who mentioned having a worsened relationship because of dementia needs and symptoms. Caregivers found it difficult to live with the person with dementia, grappling with the “huge personality change.” Similarly, 40% (2/5) of persons with dementia expressed negative changes in their relationships with their loved ones. As the condition of persons with dementia declines, they lose their abilities and have changes in their personalities, resulting in worsened relationships. It would be important for interventions to improve social connectedness between the dyads. Dyadic dynamics needs to be carefully designed to encourage a stronger bond between the pair, as mentioned by mindfulness experts M4 and M5.

**Quality of the relationship between caregiver and the person with dementia is vital.** [Mindfulness expert M2]

My husband says you have become a liability to me and that really, really hurt because we had been a team. [Person with dementia D1]

### Cultivating Positivity

A total of 40% (2/5) of dementia experts, all (5/5, 100%) mindfulness experts, 80% (4/5) of caregivers, and 40% (2/5) of persons with dementia identified the need to focus on the positive. They recommended activities that provided calm, appreciation, and loving kindness to enable the dyads to cultivate positivity. First, activities promoting calmness in the dyads were identified as essential by 40% (2/5) of dementia experts and 40% (2/5) of mindfulness experts. This resulted in a sense of peace, which was positive. Second, all (5/5, 100%) mindfulness experts, 40% (2/5) of caregivers, and 40% (2/5) of persons with dementia recommended activities that had elements of appreciation. Appreciating the “duality of human nature” and embracing situations in their entirety according to mindfulness expert M2 could be a way to bring some positivity to the lives of the dyads. Moreover, 40% (2/5) other caregivers mentioned noticing and appreciating the little moments of positivity instead of ruminating or dwelling on negative thoughts. In addition, 40% (2/5) of persons with dementia had a positive outlook on life despite having dementia and suggested that it was important to see the positive in the negative and learn to still be joyful and appreciative in the face of adversity. According to the experts and users, encouraging appreciation is essential for cultivating positivity in the dyads. Finally, 40% (2/5) of mindfulness experts recommended activities that had elements of loving kindness. Practicing loving kindness may also help to cultivate positivity. In essence, elements of positivity were recommended to be incorporated into interventions designed for dyads, enabling them to better cope with the challenges that come with dementia.

What I would really recommend building in would be something on using mindfulness in terms of appreciation and in terms of gratitude, so I think this sort of practices can be really helpful for people. [Mindfulness expert M5]

I’ve re-evaluated and rediscovered there’s a joy and a lightness in this new as well. I’ve looked death in the face and decided that I can dance and do and stand. [Person with dementia D1]

### Providing Simplified Tangible and Thought-Based Activities

According to experts and users, dyad activities could be tangible or thought based. In all, 60% (3/5) of dementia experts and 80% (4/5) of mindfulness experts recommended using mindful breathing, where the focus is on noticing the breath, as it is tangible and easy to grasp, which may be particularly useful for dyads. All (5/5, 100%) caregivers and all (5/5, 100%) persons with dementia would similarly do tangible activities like deep breathing, as it was a tangible and effective way to calm them down. In all, 20% (1/5) of mindfulness experts, 60% (3/5) of caregivers, and 60% (3/5) of persons with dementia would recommend or do other forms of tangible activities such as body scan meditation, where the focus is on the sensations one feels in the body. However, 40% (2/5) of caregivers and 40% (2/5) of persons with dementia mentioned not wanting to do body scan meditation, citing reasons of not understanding how to do the activity, lack of interest, and potential of feeling overwhelmed.

A total of 40% (2/5) of dementia experts and 60% (3/5) of mindfulness experts conducted thought-based activities that used reflection to help dyads generate insights into self. Mindfulness expert M4 explained that through mindful reflection, one would be aware of their thoughts, feelings, and sensations, enabling them to prevent automatic ruminating of thoughts and breaking the cycle of how fear-based thoughts affect unhealthy behavior. In addition, all (5/5, 100%) caregivers and 80% (4/5) of persons with dementia would do thought-based activities, focusing on reflection through dyadic gratefulness, where one reflects on positive things with a partner. Moreover, 80% (4/5) of caregivers and 60% (3/5) of persons with dementia were open to other thought-based activities such as letting go,
which helps users to be comfortable with discomfort and feel less overwhelmed. The activity encourages being open and accepting to difficult emotions and experiences, rationalizing the process using metaphors. However, 40% (2/5) of persons with dementia had difficulties understanding the abstract components that had metaphors. Overall, certain simplified tangible and thought-based activities may be suitable for dyads, depending on their individual needs.

Comparison With Prior Work
When compared with design recommendations identified in the literature—(1) dyadic mindfulness programs for persons with dementia and their caregivers as well as (2) conversational agents for persons with dementia—most of the themes uncovered were novel and reflected the needs from a holistic perspective, based on user and expert feedback. The existing design recommendations may have been created from the researchers’ perspectives, which may explain why most of them did not coincide with the user and expert themes, as seen in Table 2.

Table 2. Comparison of the dyadic mindfulness conversational agent design needs highlighted by users and experts with the existing recommendations from the literature.

<table>
<thead>
<tr>
<th>Expert themes</th>
<th>User themes</th>
<th>Design recommendation (literature)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common themes</td>
<td>common themes</td>
<td></td>
</tr>
<tr>
<td>Enhancing accessibility</td>
<td>Enhancing accessibility</td>
<td>Quality of speech recognition needs to be improved [8]</td>
</tr>
<tr>
<td>Encouraging mindful mindset shift</td>
<td>Encouraging mindful mindset shift</td>
<td>N/Aa</td>
</tr>
<tr>
<td>Enhancing relationships</td>
<td>Enhancing relationships</td>
<td>N/A</td>
</tr>
<tr>
<td>Cultivating positivity</td>
<td>Cultivating positivity</td>
<td>N/A</td>
</tr>
<tr>
<td>Providing simplified tangible and thought-based activities</td>
<td>Providing simplified tangible and thought-based activities</td>
<td>N/A</td>
</tr>
<tr>
<td>Embracing ability</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Evoking the 5 senses</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Fostering engagement</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Creating habits</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>N/A</td>
<td>Alleviate stress</td>
<td>N/A</td>
</tr>
<tr>
<td>N/A</td>
<td>Alleviate worries</td>
<td>N/A</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td>Allow separate interventions on different occasions [3]</td>
</tr>
<tr>
<td>N/A</td>
<td>Mindfulness guides need to be able to manage participants’ negative emotions [20]</td>
<td></td>
</tr>
<tr>
<td>N/A</td>
<td>Ensure that the needs of caregivers are met [3]</td>
<td></td>
</tr>
</tbody>
</table>

aN/A: not applicable.

Designing for Dyadic Mindfulness Conversational Agents—Expert Themes That Were Not Present in User Themes
A total of 4 themes were identified by experts and were not present in the user-identified themes. These include embracing ability, evoking the 5 senses, fostering engagement, and creating habits.

Embracing Ability
A total of 60% (3/5) of dementia experts and 20% (1/5) of mindfulness experts highlighted the importance of seeing dyads under a different light, focusing on abilities. They thought that after a diagnosis, dyads may focus on the negative aspects of dementia, and learning to see themselves and each other under a different lens could be useful. Similar sentiments were noted in the subthemes that identified role reversal, autonomy, and empowerment of dyads as important. Role reversal was brought up by 40% (2/5) of dementia experts and 20% (1/5) of mindfulness experts, where modifying the role for the persons with dementia could better engage them, preventing passivity. Role reversal was also used to help caregivers step out of their caring role to take a break. Second, encouraging autonomy was essential, and according to 40% (2/5) of dementia experts, it could come in the form of providing options for choices to be made by dyads. Third, according to 60% (3/5) of dementia experts, empowering dyads can help them to see that, although they are affected by the disease, there are many things that they can still do. A total of 60% (3/5) of dementia experts and 20% (1/5) of mindfulness experts mentioned the use of activities that stimulated the 5 senses. Activities used by dementia and mindfulness experts had an emphasis on the different senses, allowing for nonverbal communication needs to be met. Although it may be beneficial for activities designed for dyads

Actually he did tell us as well, that he started to see her in a different light... [Dementia expert D1]
to incorporate the 5 senses, care must be taken to prevent overloading their sensory capabilities.

We just use everything...all the senses and the garden and nature... [Dementia expert D5]

Fostering Engagement
According to 5 (100%) dementia experts, activities have to be engaging. It is important to create a positive association with the activity. Moreover, 60% (3/5) of dementia experts identified providing variety as an important component for engaging the users. Dementia experts D1, D3, and D4 provided a wide scope of activities with varying difficulties, providing flexibility for the users to choose what they preferred. Providing choices of format, frequency, and duration may also be important for fostering engagement.

It needs to be fun. [Dementia expert D2]

Creating Habits
Mindful habits can be created by incorporating mindfulness into current routines and using reminders to aid the practice. According to 40% (2/5) of dementia experts and 60% (3/5) of mindfulness experts, activities could be woven into existing routines to introduce new habits, rather than having them as separate activities. Reminders on how to perform mindfulness activities, on what the activities comprise, and to do the activities could also help to facilitate the formation of mindful habits, as mentioned by 40% (2/5) of dementia experts.

What is their routine and then build it into that, the activities list rather than separate because separate will never happen... [Dementia expert D3]

Designing for Dyadic Mindfulness Conversational Agents—User Themes That Were Not Present in Expert Themes
A total of 2 themes were identified by users and were not present in the expert-identified themes. These include alleviating stress and alleviating worries.

Alleviate Stress
Caregivers and persons with dementia are highly stressed, and interventions should address this need. A total of 60% (3/5) of caregivers had other caregiving duties in addition to caring for persons with dementia, and 40% (2/5) of caregivers also had other personal illnesses. These factors may have contributed to why 80% (4/5) of the caregivers felt highly stressed and unable to cope. In addition, 80% (4/5) of persons with dementia experienced high levels of stress because of their dementia symptoms, 2 of which also had other accompanying illnesses, adding on to their stress levels. Most caregivers and persons with dementia experienced high levels of stress, and 80% (4/5) of caregivers also had other caregiving duties in addition to caring for persons with dementia, and 40% (2/5) of caregivers also had other personal illnesses. These factors may have contributed to why 80% (4/5) of the caregivers felt highly stressed and unable to cope. In addition, 80% (4/5) of persons with dementia mentioned losing their abilities as one of the main challenges they faced, causing them to feel worried. These challenges reflected high levels of worries, and it is important for interventions to ensure that dyads are better able to manage their worries.

 Mostly I worry about him. [Caregiver C5]

Discussion
Principal Findings
A total of 5 themes outlining the user needs for designing dyadic mindfulness conversational agents were gathered from expert and user perspectives in this study. These themes could be used to design future dyadic mindfulness conversational agents.

Studies on dyadic mindfulness [15,16,18] and conversational agents for persons with dementia [23,24] were relatively scant, with a focus on assessing effectiveness rather than on the design and user experience of the intervention. Only 4 user needs were gathered from past studies on dyadic mindfulness and those on conversational agents for dementia, and it was unclear whether the needs were derived based on user feedback. Furthermore, there were no studies on dyadic mindfulness conversational agents; therefore, user needs could not be found. This revealed the lack of understanding of needs from the users’ perspective, which is what this study sought to address. The resulting themes, derived from an inductive process, provided interesting insights that helped meet the study’s objective—to fill the gap in the literature concerning the design needs of persons with dementia and their caregivers for a dyadic mindfulness conversational agent approach from expert and user perspectives.

The 5 common themes identified from expert and user feedback inform the design of mindfulness conversational agents for dyads. These themes incorporated the voices of dementia experts, mindfulness experts, persons with dementia, and caregivers. The value of approaches incorporating user opinions has been amply demonstrated in the literature, notably in the field of designing for persons with dementia and caregivers [21,25]. Gathering feedback from a holistic range of stakeholders had several advantages, including knowledge co-creation across dementia expertise, mindfulness expertise, and lived experiences. In this process, inclusivity and feasibility were enhanced by integrating user and expert perspectives, as they provided an insight into important considerations that no single stakeholder group could have identified alone.

From the 5 common themes, we noted a few important considerations. First, enhancing accessibility for both persons with dementia and caregivers is essential. This can be achieved by ensuring that the activities are simple and straightforward,

something more quickly than I have in the past. [Caregiver C2]

Alleviate Worries
Caregivers and persons with dementia had high levels of worries because of dementia, and interventions should address this issue. A total of 80% (4/5) of caregivers mentioned struggling with worries they had about the person they care for, revealing an inability to cope with the situation. In addition, all (5/5, 100%) persons with dementia mentioned losing their abilities as one of the main challenges they faced, causing them to feel worried. These challenges reflected high levels of worries, and it is important for interventions to ensure that dyads are better able to manage their worries.

Mostly I worry about him. [Caregiver C5]
where guidance is provided. Ensuring that they would still be able to perform the intervention as the condition of the person with dementia deteriorates should also be considered. However, what may be simple for designers may not be simple for persons with dementia, so there is a need to test and iterate the design of mindfulness conversational agents with dyads. Care must be taken to ensure that activities can be used by persons with dementia, even as their condition declines. Providing tangible activities that are less focused on cognitive capabilities could enable dyads to stay engaged as their condition deteriorates. Furthermore, to enhance accessibility, by providing guidance, technological implementation formats could be used for dyadic mindfulness. However, dyads may possess a wide spectrum of technological capabilities. For older adults who lack the high comfort and digital literacy skills of digital-native generations [12], it may be useful to provide more accessible and user-friendly technologies. This could be implemented using conversational agents, as users need not learn how to navigate a digital interface and simply have to talk to the conversational agent, which would guide them through an activity. There may be a learning curve to use conversational agents as well, but having a dyadic arrangement would be helpful as caregivers would be able to guide persons with dementia if needed. However, activities should be designed such that persons with dementia would be able to use them independently, as the activities should also benefit the caregiver, instead of having them worry about the person with dementia. This can be done by considering the difficulty of the activities and the phrasing of words used to convey the activities and ensuring that there is enough time for dyads to process and respond.

Second, encouraging a mindful mindset shift involving (1) staying in the present moment and (2) having no attachment or aversion to experiences is recommended. Present-moment activities were used by almost all the experts and users and should be incorporated into mindfulness interventions for dyads. Although having no attachment or aversions to experiences was identified by only 2 groups—80% (4/5) of mindfulness experts and 40% (2/5) of persons with dementia—it may be useful for dyads as 80% (4/5) of caregivers expressed negative situations where they felt like they were unable to cope, and no attachment and aversion in these situations could help them to manage. Similarly, persons with dementia D1, D3, D4, and D5 mentioned concerns about not being able to do things they were able to do before, and it may be essential to alleviate worries by promoting no attachment and aversion to their current situation. Interventions designed for dyads to promote no attachment and aversion, helping them to manage negative situations by learning to be comfortable with discomfort, may be a useful tool to cope as the condition of the person with dementia declines. Although both mindful mindset shifts may be useful, it may be easier to stay in the present moment with tangible-based activities such as deep breathing, which is easy to follow and focus on throughout an activity session. Having no attachment or aversion to experiences requires self-control, as it is not a typical response to experiences. When implementing activities of both kinds, designers must ensure that the activities are accessible. In addition, to embody a mindful mindset shift, where one lives in the present moment and has no attachment or aversion to experiences, extensive practice and habit formation are required. This could be facilitated through a technological approach such as conversational agents, where daily reminders and guided practices could aid habit formation, leading to a mindful mindset shift.

Third, there is a need to enhance relationships, as social connectedness was brought up by almost all the experts and users as essential. This may be particularly important between dyads, as the participants mentioned having worsened relationships because of dementia needs and symptoms. Most caregivers interviewed cared for parents with dementia, whereas most persons with dementia interviewed had spouses who cared for them. Whether it was a spousal relationship or child-parent relationship, persons with dementia and their caregivers had worsened relationships, which a dyadic mindfulness conversational agent could potentially address. In a previous study, dyadic mindfulness was facilitated by having one partner share whatever came to their mind and another partner listen to the other’s contemplation, which increased social connectedness [13]. Dyadic interaction could be further explored and integrated, for example, where dyads share with each other about something that they are grateful to each other for. However, it is important to note that this would encompass creating interventions for 2 different individuals and care must be taken to ensure that both their needs can be met. Although dyadic mindfulness conversational agents may provide opportunities to enhance relationships, dyadic arrangements may also prevent individuals from sharing their true concerns. For instance, mindfulness experts M4 and M5 mentioned that caregivers may not be willing to share the exasperation they feel from caregiving and persons with dementia may not be willing to share the fears they experience with dementia. Thought needs to be given on how to navigate both paths to manage the pros and cons of dyadic arrangements. Allowing separate interventions on different occasions, as suggested by the existing literature, may be able to mitigate this. Similarly, although dyadic mindfulness user preferences revealed that most users, 4 caregivers and 2 persons with dementia (6/10, 60%) would use mindfulness as a pair, 40% (4/10) of users, 1 caregiver and 3 persons with dementia, would prefer to do mindfulness individually.

Fourth, cultivating positivity is important and can be accomplished by providing calm, appreciation, and loving kindness. However, according to the 7 attitudes of mindfulness [14], cultivating positivity is not one of the attitudinal foundations of mindfulness practice. Nevertheless, dementia experts, mindfulness experts, persons with dementia, and caregivers mentioned the need to cultivate positivity, and it should still be incorporated into mindfulness interventions for dyads. It is also important to note that cultivating positivity does not necessarily mean to remove negativity, as mindfulness also promotes learning to be comfortable with discomfort. Instead, mindfulness for dyads should provide a balance of activities that cultivates a mindful mindset shift while also incorporating elements of positivity, as dyads face high levels of stress and worries with the diagnosis.

Finally, although it was recommended by both experts and users that tangible activities be provided, not all tangible activities are straightforward. Mindful breathing, a tangible activity, was...
easy to grasp and favored by majority of experts as well as all users. By contrast, body scan meditation was tangible, but not as straightforward. Some users mentioned feeling overwhelmed when practicing body scan meditation, whereas others did not know how to perform the activity. For example, when caregiver C2 tried body scan meditation, she mentioned feeling overwhelmed by the existing pain she felt from her body. “By focusing on pain...it was amplifying it,” making her “realise how bad it is and how widespread it is” when she focused on her body. Different activities may elicit different responses from the users, with some unexpected and unintended consequences. It may be important to provide choices of tangible activities from which the users may choose or provide alternative paths within the activity itself so that the users have alternatives, should a path not work well for them. It is also essential to note that although tangible activities, according to experts, may be easier for users to grasp, certain tangible activities could not be understood by persons with dementia. When abstract components and multiple concepts were introduced to persons with dementia, 40% (2/5) of them were unable to understand the activity. Even though tangible activities may be easier to grasp, the complexity of the activity and the phrasing used to explain each activity should be simplified, in particular, to cater to the needs of persons with dementia. Similarly, for thought-based activities, simplified versions were also recommended, as metaphors used to explain concepts were too difficult for 40% (2/5) of persons with dementia to follow. Care must be taken to ensure that the users are not overloaded with cognitive tasks that are too challenging.

By comparing and identifying the common themes mentioned by the 4 groups—dementia experts, mindfulness experts, persons with dementia, and caregivers—this study provided a comprehensive view of the themes needed to design dyadic mindfulness conversational agents. Most of the existing guidelines from the literature did not coincide with the common themes highlighted by experts and users.

**Conversational Agent Implementation Method**

Dyadic MBIs have been implemented in physical group settings, home settings through recordings, and guided telephone calls [15-19] but have yet to be implemented using more accessible formats and more advanced technologies such as conversational agents. Current implementation methods such as physical groups have accessibility limitations, as they require dyads to travel to a physical location, limiting the participants who are not able to be physically present. Situations such as the ongoing COVID-19 pandemic also make it difficult for people to attend in-person interventions. Physical groups and guided telephone calls also have scalability limitations and limited access, as they require a facilitator to be present, which limits the number of people who can access the intervention at each point in time. The users would also have to schedule a meeting to practice mindfulness but are not able to do it whenever they need it. By contrast, recordings of mindfulness allow for greater scalability and accessibility, as it can be used anywhere with 24/7 access, given that the users have adequate digital literacy, but do not allow for a personalized or guided experience, which physical group settings and guided telephone calls provide.

Using conversational agents may be able to address these needs while also providing for accessibility and scalability needs. Mindfulness conducted through conversational agents would be available to access in the users’ homes 24/7, not requiring them to travel to a physical location and allowing them to access it whenever needed. Conversational agents are also more accessible in terms of digital literacy needs, allowing dyads to simply speak to the conversational agent. The technology is scalable, allowing many people to simultaneously access mindfulness activities. Furthermore, it is also able to provide personalized and guided mindfulness activities, which recordings are not able to do.

This study revealed that dyadic mindfulness was of interest, and using novel implementation formats such as conversational agents, ideally one with a more personal touch, was preferred by most users. It was particularly interesting for conversational agents to be preferred by the users, as they were generally older. However, it is important to note that the users interviewed were able to navigate the internet space, use software such as Microsoft Teams independently, and have strong digital literacy skills, which may not be representative of the entire population. Moreover, 60% (3/5) of caregivers and 80% (4/5) of persons with dementia also had prior experience using conversational agents.

**Strengths and Limitations**

This study had multiple strengths, where it provided an understanding of the preferences and needs of persons with dementia and their caregivers with regard to dyadic mindfulness conversational agents. The insights gathered were novel, as dyadic mindfulness conversational agents had not been created before, and, therefore, the needs and preferences were not known. In addition, using both expert and user perspectives strengthened the study, providing insights from lived-experience pairs, as well as experts who had planned and executed dementia and mindfulness programs. However, there were a few limitations.

First, the results may not be representative of the target population, as persons with dementia and caregivers who volunteered to participate in the study, as well as mindfulness and dementia experts who were identified, were tech-savvy and able to navigate videoconferencing tools and social media. As the interviews were conducted digitally because of the COVID-19 pandemic, only interviewees who had adequate competence and confidence in using the digital tools were recruited for the study, limiting the generalizability of the results. As there would likely be interviewees who are not as technologically savvy, this resulted in the study’s sample being biased. It may be beneficial to include a more comprehensive range of users in future studies.

Second, as interviewees had no prior experience with mindfulness conversational agents for dyads, it was difficult to determine their needs when designing such interventions. To mitigate this, dyadic mindfulness conversational agents were explained to the participants and we had to rely on their knowledge and past experiences with conversational agents, dementia, and mindfulness programs. Consequently, the responses from the users did not have a strong emphasis on
conversational agent–specific feedback. This could be because the users did not have a prototype to experiment with and provided feedback based on their past experiences. Future research should use prototypes of mindfulness conversational agents for dyads to experience before obtaining their feedback. This would ensure that their needs are appropriately identified.

Third, it was challenging to recruit persons with dementia and caregivers for this study. Attempts to mitigate this were made by recruitment through multiple research networks and social network groups. This resulted in a small sample size of participants who volunteered to participate, which may have affected the validity of the study. Nevertheless, the richness of insight resulting from the thematic analysis of the interviews was satisfactory for the purpose of the study. For future studies, larger sample sizes should be recruited.

Fourth, there may have been a volunteer bias, as the persons with dementia and caregivers in this study volunteered to participate through the advertisements posted. They may not represent the general population that includes less empowered or motivated persons and, therefore, would affect the validity of the study. It may be useful for future studies to increase the number of volunteers to prevent volunteer bias.

Conclusions
This study helped fill the gap in the literature concerning the needs of persons with dementia and their caregivers for dyadic mindfulness conversational agents. The results of the semistructured interviews suggested that dyadic mindfulness for persons with dementia and their caregivers was preferred, with potential implementation formats such as conversational agents. Using technologies such as conversational agents could potentially enhance the accessibility, scalability, personalization, and guidance of dyadic mindfulness interventions. These interventions may be particularly important during circumstances such as the COVID-19 pandemic where digital technologies could enable mindfulness services to be provided to dyads. This study also revealed the needs to be considered when designing dyadic mindfulness conversational agents. A total of 5 themes were identified through expert and user interviews. The five themes, defined through an inductive process, included (1) enhancing accessibility, (2) cultivating positivity, (3) providing simplified tangible and thought-based activities, (4) encouraging mindful mindset shift, and (5) enhancing relationships. The development of mindfulness conversational agents for dyads should follow recommendations from both expert and user perspectives to ensure that dyadic needs can be met.

Acknowledgments
This study was funded by the Imperial College London President PhD Scholarship and supported by the Dementia Research Institute. Imperial College London was the sponsor for this study.

Data Availability
The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request. Please send an email to CELS.

Authors' Contributions
CELS (lead author) conceived the study, which was further refined by RAC and TP. CELS developed the detailed methodology and study design and obtained ethics approval. The recruitment of participants was led by CELS with assistance from EW and SD. Data collection was done by CELS, and data analysis was conducted by CELS, ZZ, and SS. Discrepancies between CELS and ZZ were resolved by RAC and TP. All authors (CELS, ZZ, SS, EW, SD, TP, and RAC) reviewed, commented, and edited the manuscript and approved the final version.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Expert and user topic guide focus.
[PDF File (Adobe PDF File), 120 KB - aging_v5i4e40360_app1.pdf ]

References


Abbreviations

MBI: mindfulness-based intervention
Lessons Learned From an Effectiveness Evaluation of Inlife, a Web-Based Social Support Intervention for Caregivers of People With Dementia: Randomized Controlled Trial

Hannah Liane Christie1*, BSc, MSc, PhD; Alieske Elisabeth Henrique Dam1*, BSc, MSc, PhD; Martin van Boxtel1, MD, PhD; Sebastian Köhler1, BSc, MSc, PhD; Frans Verhey1, MD, PhD; Marjolein Elisabeth de Vugt1, BSc, MSc, PhD

Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University, Maastricht, Netherlands

*these authors contributed equally

Corresponding Author:
Hannah Liane Christie, BSc, MSc, PhD
Department of Psychiatry and Neuropsychology
School for Mental Health and Neuroscience
Maastricht University
Postbus 616
Maastricht, 6200 MD
Netherlands
Phone: 31 456213078
Email: hannah.christie@maastrichtuniversity.nl

Abstract

Background: Informal care for people with dementia not only affects the well-being of the primary caregiver but also changes their roles and interactions with the social environment. New online interventions might facilitate access to social support. Recently, an online social support platform, Inlife, was developed in the Netherlands and aims to enhance social support and positive interactions in informal support networks.

Objective: This study aimed to evaluate the effectiveness of Inlife for caregivers of people with dementia.

Methods: A randomized controlled trial with 96 caregivers of people with dementia was performed. Participants were randomly assigned to the Inlife intervention or the waiting list control group. After 16 weeks of Inlife use, the waiting list control group could start using Inlife. Effects were evaluated at baseline (T₀), 8 weeks (T₁), and 16 weeks (T₂). The 16-week follow-up assessment (T₂) served as the primary endpoint to evaluate the results for the primary and secondary outcome variables evaluated with online self-report questionnaires. The primary outcomes included feelings of caregiver competence and perceived social support. The secondary outcomes included received support, feelings of loneliness, psychological complaints (eg, anxiety, stress), and quality of life.

Results: No significant improvements were demonstrated for the intervention group (n=48) relative to the control group (n=48) for the primary outcomes (feeling of carer competence: b=−0.057, 95% CI −0.715 to 0.602, P=.87; perceived social support: b=−15.877, 95% CI −78.284 to 46.530, P=.62) or any secondary outcome. This contrasts with our qualitative findings showing the potential of Inlife to facilitate the care process in daily life. Adherence was not optimal for all Inlife users. Additional per-protocol and sensitivity analyses also revealed no beneficial results for high active Inlife users or specific subgroups. Inlife users were more active when part of a larger network.

Conclusions: Researchers should be modest regarding the effectiveness of online caregiver interventions in terms of quantitative measures of well-being and quality of life. Nevertheless, online tools have the potential to facilitate the caregiver process in daily life. Lessons learned include the importance of harnessing the power of human interaction in eHealth, making use of the user’s social capital, and the need to develop research methods that can identify benefits in daily life that are ecologically valid for caregivers.

Trial Registration: Netherlands Trial Register NTR6131; https://trialsearch.who.int/Trial2.aspx?TrialID=NTR6131

International Registered Report Identifier (IRRID): RR2-10.1186/s13063-017-2097-y
dementia; Alzheimer's; neurodegenerative; caregiver; caregiving; digital health; eHealth; mHealth; Information communication technology; RCT; randomized controlled trial; social support; support platform; online platform; web-based; internet-based; peer-support; informal support; social interaction; support network

**Introduction**

**Dementia and Caregiving**

A substantial part of the care for people with dementia is provided by informal caregivers such as family, friends, and relatives [1]. Consequently, informal caring not only affects the well-being of the primary caregiver but also changes their roles and interactions with the social environment [2,3]. Although some involved caregivers report the benefits of caring, such as the strengthening and enrichment of mutual relationships, family cohesion, or personal growth [4], others face negative consequences of caring on their physical and mental health or experience increased burden due to the growing dependence of the people with dementia on their environment [2,5].

During the disease process, caregivers of people with dementia are at risk of becoming socially isolated since they might become homebound due to decreased mobility, memory problems, behavioral problems, denial of the disease, or experienced stigma [6-8].

To prevent social isolation and loneliness in caregivers and people with dementia, support is needed after a diagnosis has been made. Previous qualitative studies with carers of people with dementia have shown that there is a postdiagnostic care mismatch between the supply and demand of informal support. The authors advocated introducing early access to tools to improve open communication and facilitate positive social engagement in dementia care networks before carers might become overburdened and while they still have the resources to learn new caregiving skills [9,10].

Existing psychosocial interventions focus on psychoeducation, skill building, and psychotherapeutic counselling based on techniques derived from cognitive behavioral therapy [11-13]. It has been demonstrated that individually tailored, multicomponent interventions offered to both the caregiver and the people with dementia have positive effects on burden, anxiety, and depression [11-13]. However, the effects are generally small to moderate, and available studies are limited in their methodological quality [14].

**Online Interventions to Support Caregivers**

With the introduction of the Internet and social media in daily life in recent decades, online interventions have become a new avenue for caregiver support. Recent online intervention studies that contained multiple individual tailored elements of psychoeducation, peer support, skills training, and health assessment have demonstrated improved caregiver well-being, including confidence, self-efficacy, and lower levels of depression [15-17]. Therefore, online caregiver support might be an alternative to traditional counselling and support interventions for several reasons. Online tools are always available, regardless of distance, time, and mobility constraints, and provide easily accessible, low-cost support and caregiver empowerment [18]. Using online tools might elicit support-seeking behavior and engagement in social activities to cope better with stressful situations [19] or enhance feelings of competence [20,21] to deal with the challenges faced in the caregiving process. Additionally, through accessibility and widespread reach, online interventions can lower the threshold to involve the caregiver's social network and either seek or provide social support by increasing openness and positive interaction [9]. This might prevent social isolation and increase feelings of competence in caregivers. Although results from online network interventions are promising [22], the use in informal caregiver social networks is relatively new and not yet studied.

**Inlife**

Therefore, the online social support intervention “Inlife” was developed for caregivers of people with dementia and made available in the Netherlands [23]. Inlife intends to help caregivers of people with dementia overcome barriers to seeking help while also removing barriers for loved ones and other individuals involved in the dementia care network to offer help. Using the “Inlife platform,” the primary caregiver is encouraged to invite family, friends, and significant others into their personalized support circles. The functionalities in the intervention include Profile (personal information), Circles (layers of caregivers with different privileges), Helping (general overview to place and receive responses to help requests), Timeline, Calendar, Personal Messages, Care Book, and Compass (information about dementia-related topics). Details of these functionalities are described elsewhere [24].

Inlife is currently being called “Myinlife” after being adopted by a societal partner, the Dutch Alzheimer’s association (Alzheimer Nederland).

**Aim**

The aim was to evaluate the effectiveness of Inlife in a randomized controlled trial (RCT) over a 16-week period [24]. Primary outcomes were feelings of competence and perceived social support. Secondary outcomes were received support, feelings of loneliness, psychological complaints, and quality of life. We hypothesized that, compared with care as usual, use of the Inlife intervention would lead to change in both the primary and secondary outcome measures.

**Methods**

**Participants and Design**

Between June 2016 and June 2017, informal primary caregivers of people with dementia of all subtypes and stages were recruited via online advertisement, newsletters, and social media channels of the Dutch Alzheimer Association; regional dementia
community services; and memory clinics or other relevant care institutions. The inclusion criteria were (1) being a primary, informal caregiver of a person diagnosed with dementia of any subtype; (2) having Internet access; and (3) having basic (tablet) computer knowledge as assessed by the researcher. Participants were excluded if they were unavailable for longer than 4 weeks during the study period or had serious health problems incompatible with participation as assessed by the study staff.

After the baseline assessment, participants were randomly assigned to either the intervention or a waiting list control group. The intervention group participated in the Inlife intervention. The waiting list control group received care as usual and was able to start with Inlife after 16 weeks.

Randomization was performed using a computerized sequence generator for block randomization with variable sizes of 4, 6, and 8 (for details, see [24]). The follow-up assessments were completed online using a secure, custom-designed query system.

Ethical Approval
The study was approved by Ethical Committee of the Faculty of Psychology & Neuroscience of Maastricht University (ERCPN-172_20_03_2016_A1; Dutch Trial Register trial number: NTR6131) More detailed information about the study design is presented elsewhere [24]. The CONSORT eHEALTH checklist is presented as Multimedia Appendix 1.

Procedure
Participants were screened by telephone to check eligibility. Subsequently, the participating caregivers provided online informed consent. Assessments were collected online at 3 time points: pre-intervention (T₀), 8 weeks (T₁), and 16 weeks (T₂). The 16-week follow-up assessment served as a primary endpoint to compare group effects [24]. Because 2 weeks was deemed an adequate amount of time to become familiar with the platform, 2 weeks after registration on the Inlife platform, participants were contacted by phone to reflect on user experiences. This was done to facilitate engagement, stave off attrition, and resolve any initial queries about the platform.

Conditions
The intervention group had access to Inlife, an online social support platform for informal caregivers and people with dementia aimed at strengthening positive interactions and social support. All users had a secure username and password combination to access the website and complementary app for smartphones and tablets. Participants could use Inlife in a flexible manner and at their own pace. The platform remained accessible to them after the 16-week study period. Participants in the control group remained on the waiting list for 16 weeks and received care as usual. Care as usual entailed that the participants continued with any routine care they may receive, such as consultations with their general practitioner or dementia case manager. After the 16-week follow-up assessment (T₂), they had the opportunity to register on the Inlife platform.

Measures
A short overview of applied self-reported measurement instruments is provided in this section. More details can be found elsewhere [24].

Primary Outcomes
First, caregiver sense of competence was assessed using the Short Sense of Competence Questionnaire (SSCQ), which consists of 7 items that refer to caregivers’ feelings of being capable of caring for the people with dementia. The total score ranges from 0 to 7. The SSCQ has been evaluated as a valid and reliable instrument in caregiver research [25]. Next, perceived support was measured using the 12-item Multidimensional Scale of Perceived Social Support (MSPSS). The total score ranges from 12 to 84. The psychometric characteristics of the scores were sufficient in clinical [26] and nonclinical populations [27].

Secondary Outcomes
Received support was measured using the 12-item Social Support List-Interactions (SSL12-I); the total score ranges from 12 to 48. Good internal reliability has been previously demonstrated [28]. Feelings of loneliness were measured using the 11-item Loneliness Scale (LS), with total scores ranging from 11 to 33. The psychometric properties were sufficient [29]. The 6-item Lubben Social Network Scale (LSNS-6) was used to assess the number of friends and family ties. The total score ranges from 0 to 30. The LSNS-6 has been validated in an older sample [30]. Anxiety and depression symptoms were assessed using the Hospital Anxiety and Depression Scale (HADS). The total score ranges from 0 to 42, and the Dutch version of the HADS has demonstrated good reliability and validity [31]. Perceived stress was measured with the 10-item Perceived Stress Scale (PSS), with a total score ranging from 0 to 40. Sufficient psychometric quality has been demonstrated [32]. Perseverance time for the caregiver was measured by a single item, with scores ranging from 1 to 4. The scale was specifically validated for informal caregivers of people with dementia [33]. Domains of quality of life or capability were measured using the Investigating Choice Experiments for the Preferences of Older People Capability Measure for Older People (ICECAP-O). The summary score ranges from 0 (no capability) to 1 (full capability), and the scale has been sufficiently validated [34]. The impact of caring on quality of life was assessed using the Care-Related Quality of Life Scale (CarerQol). Scores range between 7 and 21. The psychometric properties were sufficient [35,36]. Furthermore, at baseline, (socio-)demographics of the caregivers and care recipients were collected including age, sex, education, and hours of contact with and hours caring for the people with dementia. Additionally, the number of clicks on the Inlife website was collected to measure actual usage of the platform (results reported elsewhere) [24].

Sample Size Calculation
The sample size calculation was based on previous intervention studies of caregivers of people with dementia with the SSCQ as an outcome measure [16,35], using differences between intervention and control groups at follow-up with an assumed mean effect size of Cohen d of 0.5 (medium effect). With an
alpha of .05 and power of 80%, we aimed to include 102 primary caregivers (51 participants per group). Allowing for a 20% loss to follow-up, we aimed to enroll a total of 122 caregivers into the study.

Statistical Analysis
Statistical analyses were conducted in SPSS Version 24.0 (IBM Corp, Armonk, NY). Before analysis, data were examined for missing values, outliers, and normality. Potential differences between the intervention and control groups in baseline characteristics and outcome variables at baseline and the 16-week follow-up, which might require adjustment for such differences, were tested using either t tests for continuous variables or Chi-square tests for categorical variables. Since there were missing values, we compared the baseline characteristics of study completers and the participants with missing values. A separate analysis revealed that missingness was related to the sex of the person with dementia. Since missing values were not completely at random, data were analyzed according to intention-to-treat (ITT) principles applying a multiple imputation strategy. We used the Markov chain Monte Carlo method in SPSS to produce 10 data sets. These were subsequently analyzed, and estimates were pooled using the Rubin rule [36]. Subsequently, we performed a per-protocol (PPT) analysis including only the caregivers in the analyses that used Inlife until 16 weeks in the intervention group. A subsequent sensitivity analysis was conducted by contrasting the high active versus the low active Inlife users with the control group. The intervention group was split into high active and low active Inlife users based on the median total number of clicks on the platform [24].

To test the differences in the outcome variables in the intervention group and waiting list control group, we performed linear regression analysis on the imputed data sets, with outcomes from the T2 assessment as dependent variables (ie, the primary endpoint at 16 weeks; after this period, the waiting list control group was able to start using Inlife). The primary and secondary outcome variables at the T2 follow-up were included in the model as dependent variables, and group was included as the between-subjects variable. Statistically significant baseline differences between the treatment arms (eg, age of the person with dementia) were included as covariates. Each outcome measure was assessed as a dependent variable in separate analyses. For a variable that was positively skewed, a cubic transformation was applied to better approximate a normal distribution. Subsequently, to test the changes in the primary outcome measures over time, data were analyzed performing a linear mixed model (LMM) on the nonimputed data set. This analysis estimates the fixed effects of the regression slopes, indicating the changes during the intervals \(T_0-T_1\) and \(T_1-T_2\) in the intervention and waiting list control groups. This procedure allows for modelling the rate of change in the primary outcome variables for the caregivers who did not receive the intervention compared with the caregivers who received the intervention \(T_0-T_2\). This analysis accounts for within-subject correlations between repeated measures using random (ie, individual-specific) effects, thus accounting for the hierarchical structure in the data (ie, time nested in individuals). Additionally, LMM handles missing values efficiently under the missing at random assumption if variables that are associated with missingness are included in the analyses using maximum likelihood estimates for the missing observations. Hence, it is suitable for ITT analysis [37]. Random effects for the intercept only were specified because likelihood ratio testing revealed that this model fit the data better than adding a random slope or adjusting for correlated residuals. To model the effect of the intervention on the primary outcome variables over time, we entered a group-by-time interaction term as a dummy variable for each of the follow-ups to allow for nonlinear effects. The model was adjusted for baseline differences (eg, age of the person with dementia) and associations with missingness (eg, sex of the person with dementia). All tests were 2-tailed with an alpha level of .05.

Results
Participants and Descriptive Statistics
A total of 475 caregivers were approached to participate in the study. In total, 379 people were excluded: 124 people were excluded due to ineligibility, and 255 people were excluded due to the fact that they declined to participate. Subsequently, 96 informal caregivers who signed informed consent were included. In total, 96 caregivers signed informed consent and were randomly assigned to either the Inlife intervention group \((n=48)\) or the waiting list control group \((n=48)\). Reasons for declining participation are described elsewhere [24]. Of the 96 randomized participants, 89 completed the 16-week follow up \((T_2)\). Figure 1 depicts the flowchart of study participation. The baseline characteristics for completers and noncompleters did not differ significantly. The baseline characteristics of the participants are presented in Table 1. For the majority of variables, there were no significant differences between the groups, except for age of the person with dementia \((t_{47}=-2.05, P=.04)\) and baseline scores on the SSCQ \((t_{47}=-2.65, P=.01)\) and ICECAP-O \((t_{47}=-2.81, P=.006)\), which were significantly different. Therefore, these variables were included as covariates in the analysis. The high active Inlife users had a larger number of circle members in their network (mean 9.4, SD 5.2) compared with the low active users (mean 3.3, SD 3.7).

---

https://aging.jmir.org/2022/4/e38656

JMI Aging 2022 | vol. 5 | iss. 4 | e38656 | p.213

(page number not for citation purposes)
Figure 1. Participant flow diagram according to Consolidated Standards of Reporting Trials (CONSORT).
Table 1. Characteristics of the intervention and control groups.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Intervention (n=48)</th>
<th>Control (n=48)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caregiver age (years), mean (SD)</td>
<td>58.1 (11.8)</td>
<td>55.7 (13.6)</td>
<td>.35</td>
</tr>
<tr>
<td>Caregiver age (years), range</td>
<td>26-84</td>
<td>22-82</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>17 (35.4)</td>
<td>14 (29.2)</td>
<td>.51</td>
</tr>
<tr>
<td>Female</td>
<td>31 (64.6)</td>
<td>34 (70.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Caregiver education (years), mean (SD)</strong></td>
<td>13.1 (5.1)</td>
<td>14.1 (5.7)</td>
<td>.36</td>
</tr>
<tr>
<td><strong>Length of caregiving (years), mean (SD)</strong></td>
<td>6.2 (7.8)</td>
<td>4.2 (4.6)</td>
<td>.14</td>
</tr>
<tr>
<td><strong>Weekly caregiving (hours), mean (SD)</strong></td>
<td>33.9 (44.9)</td>
<td>27.9 (48.0)</td>
<td>.53</td>
</tr>
<tr>
<td><strong>Caregiver relationship (n, %)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spouse/partner</td>
<td>21 (43.8)</td>
<td>13 (27.1)</td>
<td>.30</td>
</tr>
<tr>
<td>Son or daughter (-in law)</td>
<td>24 (50.0)</td>
<td>33 (68.7)</td>
<td></td>
</tr>
<tr>
<td>Brother or sister</td>
<td>1 (2.1)</td>
<td>1 (2.1)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2 (4.2)</td>
<td>1 (2.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Number of other involved caregivers, mean (SD)</strong></td>
<td>2.2 (2.2)</td>
<td>2.5 (2.3)</td>
<td>.59</td>
</tr>
<tr>
<td><strong>Care recipient age (years), mean (SD)</strong></td>
<td>75.0 (10.8)</td>
<td>79.1 (8.8)</td>
<td>.04</td>
</tr>
<tr>
<td><strong>Care recipient age (years), range</strong></td>
<td>47-91</td>
<td>55-92</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Care recipient education (years), mean (SD)</strong></td>
<td>10.2 (4.9)</td>
<td>11.1 (6.5)</td>
<td>.43</td>
</tr>
<tr>
<td><strong>Type of dementia (n, %)</strong></td>
<td></td>
<td></td>
<td>.47</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>25 (52.1)</td>
<td>19 (39.6)</td>
<td></td>
</tr>
<tr>
<td>Frontotemporal dementia</td>
<td>5 (10.5)</td>
<td>2 (4.2)</td>
<td></td>
</tr>
<tr>
<td>Vascular dementia</td>
<td>6 (12.5)</td>
<td>9 (18.6)</td>
<td></td>
</tr>
<tr>
<td>Dementia with Lewy bodies</td>
<td>1 (2.1)</td>
<td>3 (6.3)</td>
<td></td>
</tr>
<tr>
<td>Mixed dementia</td>
<td>2 (4.2)</td>
<td>5 (10.5)</td>
<td></td>
</tr>
<tr>
<td>Dementia not otherwise specified</td>
<td>9 (18.6)</td>
<td>10 (20.8)</td>
<td></td>
</tr>
<tr>
<td><strong>Living situation (n, %)</strong></td>
<td></td>
<td></td>
<td>.13</td>
</tr>
<tr>
<td>Home</td>
<td>41 (85.4)</td>
<td>35 (72.9)</td>
<td></td>
</tr>
<tr>
<td>Nursing home</td>
<td>7 (14.6)</td>
<td>13 (27.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Outcome variables at baseline and the 16-week follow-up</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SSCQ</strong>b, mean (SD)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=95)</td>
<td>3.8 (1.9)</td>
<td>4.7 (1.6)</td>
<td>.01</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>3.7 (1.9)</td>
<td>4.5 (1.7)</td>
<td>.05</td>
</tr>
<tr>
<td><strong>MSPSS</strong>c, mean (SD)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=96)</td>
<td>63.3 (13.0)</td>
<td>68.3 (11.7)</td>
<td>.055</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>63.8 (16.5)</td>
<td>66.5 (12.8)</td>
<td>.38</td>
</tr>
<tr>
<td><strong>SSL-12d, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=96)</td>
<td>29.8 (7.1)</td>
<td>31.8 (8.3)</td>
<td>.20</td>
</tr>
<tr>
<td>16-week follow-up (n=90)</td>
<td>29.6 (7.6)</td>
<td>32.6 (7.5)</td>
<td>.06</td>
</tr>
<tr>
<td><strong>LS</strong>e, mean (SD)**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=96)</td>
<td>3.9 (3.4)</td>
<td>3.6 (3.5)</td>
<td>.64</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>4.4 (3.8)</td>
<td>3.5 (3.9)</td>
<td>.28</td>
</tr>
</tbody>
</table>
### Variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Intervention (n=48)</th>
<th>Control (n=48)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LSNS-6(^f), mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=95)</td>
<td>17.0 (6.6)</td>
<td>18.4 (4.9)</td>
<td>.26</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>17.7 (6.6)</td>
<td>18.4 (6.1)</td>
<td>.33</td>
</tr>
<tr>
<td><strong>PSS(^g), mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=95)</td>
<td>14.2 (6.6)</td>
<td>13.7 (6.6)</td>
<td>.70</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>15.5 (6.9)</td>
<td>14.1 (8.0)</td>
<td>.40</td>
</tr>
<tr>
<td><strong>HADS(^h), mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=95)</td>
<td>22.8 (2.2)</td>
<td>22.0 (2.9)</td>
<td>.15</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>22.4 (2.6)</td>
<td>22.9 (2.5)</td>
<td>.37</td>
</tr>
<tr>
<td><strong>PT(^i), mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=95)</td>
<td>5.4 (0.9)</td>
<td>5.2 (1.2)</td>
<td>.42</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>4.6 (1.2)</td>
<td>4.5 (1.5)</td>
<td>.84</td>
</tr>
<tr>
<td><strong>CarerQol(^j), mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=95)</td>
<td>5.6 (1.8)</td>
<td>5.9 (2.1)</td>
<td>.47</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>6.0 (2.0)</td>
<td>6.0 (2.0)</td>
<td>.99</td>
</tr>
<tr>
<td><strong>ICECAP-O(^k), mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline (n=95)</td>
<td>0.78 (0.13)</td>
<td>0.84 (0.10)</td>
<td>.006</td>
</tr>
<tr>
<td>16-week follow-up (n=89)</td>
<td>0.77 (0.14)</td>
<td>0.83 (0.12)</td>
<td>.03</td>
</tr>
</tbody>
</table>

\(^a\)N/A: not applicable.
\(^b\)SSCQ: Short Sense of Competence Questionnaire.
\(^c\)MSPSS Multidimensional Scale of Perceived Social Support.
\(^d\)SSL-12: Social Support List 12-Interactions.
\(^e\)LS: Loneliness Scale.
\(^f\)LSNS-6: Lubben Social Network Scale.
\(^g\)PSS: Perceived Stress Scale.
\(^h\)HADS: Hospital Anxiety and Depression Scale.
\(^i\)PT: perseverance time.
\(^j\)CarerQol: Care and Quality of Life scale.
\(^k\)ICECAP-O: Investigating Choice Experiments for the Preferences of Older People Capability measure for Older People.

### Intervention Effects

Primary and secondary outcome measures were compared between groups (intervention vs waiting list control) after 16 weeks (T\(^2\)). Table 2 shows the results of the ITT analysis. Overall, no significant effects in favor of the intervention group compared with the control group were found. The PPT analysis also did not yield any significant effects. Considering our heterogeneous group, we performed a post hoc sensitivity analysis splitting the data into low active and high active users (Multimedia Appendix 2), caregiver relationship status (spouse vs children or others at distance), and living situation of the person with dementia (home vs institution). This analysis revealed no significant differences between the caregiver groups at T\(^2\) (see Multimedia Appendix 2).

For the primary outcome variables, we explored the rate of change over time, as shown in Table 3. For caregiver feelings of competence (SSCQ), the analysis demonstrated a significant group difference at baseline, but no significant overall interaction between group and time was found, indicating that change over time in caregiver feelings of competence was not explained by the intervention (see Multimedia Appendix 3). Similarly, the caregiver perceived support (MPSSS) analysis revealed no overall interaction between group and time, indicating that change over time in caregiver-perceived social support was not explained by the intervention (see Multimedia Appendix 4).
Table 2. Intention-to-treat (ITT) analyses comparing the intervention group (n=48) and control group (n=48) at the 16-week follow-up, showing the pooled statistics of the linear regression analyses of the 16-week follow-up outcome measures adjusted for the age of the person with dementia.

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>B</th>
<th>SE</th>
<th>Group P value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSCQ</td>
<td>-0.057</td>
<td>0.335</td>
<td>.87</td>
<td>-0.715 to 0.602</td>
</tr>
<tr>
<td>MSPSS</td>
<td>-15.877</td>
<td>31.841</td>
<td>.62</td>
<td>-78.284 to 46.530</td>
</tr>
<tr>
<td>SSL-12</td>
<td>-2.511</td>
<td>1.551</td>
<td>.11</td>
<td>-5.550 to 0.528</td>
</tr>
<tr>
<td>LSS</td>
<td>0.576</td>
<td>0.792</td>
<td>.47</td>
<td>-0.976 to 2.128</td>
</tr>
<tr>
<td>LSNS-6</td>
<td>-1.586</td>
<td>1.495</td>
<td>.29</td>
<td>-4.524 to 1.352</td>
</tr>
<tr>
<td>PSS</td>
<td>0.893</td>
<td>2.074</td>
<td>.67</td>
<td>-3.196 to 4.982</td>
</tr>
<tr>
<td>HADS</td>
<td>-0.541</td>
<td>0.536</td>
<td>.31</td>
<td>-1.591 to 0.509</td>
</tr>
<tr>
<td>PTE</td>
<td>0.031</td>
<td>0.524</td>
<td>.95</td>
<td>-1.011 to 1.073</td>
</tr>
<tr>
<td>CarerQol</td>
<td>0.097</td>
<td>1.022</td>
<td>.93</td>
<td>-1.928 to 2.122</td>
</tr>
<tr>
<td>ICECAP-O</td>
<td>-0.013</td>
<td>0.024</td>
<td>.59</td>
<td>-0.060 to 0.034</td>
</tr>
</tbody>
</table>

a ITT is based on a multiple imputation Markov chain Monte Carlo method with 10 iterations.
bAdjusted for baseline scores.
cSSCQ: Short Sense of Competence Questionnaire.
dThis variable was negatively skewed, and a cubic transformation was used.
eMSPSS: Multidimensional Scale of Perceived Support.
fSSL-12: Social Support List 12-Interactions.
gThis variable was skewed, but no transformation could better approach a normal distribution; therefore, raw data are presented.
hLSS: Loneliness Scale.
iLSNS-6: Lubben Social Network Scale.
jPSS: Perceived Stress Scale.
kHADS: Hospital Anxiety and Depression Scale.
lPT: perseverance time.
mCarerQol: Care and Quality of Life scale.
nICECAP-O: Investigating Choice Experiments for the Preferences of Older People Capability measure for Older People.

Table 3. Differences in rates of change over time for the primary outcome measures between the intervention and control groups (n=96), assessed using a linear mixed model (group, time, group x time, age of person with dementia, sex of person with dementia).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline B</th>
<th>95% CI</th>
<th>8-week follow-up B</th>
<th>95% CI</th>
<th>16-week follow-up B</th>
<th>95% CI</th>
<th>F statistic for group x time (df)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSCQ</td>
<td>-0.96</td>
<td>-1.65 to -0.26</td>
<td>0.32</td>
<td>-0.24 to 0.88</td>
<td>0.20</td>
<td>-0.37 to 0.77</td>
<td>0.66 (2,183)</td>
</tr>
<tr>
<td>MSPSS</td>
<td>-4.6</td>
<td>-10.35 to 1.07</td>
<td>-0.17</td>
<td>-4.39 to 4.05</td>
<td>1.17</td>
<td>-3.11 to 5.44</td>
<td>0.22 (2,181)</td>
</tr>
</tbody>
</table>

aTest of overall interaction between group (intervention, control) and time (baseline, 8-week follow-up, 16-week follow-up).
bSSCQ: Short Sense of Competence Questionnaire.
cThe control group is the reference group.
dMSPSS: Multidimensional Scale of Perceived Social Support.

Discussion

Main Findings

This RCT evaluated Inlife, an online social support intervention for caregivers of people with dementia. No significant improvements in the primary or secondary outcome variables were demonstrated for the intervention group relative to the control group. Additional PPT and sensitivity analyses revealed no beneficial results for the high active Inlife users or specific subgroups of caregivers (ie, spouse vs children or community dwelling vs institutionalized). However, the results indicated that users in general were more active when they had a larger...
number of people in their Inlife network. Furthermore, active users tended to have slightly longer care duration.

Lessons Learned

Care Circle Size Is Linked to Inlife Engagement

A first important lesson learned was how Inlife care circle size is linked to engagement with the platform. High active users tended to have a larger number of circle members on Inlife. Previous research also showed more beneficial results of an online support intervention for caregivers with larger informal social networks [37]. Similarly, Inlife might be especially helpful for individuals with an already large social network facilitating openness, involvement, and information flow in their daily life.

Conversely, it is possible that the Inlife intervention could have unintentionally induced a heightened awareness of one’s lack of available support in circles with a low level of responsiveness, which may otherwise not have been salient. This could be a contributing factor to the suboptimal compliance to Inlife. This finding is in line with our qualitative results described elsewhere [38]. Compliance issues are not uncommon in eHealth intervention studies and are likely to lessen their effectiveness [39].

Harnessing Social Capital and Embedding Inlife in the Care Context

It is important to note that the observed circle size across Inlife users is probably not representative of the actually available social capital. Indeed, the number of people in the social networks of the low and high active Inlife users was not significantly different at baseline, as measured by the LSNS-6 (see Multimedia Appendix 5). This indicates that some caregivers might still experience difficulties in recruiting people in their social network to join the Inlife platform. This hinders the full use of Inlife, and it raises the question of how we could help Inlife users to involve and expand their care circles and social capital. Additional offline guidance and information could help to extend access to available social capital and to overcome the existing threshold, stigma, and barriers to seeking support [9,23]. Health care professionals could help increase awareness of caregivers’ social support needs and already existing social capital, potentially also by making the link to local peer support services and offline networks, such as Alzheimer Cafés. This also provides the opportunity to connect online support to offline support, where potential Inlife users could be introduced to both health care professionals and peers who might provide upfront support as well as aid with actual Inlife usage to increase compliance rates and alleviate potential negative effects of low care circle responsiveness. Previous research has demonstrated that guidance by a coach could be a valuable contribution to online interventions, as blended eHealth interventions (that is, eHealth interventions that combine online and offline support elements) for caregivers appear to be more effective than nonblended interventions [15]. Moreover, integrating eHealth interventions for caregivers of people with dementia into existing (dementia) care organizations is an important determinant of their sustained implementation success [40]. This approach would necessitate thorough training and monitoring of the implementing health care professionals, as research has shown that implementer self-efficacy and sense of ownership are important predictors of sustainable implementation of online interventions for caregivers of people with dementia [41].

Considering Innovative Research Designs

Recently, researchers have been more critical of the gold standard of the RCT design for the evaluation of eHealth interventions [42]. Although the RCT is an established and proven method to gain insight into eHealth effectiveness and mechanisms, they are time and resource-intensive and often result in a lack of important, qualitative implementation data [43]. Currently, staying up to date with technological advancements is a challenge due to the expansive time frame of typical effectiveness studies. One way of developing eHealth interventions that are suitable to implementation when proven effective is by using more flexible research designs [44]. Inspiration for methods to evaluate (new functionalities to) eHealth interventions can be gained from industry, where many commercial platforms use real-time evaluations to gain feedback from users. These can include pop-ups, which ask the user to rate their experiences, or the launch of different versions of the same functionality in order to assess which of the 2 versions is more successful [40]. It is possible that the retrospective measurements used in this study could not capture the kinds of practical benefits (ie, increased time savings and positive interactions) that are highly important and ecologically valid for informal caregivers. Previous research has demonstrated that experience sampling methodology can provide both caregivers and clinicians with more detailed, ecologically valid information about caregiver well-being in real time [21]. This is a promising measurement method for the evaluation of online tools such as Inlife, as they are applied during daily life and show a more complete, variable picture, rather than a retrospective summary.

Strengths and Limitations

A first important strength of Inlife includes its development through co-design with potential users and its feasibility testing in a pilot evaluation, as recommended by the Medical Research Council framework [45]. This improved the usability and face validity of the platform. Second, the effects of Inlife were evaluated with a robust research design and statistical approach via an RCT that applied both ITT analysis and PPT analysis. However, this study also has several limitations (in addition to the issues relating to the study’s RCT design and retrospective measurements discussed in the section “Considering Innovative Research Designs”). First, the waiting list control design of the study might have affected group differences, as the waiting list control group could have had a longer anticipatory experience than the intervention group. However, eventually, the design enabled all interested caregivers to gain access to the Inlife platform. Second, this study’s sample was heterogeneous in nature, consisting of both spousal caregivers and children of people with dementia living either in the community or in care institutions. We selected this broad population considering the exploratory nature of our study to increase the generalizability of our findings to the general population. However, the power
of the present study was insufficient to reveal the effectiveness of Inlife for specific subsamples in separate analyses. This was because we were unable to recruit the intended (n=122) number of caregivers. Finally, although participants were recruited on a national level, they may reflect a distinct subgroup from the general population. For instance, the online nature of the study inevitably resulted in a highly educated sample with a relatively high computer literacy and with highly motivated individuals who potentially already had a special interest in online tools [46,47]. This recruitment method that relied on self-selection probably introduced a selection bias, in which individuals who were more motivated and proficient regarding the use of digital tools than average were sampled, potentially resulting in more positive effects than would be found in a more representative sample. However, given the lack of positive results in this study, the potential impact of this selection bias appears to be minimal.

Future Research Directions
First, future studies should determine methods to identify practical benefits, such as with qualitative research methods or momentary assessments in the flow of daily life. Efforts should be undertaken to develop research methods that can identify benefits in daily life that are ecologically valid for caregivers of people with dementia. Second, future studies should incorporate contextual factors, such as organizational implementation determinants and available social capital into the intervention design and implementation, to facilitate uptake and make use of the benefits of human interaction. It would be worthwhile to include other Inlife circle members and the people with dementia in the evaluation. Considering our promising qualitative findings in a caregiver subsample [38], we suggest adding more extensive qualitative research methods to gain more insight into the circumstances and factors that are required to make Inlife use effective. This might also enable tailoring the Inlife platform to individual needs by integrating persuasive design features, such as by providing personalized functionalities and tailored notifications that are relevant to the individual caregivers’ needs and are specific to their current stage in the caregiver process [48]. In this way, caregivers can become acquainted with the opportunities of the Inlife platform in accordance with their own needs and at their own pace.

Conclusions
The present RCT demonstrated no significant effects of Inlife on feelings of caregiver competence, social support, and measures of caregiver well-being. Nevertheless, online tools such as Inlife show promise to facilitate the care process in daily life, though researchers should be modest regarding their effectiveness in terms of quantitative measures of well-being and quality of life. Future eHealth studies should (1) exploit the power of human interaction in eHealth and facilitate use of the user’s social capital, (2) apply extensive qualitative process evaluations to unravel beneficial effects for specific subgroups of caregivers and gain insight into potential barriers for implementation in clinical practice, and (3) from the start of the intervention’s development, carefully consider how interventions should be implemented by including contextual factors into the design and evaluation process. Applying these lessons can help researchers develop eHealth interventions for caregivers of people with dementia (such as Inlife) that are better suited to both carer needs and their wider implementation contexts.

Acknowledgments
This study was supported in part by the Alzheimer Research Fund Limburg and ZonMw (70-73305-98-611). We thank the caregivers who participated in this study.

Data Availability Statement
The data that support the findings of this study are available on request from the corresponding author, HLC. The data are not publicly available due to restrictions (eg, their containing information that could compromise the privacy of research participants).

Authors’ Contributions
AEHD, MEvD, MVb, and FV designed the study. AEHD collected the data and performed the analyses. HLC updated, reworked, and submitted the manuscript for publication, after AEHD’s departure from the research group. SK assisted with the data analysis. MEvD and MVb supervised the data collection and analysis. MEvD, MVb, SK, HLC, AEHD, and FV reviewed and approved this manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 378 KB - aging_v5i4e38656_app1.pdf ]

Multimedia Appendix 2
Sensitivity analyses of psychosocial outcome measures by group.
[DOCX File - aging_v5i4e38656_app2.docx ]
Multimedia Appendix 3
Change over time for feelings of competence (Short Sense of Competence Questionnaire [SSCQ]) for the intervention and control groups.

Multimedia Appendix 4
Change over time for perceived social support (Multidimensional Scale of Perceived Social Support [MSPSS]) for the intervention and control groups.

Multimedia Appendix 5
Baseline characteristics of the high active and low active Inlife users in the intervention group.

References


Abbreviations

- **CarerQol**: Care-Related Quality of Life Scale
- **HADS**: Hospital Anxiety and Depression Scale
- **ICECAP-O**: Investigating Choice Experiments for the Preferences of Older People Capability Measure for Older People
- **ITT**: intention to treat
- **LMM**: linear mixed model
- **LS**: Loneliness Scale
- **LSNS-6**: 6-item Lubben Social Network Scale
- **MSPSS**: Multidimensional Scale of Perceived Social Support
- **PPT**: per-protocol
- **PSS**: Perceived Stress Scale
- **RCT**: randomized controlled trial
- **SSCQ**: Short Sense of Competence Questionnaire
- **SSL12-I**: 12-item Social Support List-Interactions
Evolving Hybrid Partial Genetic Algorithm Classification Model for Cost-effective Frailty Screening: Investigative Study

John Oates1*, PhD; Niusha Shafiabady2*, BEng (Hon), MSc, PhD; Rachel Ambagtsheer3, PhD; Justin Beilby3, MD, PhD; Chris Seiboth4, BA; Elsa Dent3, PhD

1Torrens University, Ultimo, Australia
2College of Engineering, Information Technology and Environment, Charles Darwin University, Haymarket, NSW, Australia
3Torrens University, Adelaide, Australia
4Lutheran Services, Milton, Australia
*these authors contributed equally

Corresponding Author:
Niusha Shafiabady, BEng (Hon), MSc, PhD
College of Engineering, Information Technology and Environment
Charles Darwin University
815 George Street
Haymarket, NSW, 2000
Australia
Phone: 61 80474147
Email: niusha.shafiabady@cdu.edu.au

Abstract

Background: A commonly used method for measuring frailty is the accumulation of deficits expressed as a frailty index (FI). FIs can be readily adapted to many databases, as the parameters to use are not prescribed but rather reflect a subset of extracted features (variables). Unfortunately, the structure of many databases does not permit the direct extraction of a suitable subset, requiring additional effort to determine and verify the value of features for each record and thus significantly increasing cost.

Objective: Our objective is to describe how an artificial intelligence (AI) optimization technique called partial genetic algorithms can be used to refine the subset of features used to calculate an FI and favor features that have the least cost of acquisition.

Methods: This is a secondary analysis of a residential care database compiled from 10 facilities in Queensland, Australia. The database is comprised of routinely collected administrative data and unstructured patient notes for 592 residents aged 75 years and over. The primary study derived an electronic frailty index (eFI) calculated from 36 suitable features. We then structurally modified a genetic algorithm to find an optimal predictor of the calculated eFI (0.21 threshold) from 2 sets of features. Partial genetic algorithms were used to optimize 4 underlying classification models: logistic regression, decision trees, random forest, and support vector machines.

Results: Among the underlying models, logistic regression was found to produce the best models in almost all scenarios and feature set sizes. The best models were built using all the low-cost features and as few as 10 high-cost features, and they performed well enough (sensitivity 89%, specificity 87%) to be considered candidates for a low-cost frailty screening test.

Conclusions: In this study, a systematic approach for selecting an optimal set of features with a low cost of acquisition and performance comparable to the eFI for detecting frailty was demonstrated on an aged care database. Partial genetic algorithms have proven useful in offering a trade-off between cost and accuracy to systematically identify frailty.

(JMIR Aging 2022;5(4):e38464) doi:10.2196/38464

KEYWORDS

machine learning; frailty screening; partial genetic algorithms; SVM; KNN; decision trees; frailty; algorithm; cost; model; index; database; ai; ageing; adults; older people; screening; tool
Introduction

Genetic algorithms (GA) are a general-purpose computational optimization method inspired by the evolution mechanism in nature. They are one of the most popular metaheuristic search algorithms and have been used for a variety of applications, including synthetic data generation, feature selection, and to solve complex equations [1]. In this study, genetics algorithms have been applied to identify features that offer a suitable trade-off between cost and accuracy.

Within the context of global population aging, the number of older people who will live a significant proportion of their lives with frailty is growing rapidly [2]. Frailty is problematic for older people and the societies in which they live due to the elevated risks associated with the syndrome, including poor health outcomes [3] and additional use of health and aged care services [4-7], leading to inflated health care costs [8-10]. However, emerging research suggests that frailty is a highly dynamic [11,12] and potentially modifiable state with appropriate intervention [13,14]. Screening for early detection is proposed to increase the likelihood that the worst impacts of frailty can be lessened [4,15,16].

There are 2 main approaches to identifying frailty: the frailty phenotype (FP) and the frailty index (FI) [17]. However, these established approaches have known drawbacks, requiring significant time investment, face-to-face interaction, and specific data items to be collected [18]. Recently, an electronic frailty index (eFI) was proposed [19] that has the potential to achieve greater efficiencies over face-to-face models when applied to administrative data sets, but the need to ensure a minimum set of items adhering to prespecified criteria remains a barrier to implementation. For example, previous research has shown that although it is possible to calculate and construct an eFI based on an aged care administrative data set, a significant proportion of the items require manual calculation to ensure accuracy and improve quality [20]. Clearly, it would be preferable to identify automated techniques capable of delivering comparable accuracy and quality but with greater efficiency. Consequently, this study aimed to apply a sophisticated genetic algorithm technique to identify an optimal predictor of the calculated eFI.

Methods

Study Design, Participants, and Setting

This retrospective study utilized a data set previously compiled [21] from the administrative database of 10 residential aged care facilities located in Queensland, Australia. Participants were included in the study if they were aged 75 years or older and had completed an Aged Care Funding Instrument (ACFI) assessment within the previous 3 years.

Ethical Considerations

A waiver of consent for the initial study was obtained from the Human Research Ethics Committee of Torrens University Australia (application H11/19), which declared the study exempt under National Statement 5.1.22 (secondary use of deidentified administrative data) due to the pragmatic nature of the study. Because this is a secondary study of the same data, the approval extends to this study. Moreover, this study adheres to the Australian National Statement on Ethical Conduct in Human Research.

Frailty Outcome Measure

An eFI was previously calculated for this data [21] based on a formulation originally specified by Clegg et al [22]. Care was taken to ensure the included deficits adhered to the criteria recommended by Searle and colleagues [23], which resulted in 32 of the 35 deficits being extracted from unstructured patient notes and only 3 being derived from the ACFI data. The binary frailty classification was derived using a threshold of 0.21 (ie, frailty defined as >0.21) [24].

Screening Test Construction

Genetic algorithms are an optimization technique [1] applied in machine learning to filter a set of features that are used to construct a classification model. During training, a classification algorithm is tuned on a training set, and the success of attaining a generalized predictive algorithm is then verified by measuring the classification errors in the test set.

Genetic algorithms leverage the observation that classification models often perform better when they are trained on a subset of the available features. Which subset of features to use, however, is not obvious. Genetic algorithms start with a population of randomly generated subsets of features, or chromosomes, that are all independently used to generate classification models. The chromosomes from the population that generated the best performing models are allowed to combine, or breed, to form a new generation of the population, while the worst performing ones are removed completely. The process continues until either a predefined number of generations have been trained or the performance of the models has plateaued. Once training is complete, the best-performing model is deployed using only the naturally selected subset of the available features.

While genetic algorithms are good at selecting an optimal subset of features, they select the features based on maximizing the classification accuracy of a generated model. The cost of acquiring the various features is not factored into the choice of features, even if the performance of less expensive features is close to that of their more expensive counterparts. In this study, the cost of a feature is the combination of the effort, monetary cost, and patient risk involved in capturing the values. We want to minimize the number of expensive features chosen to form the model but allow as many low-cost features to be used as is necessary to gain acceptable performance of the model.

To achieve the inclusion of low-cost features in the classification model, the standard genetic algorithm training configuration illustrated in Figure 1 is modified as illustrated in Figure 2.
This modification is performed every time a model is trained for every member of the population trialed by the genetic algorithm. When the genetic algorithm trains a model, it passes a subset of the available training records to the classification model’s training algorithm. The low-cost feature values for each record need to be added to the selected training records before commencing the training. The genetic algorithm trains the classification model for each chromosome multiple times with different subsets of the training records and determining the performance of each model using records not used in training that instance. As with the training records, the low-cost features need to be added to the records used to determine a model’s performance. The performance of the chromosome is calculated as the average performance of all the models built from different subsets of the training records. This process is called n-fold cross validation, where n is the number of models built. In this study, 3-fold cross validation was used because it ensured a good balance between performance and the time it took to build the models.

Four types of classification models were optimized using partial genetic algorithms: logistic regression, support vector machines, random forest, and decision trees. These algorithms are popular choices for classification because they have proven successful in generating generalized models for a wide range of applications [20]. Logistic regression is a statistical modeling technique whereby a linear combination of the input features is found during training, which models the logarithm of the odds that a binary outcome is in the true state. A support vector machine (SVM) aims to learn a multidimensional hyperplane that separates the set of records given to it for training. Predictions are made by placing the candidate record in the same multidimensional classification space and determining which side of the hyperplane it maps to. SVM was developed in the 1990s and has since enjoyed success in many real-world
applications, including pattern recognition [25], text classification [26], and bioinformatics. Decision trees employ a divide and conquer strategy. A tree is formed of nodes, and each node performs a comparison of a single input feature and a threshold if the variable is continuous or a state if the feature is discrete. The outcome of the comparison determines the choice of the next node, which either performs a new comparison or terminates the tree with a given classification. During training, the set of training records are used to find comparisons at each node that gain the most information by reducing entropy in the outcomes by the greatest amount. Subsequent training predictions are made by feeding records into the root node and determining the classification of the terminating node where the record exits the tree. Random forest is a meta form of decision trees, where the output is determined by a vote between many trees. The trees are built using different methods to ensure they are not replicas of each other.

The software was written in Python and the models were built using the sklearn module (version 1.0.2) and the genetic_selection module from sklearn-genetic (version 0.5.1).

Results

Model Generation

Of the 69 features considered, 34 were extracted directly from the ACFI assessment and 35 were the values used to calculate the eFI. Two of the ACFI features, Psychogeriatric Assessment Scales (PAS) score and Cornell Scale, were excluded as they had a high percentage of missing values (PAS score 36%, Cornell Scale 42%). The remaining 32 ACFI assessment features had no missing values and were categorized as low cost of acquisition features. Of the 35 features used to calculate the eFI, 32 were extracted by an automated search for key words in the unstructured patient notes, followed by manual inspection and verification by a clinician. These were categorized as having a high cost of acquisition. The remaining 3 features used to calculate the eFI were direct combinations of ACFI features. As the calculation of these features could be fully automated, they were included with the low-cost features. A total of 4 sets of low-cost features were considered: (1) ACFI features + the low-cost eFI features; (2) the low-cost eFI features; (3) no low-cost features; and (4) a set of features chosen from the low-cost features using genetic algorithms. A different set was found for each of the classification algorithms.

Sixteen scenarios were trialed, comprising each of the aforementioned 4 sets of low-cost features for each of the 4 classification algorithms. For each scenario, the partial genetic algorithm was used to optimize the classification algorithm with different limits placed on the number of high-cost features. The limits were varied sequentially from 1 to 32, which was the number of candidate high-cost features. The performance of each of the 32 algorithms generated for each scenario were plotted on a single graph. The graphs for each scenario are plotted in Figures 3-6.

When comparing the graphs for each classification model, logistic regression outperformed decision trees in every scenario and SVM and random forest in almost all scenarios. Tables 1-3 demonstrate the numeric comparison of the 16 scenarios when 5, 10, and 15 of the high cost of acquisition features were used.

The option of “No low-cost” features was provided to determine how much predictive value the low-cost features were adding to the classification. As expected, this option performed the worst for all the classification algorithms, confirming that the low-cost features were adding value. Next, models were built using only the 3 low-cost eFI features as fixed features. This improved the accuracy of the logistic regression algorithm to 97% when almost all the eFI features were included (Table 4). Although this is a good outcome, a model built using so many of the high-cost features was not the goal of this study.

A genetic algorithm works by selecting an optimal subset of all the features made available to it. This characteristic was the motivation behind building a version of the models in 2 stages. In the first stage, a standard, nonpartial, genetic algorithm was used on the low-cost features to find an optimal combination. These models performed so poorly (Table 5) that they could not be used without further improvement. The combination of features used to generate these models (Multimedia Appendices 1-3) was then employed as the fixed features in the partial genetic algorithm during the second stage. The models in the second stage performed surprisingly poorly, showing no difference from the models built without any low-cost features, regardless of the classification model used.

Using all the low-cost features in a partial genetic algorithm yielded the best overall results and matched the 97% accuracy achieved by the models that used the low-cost eFI features when the model was able to select most of the high-cost eFI features. At 10 features, however, the extra low-cost features allowed the algorithm to increase its sensitivity from 82.7% to 89.3% and specificity from 81.7% to 86.7%.
Figure 3. Logistic regression optimized with a partial genetic algorithm. ACFI: Aged Care Funding Instrument; EFI: electronic frailty index; GA: Genetic algorithm; LR: logistic regression; npa: negative percent agreement; ppa: positive percent agreement.
Figure 4. Support vector machine optimized with a partial genetic algorithm. ACFI: Aged Care Funding Instrument; EFI: electronic frailty index; GA: Genetic algorithm; npa: negative percent agreement; ppa: positive percent agreement; SVM: support vector machine.
Figure 5. Decision tree optimized with a partial genetic algorithm. ACFI: Aged Care Funding Instrument; DT: decision tree; EFI: electronic frailty index; GA: Genetic algorithm; npa: negative percent agreement; ppa: positive percent agreement.
Figure 6. Random forest optimized with a partial genetic algorithm. ACFI: Aged Care Funding Instrument; EFI: electronic frailty index; GA: Genetic algorithm; npa: negative percent agreement; ppa: positive percent agreement; RF: random forest.
<table>
<thead>
<tr>
<th>Features</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPA(^a)</th>
<th>NPA(^b)</th>
<th>Accuracy</th>
<th>F1(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACFI(^d) + low-cost eFI(^e)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>76</td>
<td>75</td>
<td>71.4</td>
<td>79.2</td>
<td>75.6</td>
<td>73.2</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>76</td>
<td>61.7</td>
<td>67.3</td>
<td>71.3</td>
<td>69.6</td>
<td>64.3</td>
</tr>
<tr>
<td>Decision tree</td>
<td>73.3</td>
<td>63.3</td>
<td>65.5</td>
<td>71.4</td>
<td>68.8</td>
<td>64.4</td>
</tr>
<tr>
<td>Random forest</td>
<td>80</td>
<td>71.7</td>
<td>74.1</td>
<td>77.9</td>
<td>76.3</td>
<td>72.9</td>
</tr>
<tr>
<td><strong>Low-cost eFI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>80</td>
<td>81.7</td>
<td>76.6</td>
<td>84.5</td>
<td>80.7</td>
<td>79</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>74.7</td>
<td>66.7</td>
<td>67.8</td>
<td>73.7</td>
<td>71.1</td>
<td>67.2</td>
</tr>
<tr>
<td>Decision tree</td>
<td>76</td>
<td>53.3</td>
<td>64.0</td>
<td>67.1</td>
<td>65.9</td>
<td>58.2</td>
</tr>
<tr>
<td>Random forest</td>
<td>72</td>
<td>70</td>
<td>66.7</td>
<td>75</td>
<td>71.1</td>
<td>68.3</td>
</tr>
<tr>
<td><strong>No low-cost features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>78.7</td>
<td>66.7</td>
<td>71.4</td>
<td>74.7</td>
<td>73.3</td>
<td>69</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>82.6</td>
<td>55</td>
<td>71.2</td>
<td>69.7</td>
<td>70.4</td>
<td>62.2</td>
</tr>
<tr>
<td>Decision tree</td>
<td>84</td>
<td>55</td>
<td>73.3</td>
<td>70</td>
<td>71.1</td>
<td>62.9</td>
</tr>
<tr>
<td>Random forest</td>
<td>76</td>
<td>65</td>
<td>68.4</td>
<td>73.1</td>
<td>71.1</td>
<td>66.7</td>
</tr>
<tr>
<td><strong>Genetically selected low-cost features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>80</td>
<td>66.7</td>
<td>72.7</td>
<td>75</td>
<td>74.1</td>
<td>69.6</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>82.7</td>
<td>58.3</td>
<td>72.9</td>
<td>71.3</td>
<td>71.9</td>
<td>64.8</td>
</tr>
<tr>
<td>Decision tree</td>
<td>76</td>
<td>68.3</td>
<td>69.5</td>
<td>75</td>
<td>77.2</td>
<td>68.9</td>
</tr>
<tr>
<td>Random forest</td>
<td>81.3</td>
<td>75</td>
<td>76.3</td>
<td>80.3</td>
<td>78.5</td>
<td>75.6</td>
</tr>
</tbody>
</table>

\(^a\)PPA: positive percent agreement.

\(^b\)NPA: negative percent agreement.

\(^c\)F1: F-score.

\(^d\)ACFI: Aged Care Funding Instrument.

\(^e\)eFI: electronic frailty index.
Table 2. Performance of the 12 scenarios with 10 high-cost features.

<table>
<thead>
<tr>
<th>Features</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPA&lt;sup&gt;a&lt;/sup&gt;</th>
<th>NPA&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Accuracy</th>
<th>F1&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACFI&lt;sup&gt;d&lt;/sup&gt; + low-cost eFI&lt;sup&gt;e&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>89.3</td>
<td>86.7</td>
<td>86.7</td>
<td>89.3</td>
<td>88.1</td>
<td>86.7</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>85.3</td>
<td>80.0</td>
<td>81.4</td>
<td>84.2</td>
<td>83</td>
<td>80.7</td>
</tr>
<tr>
<td>Decision tree</td>
<td>65.3</td>
<td>61.7</td>
<td>58.7</td>
<td>68.1</td>
<td>63.7</td>
<td>60.2</td>
</tr>
<tr>
<td>Random forest</td>
<td>80</td>
<td>70</td>
<td>73.9</td>
<td>76.9</td>
<td>75.6</td>
<td>71.8</td>
</tr>
<tr>
<td><strong>Low-cost eFI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>82.7</td>
<td>81.7</td>
<td>79</td>
<td>84.9</td>
<td>82.2</td>
<td>80.3</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>81.3</td>
<td>81.7</td>
<td>77.8</td>
<td>84.7</td>
<td>81.5</td>
<td>79.7</td>
</tr>
<tr>
<td>Decision tree</td>
<td>81.3</td>
<td>50</td>
<td>68.2</td>
<td>67</td>
<td>67.4</td>
<td>57.7</td>
</tr>
<tr>
<td>Random forest</td>
<td>84</td>
<td>63.3</td>
<td>76</td>
<td>74.1</td>
<td>74.8</td>
<td>69.1</td>
</tr>
<tr>
<td><strong>No low-cost features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>72</td>
<td>86.7</td>
<td>71.2</td>
<td>87.1</td>
<td>78.5</td>
<td>78.2</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>77.3</td>
<td>78.3</td>
<td>73.4</td>
<td>81.7</td>
<td>77.8</td>
<td>75.8</td>
</tr>
<tr>
<td>Decision tree</td>
<td>81.3</td>
<td>70</td>
<td>75</td>
<td>77.2</td>
<td>76.3</td>
<td>72.4</td>
</tr>
<tr>
<td>Random forest</td>
<td>80</td>
<td>66.7</td>
<td>72.7</td>
<td>75</td>
<td>74.1</td>
<td>69.6</td>
</tr>
<tr>
<td><strong>Genetically selected low-cost features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>82.6</td>
<td>80</td>
<td>78.6</td>
<td>83.8</td>
<td>81.5</td>
<td>79.3</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>78.7</td>
<td>78.3</td>
<td>74.6</td>
<td>81.9</td>
<td>78.5</td>
<td>76.4</td>
</tr>
<tr>
<td>Decision tree</td>
<td>77.3</td>
<td>60</td>
<td>67.9</td>
<td>70.7</td>
<td>69.6</td>
<td>63.7</td>
</tr>
<tr>
<td>Random forest</td>
<td>78.7</td>
<td>66.7</td>
<td>71.4</td>
<td>74.7</td>
<td>73.3</td>
<td>69</td>
</tr>
</tbody>
</table>

<sup>a</sup>PPA: positive percent agreement.

<sup>b</sup>NPA: negative percent agreement.

<sup>c</sup>F1: F-score.

<sup>d</sup>ACFI: Aged Care Funding Instrument.

<sup>e</sup>eFI: electronic frailty index.
Table 3. Performance of the 12 scenarios with 15 high-cost features.

<table>
<thead>
<tr>
<th>Features</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPA&lt;sup&gt;a&lt;/sup&gt;</th>
<th>NPA&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Accuracy</th>
<th>F1&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ACFI&lt;sup&gt;d&lt;/sup&gt; + low-cost eFI&lt;sup&gt;e&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>85.3</td>
<td>85.0</td>
<td>82.3</td>
<td>87.7</td>
<td>85.2</td>
<td>83.6</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>84.0</td>
<td>86.7</td>
<td>81.3</td>
<td>88.7</td>
<td>85.1</td>
<td>83.9</td>
</tr>
<tr>
<td>Decision tree</td>
<td>69.3</td>
<td>61.7</td>
<td>61.7</td>
<td>69.3</td>
<td>65.9</td>
<td>61.7</td>
</tr>
<tr>
<td>Random forest</td>
<td>84.0</td>
<td>66.7</td>
<td>76.9</td>
<td>75.9</td>
<td>76.3</td>
<td>71.4</td>
</tr>
<tr>
<td><strong>Low-cost eFI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>85.3</td>
<td>81.7</td>
<td>81.7</td>
<td>85.3</td>
<td>83.7</td>
<td>81.7</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>86.7</td>
<td>81.7</td>
<td>83.1</td>
<td>85.5</td>
<td>84.4</td>
<td>82.4</td>
</tr>
<tr>
<td>Decision tree</td>
<td>76.0</td>
<td>58.3</td>
<td>66.0</td>
<td>69.5</td>
<td>68.1</td>
<td>61.9</td>
</tr>
<tr>
<td>Random forest</td>
<td>86.7</td>
<td>65.0</td>
<td>79.6</td>
<td>75.6</td>
<td>77.0</td>
<td>71.6</td>
</tr>
<tr>
<td><strong>No low-cost features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>80.0</td>
<td>83.3</td>
<td>76.9</td>
<td>85.7</td>
<td>81.5</td>
<td>80.0</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>73.3</td>
<td>75.0</td>
<td>69.2</td>
<td>78.6</td>
<td>74.1</td>
<td>72.0</td>
</tr>
<tr>
<td>Decision tree</td>
<td>78.7</td>
<td>55.0</td>
<td>67.3</td>
<td>68.6</td>
<td>68.1</td>
<td>60.6</td>
</tr>
<tr>
<td>Random forest</td>
<td>77.3</td>
<td>76.6</td>
<td>73.0</td>
<td>80.6</td>
<td>77.0</td>
<td>74.8</td>
</tr>
<tr>
<td><strong>Genetically selected low-cost features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Logistic regression</td>
<td>81.3</td>
<td>80.0</td>
<td>77.4</td>
<td>83.5</td>
<td>80.7</td>
<td>78.7</td>
</tr>
<tr>
<td>Support vector machine</td>
<td>80.0</td>
<td>76.7</td>
<td>75.4</td>
<td>81.1</td>
<td>78.5</td>
<td>76.0</td>
</tr>
<tr>
<td>Decision tree</td>
<td>69.3</td>
<td>61.7</td>
<td>61.7</td>
<td>69.3</td>
<td>65.9</td>
<td>61.7</td>
</tr>
<tr>
<td>Random forest</td>
<td>84.0</td>
<td>66.7</td>
<td>76.9</td>
<td>75.9</td>
<td>76.3</td>
<td>71.4</td>
</tr>
</tbody>
</table>

<sup>a</sup>PPA: positive percent agreement.  
<sup>b</sup>NPA: negative percent agreement.  
<sup>c</sup>F1: F-score.  
<sup>d</sup>ACFI: Aged Care Funding Instrument.  
<sup>e</sup>eFI: electronic frailty index.

Table 4. Performance of models based on all features.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPA&lt;sup&gt;a&lt;/sup&gt;</th>
<th>NPA&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Accuracy</th>
<th>F1&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR&lt;sup&gt;d&lt;/sup&gt;</td>
<td>97.3</td>
<td>96.7</td>
<td>96.7</td>
<td>97.3</td>
<td>97.0</td>
<td>96.7</td>
</tr>
<tr>
<td>SVM&lt;sup&gt;e&lt;/sup&gt;</td>
<td>86.7</td>
<td>95.0</td>
<td>85.1</td>
<td>95.6</td>
<td>90.4</td>
<td>89.8</td>
</tr>
<tr>
<td>DT&lt;sup&gt;f&lt;/sup&gt;</td>
<td>76.0</td>
<td>63.3</td>
<td>67.9</td>
<td>72.1</td>
<td>70.4</td>
<td>65.5</td>
</tr>
<tr>
<td>RF&lt;sup&gt;g&lt;/sup&gt;</td>
<td>88.0</td>
<td>75.0</td>
<td>83.3</td>
<td>81.5</td>
<td>82.2</td>
<td>78.9</td>
</tr>
</tbody>
</table>

<sup>a</sup>PPA: positive percent agreement.  
<sup>b</sup>NPA: negative percent agreement.  
<sup>c</sup>F1: F-score.  
<sup>d</sup>LR: logistic regression.  
<sup>e</sup>SVM: support vector machine.  
<sup>f</sup>DT: decision tree.  
<sup>g</sup>RF: random forest.
Table 5. Performance of models based only on low-cost features.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>PPA(^a)</th>
<th>NPA(^b)</th>
<th>Accuracy</th>
<th>FI(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR(^d)</td>
<td>77.3</td>
<td>63.3</td>
<td>69.1</td>
<td>72.5</td>
<td>71.1</td>
<td>66.1</td>
</tr>
<tr>
<td>SVM(^e)</td>
<td>77.3</td>
<td>58.3</td>
<td>67.3</td>
<td>69.9</td>
<td>68.9</td>
<td>62.5</td>
</tr>
<tr>
<td>DT(^f)</td>
<td>61.3</td>
<td>70.0</td>
<td>59.2</td>
<td>71.9</td>
<td>65.2</td>
<td>64.1</td>
</tr>
<tr>
<td>RF(^g)</td>
<td>77.3</td>
<td>58.3</td>
<td>67.3</td>
<td>69.9</td>
<td>68.9</td>
<td>62.5</td>
</tr>
</tbody>
</table>

\(^a\)PPA: positive percent agreement.  
\(^b\)NPA: negative percent agreement.  
\(^c\)FI: F-score.  
\(^d\)LR: logistic regression.  
\(^e\)SVM: support vector machine.  
\(^f\)DT: decision tree.  
\(^g\)RF: random forest.

**Discussion**

**Principal Findings**

With AI techniques, cost-effective screening tests for frailty are possible for aged care databases that contain an ACFI assessment and unstructured patient notes. This study has shown that the ACFI assessment alone does not provide sufficient information to determine if a patient is frail. However, when ACFI data are augmented by as few as 10 additional features, an AI model can be derived that performs well enough to be used as a screening test. What this means in clinical practice is that older people with frailty can be rapidly and accurately identified in residential care using our novel AI-derived model for frailty. A rapid identification of frailty is crucial to optimally manage the condition [27]. Indeed, the recent Australian Royal Commission to Aged Care highlighted the importance of early identification of aged care residents with frailty, who require additional support [28].

The value of any AI-derived model for frailty screening can be judged by the amount it reduces the cost of acquisition of the features required to determine the value of the deficits used to construct a frailty index. Features that are routinely collected and stored in a database in a format that can be directly fed into a classification model have a low cost of acquisition. Unfortunately, as shown in this study (Table 5) and others [20], such models lack both the sensitivity and specificity to be useful screening tests. At the other extreme, models that include all the deficit features used to calculate the eFI perform extremely well [20] (Table 4), but the value of such models is marginal.

To be useful for a screening test, a model must be acceptably accurate and significantly reduce the cost of acquisition of the features required to implement a frailty index. If a model cannot be developed with acceptable accuracy without including at least some high-cost features, it is desirable to determine the optimal minimum set of high-cost features required to achieve an acceptable performance. Genetic algorithms perform well at determining the optimal subset of features required to maximize the performance of a model. Furthermore, their choice of a subset can be limited to any number of features, up to and including all the available features. This allows the trade-off between the number of features and the performance of the derived models to be determined.

This study found that if a genetic algorithm was permitted to choose any number of features from all the available features, regardless of their cost, it most frequently chose subsets that only included high-cost features. This motivated the development of the previously mentioned partial genetic algorithm, which forced the algorithm to include low-cost features as well. However, this raises the question of whether the low-cost features add any value at all. To answer this question, the results include both a fixed set that had no low-cost features and a set that included only the low-cost features used to calculate the eFI. Considering logistic regression models with 10 high-cost features, including all the low-cost features, yielded an improvement of 17% in sensitivity (89% versus 72%). This combination did not compromise specificity, which remained stable (87%) and is comparable to the scenario with no low-cost features. This improvement is significant and possibly represents the difference between a clinically useful screening test and one that is inadequate. Even if the comparison is made between models built on all the low-cost features and those that include only low-cost features used in the eFI calculation, there is a 6% improvement in sensitivity (89% versus 83%) and 5% in specificity (87% versus 82%).

Although the partial genetic algorithm–built models with 10 high-cost features use less than a third of all the high-cost features, they still require those 10 features to be extracted by screening patient notes. Recent advances in natural language processing (NLP) show promise for automating this extraction process. It is plausible that NLP could extract all the features required to calculate the eFI, but this would require a much larger data set than the one used in this study. In the meantime, the cost of acquisition of at least 10 features from every patient record remains the cost of implementing a screening test on any database similar to ours that contains an ACFI assessment and unstructured patient notes.

Partial genetic algorithms can be used to derive classification models from any database where the cost of acquisition of some parameters is higher than it is for others. Although they have been demonstrated in this study on an aged care database to
predict frailty, they could be used in any domain. They are well suited to permit AI models to be trained to implement screening tests in domains where costs are important and there is a difference in the cost of acquisition of candidate features.

Limitations
Because this study reuses the data from a previous study [20], it shares the limitations associated with the data from the first study. In particular, the data were sourced from a single aged care provider, and the data set was relatively small. This study further filtered patients based on the availability of an ACFI assessment. It is plausible that these criteria gave a skewed representation of the population that a screening test would be applied to, resulting in different model performance. The ability to reproduce AI results continues to be controversial [29,30] within medicine, so further studies should aim to reproduce these results with different data sets. A further limitation is the changing model of aged care in Australia, with a new model set to replace ACFI in the next 2 years.

Conclusion
The value of screening tests lies in their cost-effective application. The main cost of applying a model-based screening test lies in acquiring the measures fed into the model. To derive useful screening tests using AI techniques, algorithms must be employed that favor the use of cheaper features over those that require more effort or patient risk to acquire. What all aged care providers and their clinical advisers need is a screening tool that will allow the efficient planning of evidence-based interventions to older frail people who will best benefit from them. At a time where the aged care sector and all providers are being asked by governments and national quality agencies to focus on this vulnerable group, it is crucial that we employ an efficient screening tool.

This paper has shown how partial genetic algorithms can be used to determine an optimal subset of high-cost features to use with cheap features to derive AI models to classify frailty, both in terms of which parameters to use and how many to use. This technique can be applied to any database. It does not guarantee that an adequate model will be found from any database, but it does give a good indication of whether there is sufficient information in the data to derive a model.

Partial genetic algorithms were demonstrated in this paper to derive a cost-effective screening test for frailty, but the method can be applied to any screening tests where there is a disparity in the cost of measuring the required features. The outcome of this study will aid health care providers in screening for frailty with better accuracy through the proposed cost-effective method, which strikes a good balance between accuracy and cost.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Full list of features.
[DOCX File, 14 KB - aging_v5i4e38464_app1.docx ]

Multimedia Appendix 2
Selected features.
[DOCX File, 41 KB - aging_v5i4e38464_app2.docx ]

Multimedia Appendix 3
Low-cost features selected for models built with GA-selected subset.
[DOCX File, 13 KB - aging_v5i4e38464_app3.docx ]

References


Abbreviations

ACFI: Aged Care Funding Instrument
AI: artificial intelligence
eFI: electronic frailty index
FI: frailty index
FP: frailty phenotype
GA: genetic algorithm
NLP: natural language processing
PAS: Psychogeriatric Assessment Scales
SVM: support vector machine

Edited by J Wang, T Leung; submitted 03.04.22; peer-reviewed by Y Longobucco, M Kraus; comments to author 21.06.22; revised version received 01.07.22; accepted 30.07.22; published 07.10.22.

Please cite as:
Oates J, Shafiabady N, Ambagtsheer R, Beilby J, Seiboth C, Dent E
Evolving Hybrid Partial Genetic Algorithm Classification Model for Cost-effective Frailty Screening: Investigative Study
JMIR Aging 2022;5(4):e38464
URL: https://aging.jmir.org/2022/4/e38464
doi:10.2196/38464
PMID:36206042

©John Oates, Niusha Shafiabady, Rachel Ambagtsheer, Justin Beilby, Chris Seiboth, Elsa Dent. Originally published in JMIR Aging (https://aging.jmir.org), 07.10.2022. 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.