frontiers | Research Topics

Aging-related factors in digital health: Design, uptake, engagement, and outcomes

Edited by
OraLee Branch, Sarah Anne Graham, Raeanne Moore and Patricia A. Arean

Published in
Frontiers in Digital Health

ISSN 1664-8714
ISBN 978-2-83251-408-5
DOI 10.3389/978-2-83251-408-5

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Aging-related factors in digital health: Design, uptake, engagement, and outcomes

Topic editors

OraLee Branch — Lark Health (United States), United States
Sarah Anne Graham — Lark Health, United States
Raeanne Moore — University of California, San Diego, United States
Patricia A. Arean — University of Washington, United States

Citation

Branch, O., Graham, S. A., Moore, R., Arean, P. A., eds. (2023). *Aging-related factors in digital health: Design, uptake, engagement, and outcomes*.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-408-5

Table of contents

frontiers | Frontiers in Digital Health

TYPE Editorial
PUBLISHED 09 January 2023
DOI 10.3389/fdgth.2022.1124464

OPEN ACCESS

EDITED AND REVIEWED BY
Stephen Schueller,
University of California, Irvine, United States

*CORRESPONDENCE
OraLee H. Branch
oralee.branch@lark.com

SPECIALTY SECTION
This article was submitted to Human Factors and Digital Health, a section of the journal Frontiers in Digital Health

RECEIVED 15 December 2022
ACCEPTED 19 December 2022
PUBLISHED 09 January 2023

CITATION
Branch OH, Graham SA, Moore RC and Areán PA (2023) Editorial: Aging-related factors in digital health: Design, uptake, engagement, and outcomes.
Front. Digit. Health 4:1124464.
doi: 10.3389/fdgth.2022.1124464

Editorial: Aging-related factors in digital health: Design, uptake, engagement, and outcomes

OraLee H. Branch[1]*, Sarah A. Graham[1], Raeanne C. Moore[2] and Patricia A. Areán[3]

[1]Department of Clinical Research, Lark Health, Mountain View, CA, United States, [2]Department of Psychiatry, School of Medicine, University of California, San Diego, La Jolla, CA, United States, [3]Department of Psychiatry and Behavioral Sciences, School of Medicine, University of Washington, Seattle, WA, United States

KEYWORDS

mhealth, patient activation, preventive health, mental health, usability, acceptability, adherence

Editorial on the Research Topic
Aging-related factors in digital health: design, uptake, engagement, and outcomes

This Research Topic explores digital health technologies designed to facilitate chronic disease prevention and management for older adults. Adults over age 65 have a high prevalence of chronic diseases, with 86% having at least one chronic condition such as diabetes or hypertension (1) and ~12% reporting subjective cognitive decline (2). These conditions are costly: in recent estimates, despite accounting for only 15% of the US population, older adults accounted for 34% of total healthcare expenditures (3). As society ages (4) and provider availability and time dwindle (5), enhancing older adults' ability to effectively self-manage their health is paramount.

For patients to successfully manage their health, they must exhibit "patient activation"—actively demonstrating the willingness and ability to take independent actions to manage their health and care (6). Digital health technologies offer a scalable means to facilitate patient activation by increasing patient knowledge, equipping patients with self-management techniques, and offering feedback and support to improve patient confidence.

The articles in this research topic collection address three central themes: (1) Older adults are willing and able to use digital health technologies when provided the opportunity, (2) They exhibit excellent adherence when using these technologies, and (3) Their use of these technologies generates unique, actionable health information and results in positive health outcomes.

Despite the potential of digital health technologies, skepticism exists regarding their utility for older adults. A considerable body of research has focused on the "Digital Divide" as it pertains to age-related willingness and ability to use digital health

technologies (7). However, older adults' use of technology continues to increase, and the notion that older adults are unwilling or incapable of using technology is outdated.

As of 2021, 61% of older adult Americans owned a smartphone, representing a considerable increase from 27% in 2015 (8). The gaps between the oldest and youngest adults have narrowed in areas such as smartphone and tablet computer ownership and social media and internet use, with 75% of older adults now reporting internet use (8). These trends are supported by papers in this special issue by Graham et al. and Auster-Gussman et al. Older adults do engage with and have success when using technologies that require smartphone and internet use when provided with opportunities from their healthcare provider to engage in digital preventive health programs. In fact, engagement among older adults exceeded that of younger adults (Graham et al.), and older adults experienced beneficial outcomes such as weight loss while using these programs (Auster-Gussman et al.). Although contrary to outdated assumptions that older adults struggle to use technology, the high engagement and positive outcomes are consistent with the data described throughout this Research Topic Collection. Older adults may represent a group that takes full advantage of opportunities to better self-manage their health when presented with digital technologies.

Digital technologies can provide greater access to personalized health information and generate key insights that can improve care quality and activate patients to achieve better health outcomes. Quality care and improved outcomes benefit all ages, but older adults stand to benefit from digital health technologies to the greatest extent given the complexity of managing multiple health conditions simultaneously, the need for frequent and real-time care, the need to understand the dynamics of older adult behaviors in a home environment, and challenges to adhering to care management recommendations introduced by the normal aging process.

In this collection, Paolillo et al. demonstrate that older adults are willing and able to adhere to a wearable device (Fitbit) protocol that enables accurate monitoring of physical activity behaviors. Accuracy in measuring physical activity behaviors is critical in evaluating the effectiveness of lifestyle interventions, and as VandeBuente et al. demonstrate, the accuracy of subjective physical activity reporting decreases with declining memory and executive performance–common consequences of aging. Research demonstrates that wearable devices provide more valid estimates of physical activity behaviors than subjective reporting, providing actionable information to both patients and providers.

To maximize their potential, designers of digital technologies must understand the fundamentals of aging and incorporate age-related considerations (e.g., motor, perceptual, and cognitive capabilities) into the design of their products. Research must continue to focus on older adults' adoption of these technologies, engagement, and perceptions of usefulness.

In this collection, Badal et al., Klaus et al., and Moore et al. demonstrate that surveys deployed *via* smartphone technology such as ecological momentary assessment (EMA) and EMA along with mobile cognitive testing are feasible and acceptable to older adults and present unique opportunities for assessment of health. Use of smartphone technology is a valid way to collect real-time behaviors, experiences, and emotions without reliance on retrospective recall or influence of current mood state. The ability to serially collect data in the individual's environment revealed relationships among negative affect, loneliness, and adaptive behaviors (Badal et al.). EMA resulted in high adherence (Klaus et al.), even when older adults had potential barriers to engagement such as mild cognitive impairment (Moore et al.). Reaching the older adult population during the COVID-19 pandemic to facilitate the maintenance of health and well-being was of critical importance, and EMA offered a remote and scalable means to do so.

Digital health technologies enable data collection and patient activation in the natural environment. In addition to 24/7 tracking of daily behaviors and experiences, digital technologies can also support essential round-the-clock aspects of care management such as medication adherence. In this collection, Gualtiere et al. describe how medication storage location impacts adherence and discuss how technology can play an important role in promoting adherence, which can also be impacted by things like worsening memory with age.

When deploying digital health technologies, consideration must be given to older adults' perceptions and attitudes toward these technologies. In this collection, Woerner et al. observed that although digital technologies for mental health are considered valuable by all age groups, older adults prefer that they play a complementary or supportive role rather than primary role in their care. Continued research and collaborations between industry, where products are designed, built, and maintained, and academic institutions, where research designs can be optimized, should be encouraged to best reveal how digital health technologies should be deployed and leveraged in practice for older adults.

Digital health technologies represent a scalable and cost-effective opportunity to activate older adults in self-managing their health and care. As demonstrated in this special issue of Frontiers in Digital Health, older adults are willing to use digital health technologies when provided the opportunity, they exhibit excellent adherence when using these technologies, and their use of these technologies generates actionable health information and positive results. Best practices for optimizing the older adult user experience and the implementation of these technologies within the healthcare environment should remain a focus of future

research, including patterns of digital health use across racially and ethnically diverse older populations.

Author contributions

OLB, SAG, RCM, and PAA contributed to this editorial. All authors contributed to the article and approved the submitted version.

Conflict 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.

Publisher's note

References

1. Centers for Disease Control Prevention. Percent of U.S. Adults 55 and Over with Chronic Conditions. National Center for Health Q16 Statistics. (2015) (Accessed December 10, 2022).

2. Centers for Disease Control Prevention. Subjective Cognitive Decline — A Public Health Issue. Division of Population Health, National Center for Chronic Disease Prevention and Health Promotion. (2019) (Accessed December 10, 2022).

3. Mitchell EM. *Concentration of health expenditures in the US civilian noninstitutionalized population, 2014*. Rockville, MD: Agency for Healthcare Research and Quality (US) (2017).

4. Administration for Community Living. 2020 PROFILE OF OLDER AMERICANS. (2021). Retrieved from: https://acl.gov/sites/default/files/aging%20and%20Disability%20In%20America/2020Profileolderamericans.final_.pdf (Accessed December 10, 2022).

5. Zhang X, Lin D, Pforsich H, Lin VW. Physician workforce in the United States of America: forecasting nationwide shortages. *Hum Resour Health*. (2020) 18(1):1–9. doi: 10.1186/s12960-020-0448-3

6. Greene J, Hibbard JH. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. *J Gen Intern Med*. (2012) 27(5):520–6. doi: 10.1007/s11606-011-1931-2

7. McDonough CC. The effect of ageism on the digital divide among older adults. *J Gerontol Geriatr Med* (2016) 2:008. doi: 10.24966/GGM-8662/100008.

8. Pew Research Center. Share of those 65 and older who are tech users has grown in the past decade. (2022). Available from: https://www.pewresearch.org/fact-tank/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/ (Accessed December 10, 2022).

frontiers in Digital Health

ORIGINAL RESEARCH
published: 06 August 2021
doi: 10.3389/fdgth.2021.642818

Older Adults Engage With Personalized Digital Coaching Programs at Rates That Exceed Those of Younger Adults

*Sarah A. Graham, Natalie Stein, Fjori Shemaj, OraLee H. Branch, Jason Paruthi and Stephen Chad Kanick**

Lark Health, Mountain View, CA, United States

Background: The US population is aging and has an expanding set of healthcare needs for the prevention and management of chronic conditions. Older adults contribute disproportionately to US healthcare costs, accounting for 34% of total healthcare expenditures in 2014 but only 15% of the population. Fully automated, digital health programs offer a scalable and cost-effective option to help manage chronic conditions. However, the literature on technology use suggests that older adults face barriers to the use of digital technologies that could limit their engagement with digital health programs. The objective of this study was to characterize the engagement of adults 65 years and older with a fully automated digital health platform called Lark Health and compare their engagement to that of adults aged 35–64 years.

Methods: We analyzed data from 2,169 Lark platform users across four different coaching programs (diabetes prevention, diabetes care, hypertension care, and prevention) over a 12-month period. We characterized user engagement as participation in digital coaching conversations, meals logged, and device measurements. We compared engagement metrics between older and younger adults using nonparametric bivariate analyses.

Main Results: Aggregate engagement across all users during the 12-month period included 1,623,178 coaching conversations, 588,436 meals logged, and 203,693 device measurements. We found that older adults were significantly more engaged with the digital platform than younger adults, evidenced by older adults participating in a larger median number of coaching conversations (514 vs. 428) and logging more meals (174 vs. 89) and device measurements (39 vs. 28) all $p \leq 0.01$.

Conclusions: Older adult users of a commercially available, fully digital health platform exhibited greater engagement than younger adults. These findings suggest that despite potential barriers, older adults readily adopted digital health technologies. Fully digital health programs may present a widely scalable and cost-effective alternative to traditional telehealth models that still require costly touchpoints with human care providers.

Keywords: telemedicine, mobile health, engagement, chronic disease management, geriatric population, preventative care

OPEN ACCESS

Edited by:
Constantinos S. Pattichis, University of Cyprus, Cyprus

Reviewed by:
Palash Chandra Banik, Bangladesh University of Health Sciences, Bangladesh
Timo Jämsä, University of Oulu, Finland

****Correspondence:***
Stephen Chad Kanick
chad.kanick@lark.com

Specialty section:
This article was submitted to Connected Health, a section of the journal Frontiers in Digital Health

Received: *16 December 2020*
Accepted: *08 July 2021*
Published: *06 August 2021*

Citation:
Graham SA, Stein N, Shemaj F, Branch OH, Paruthi J and Kanick SC (2021) Older Adults Engage With Personalized Digital Coaching Programs at Rates That Exceed Those of Younger Adults. Front. Digit. Health 3:642818. doi: 10.3389/fdgth.2021.642818

HIGHLIGHTS

- Personalized digital health programs can meet the growing needs of a rapidly expanding population of older adults.
- Older adults over 65 years showed greater engagement than adults aged 35–64 years with a fully automated health coaching platform.
- Engagement of older adults in a fully digital health platform highlights the potential for widespread adoption, and this supports continued research to optimize digital health interventions for older adult users.

INTRODUCTION

Digital health has grown considerably in recent years, with revenue increasing from $4.4 billion in 2016 to $6 billion in 2017 and an estimated 200 new health apps being released per day (1). Since the onset of the COVID-19 pandemic in 2019, telehealth and digital health utilization rates have further increased (2, 3) with high patient satisfaction (4). The growth of digital health coincides with the US population aging, with adults 65 years and older comprising 15% of the population in 2014 and projected to grow to 21% in 2030 (5). Despite accounting for only 15% of the population, older adults accounted for 34% of total healthcare expenditures (6). Digital health innovations offer an affordable and scalable mechanism to address older adults' unique needs, helping them better manage their health and retain their autonomy (7, 8). However, to be effective for these purposes, older adults must engage with digital health technologies. A variety of digital health offerings are covered by Medicare (9, 10), but the technologies used to enable such programs may present unique barriers to older adults such as prior experience, attitudes, usability, trust, and physical and cognitive abilities (11, 12).

Older adults have a high prevalence of chronic diseases; 86% have at least one chronic condition such as diabetes or hypertension, 56% have two, and 23% have at least three (13). Self-management of these chronic conditions is essential to minimize healthcare spending. In 2016 alone, $730 billion was attributable to modifiable risk factors including high body mass index (BMI), blood pressure, and fasting glucose, and the largest fraction was for those aged 65 and older (14). Digital health programs may help older adults manage modifiable risk factors through interventions that engender positive behavior changes in physical activity, weight management, nutrition, medication adherence, and monitoring of clinical indicators like blood pressure and glucose (15, 16). Digital health programs may increase access to primary and specialty care, especially in remote or underserved areas or in populations with challenges like poor mobility (17). By increasing access to care and improving patient health, digital health programs may lessen the burden on healthcare systems (18).

Adoption and use of digital technologies by older adults are important topics, as these individuals tend to be slower than younger adults to adopt new technologies (19). However, older adults are rapidly integrating technology into their lives and are more likely to use technology when they perceive a benefit (20). Questions around older adult engagement with digital health technologies are important, as greater engagement has been associated with improved health outcomes (21, 22). Additionally, elucidating interactions between users and digital platforms helps to tailor these platforms to user preferences, which is associated with increased engagement (23). Though digital health appears promising for older adults, there are little data characterizing their engagement with digital health platforms. This knowledge gap hinders the potential for digital health programs to better serve the needs of older adults. Investigations of older adults' engagement with fully digital health platforms are necessary to help pave the way for the field and inform future studies and interventions.

The purpose of the present study was to characterize the engagement of adults 65 years and older with a mobile digital health platform called Lark Health and to compare their engagement to that of adults aged 35–64 years. We chose the comparison group of 35–64 years based upon this group having adopted digital technologies later in life, rather than having grown up with such technologies (24). The Lark digital health platform delivers personalized health coaching to promote wellness or to prevent, delay, or manage chronic diseases through promoting positive behavior changes. Lark programs are delivered *via* artificial intelligence (AI) with a responsive coaching interface on a smartphone. We analyzed data from users enrolled across four digital health programs to determine whether engagement, defined as participation in coaching conversations, meal logging, and device measurements, varied by age. We hypothesized that older adults would have less engagement in the digital platform than younger adults due to barriers to technology use common to this age group, which would be reflected by lower participation in coaching conversations and fewer meals logged and device measurements.

METHODS

Study Design

This was a longitudinal, observational study of participants who were users of the Lark Health disease management and prevention programs. We considered measures of engagement from the program start until 12 months later. The study received exemption status from Advarra (Protocol #Pro00047181) Institutional Review Board (IRB) for retrospective analyses of previously collected and de-identified data.

Participants and Recruitment

Participants in this study were individuals who qualified for any of four digital health programs (see **Figure 1**) offered through existing partnerships between Lark Health and health insurance companies, employers, or other organizations, and who owned an Android-enabled smartphone or iPhone.

Lark recruits eligible users *via* direct referrals from health plans and/or healthcare providers, digital awareness campaigns (e.g., Facebook ads), and a large managed services organization. The Lark programs are a covered service under the insurance plans of these users. Eligibility differs for each clinical program

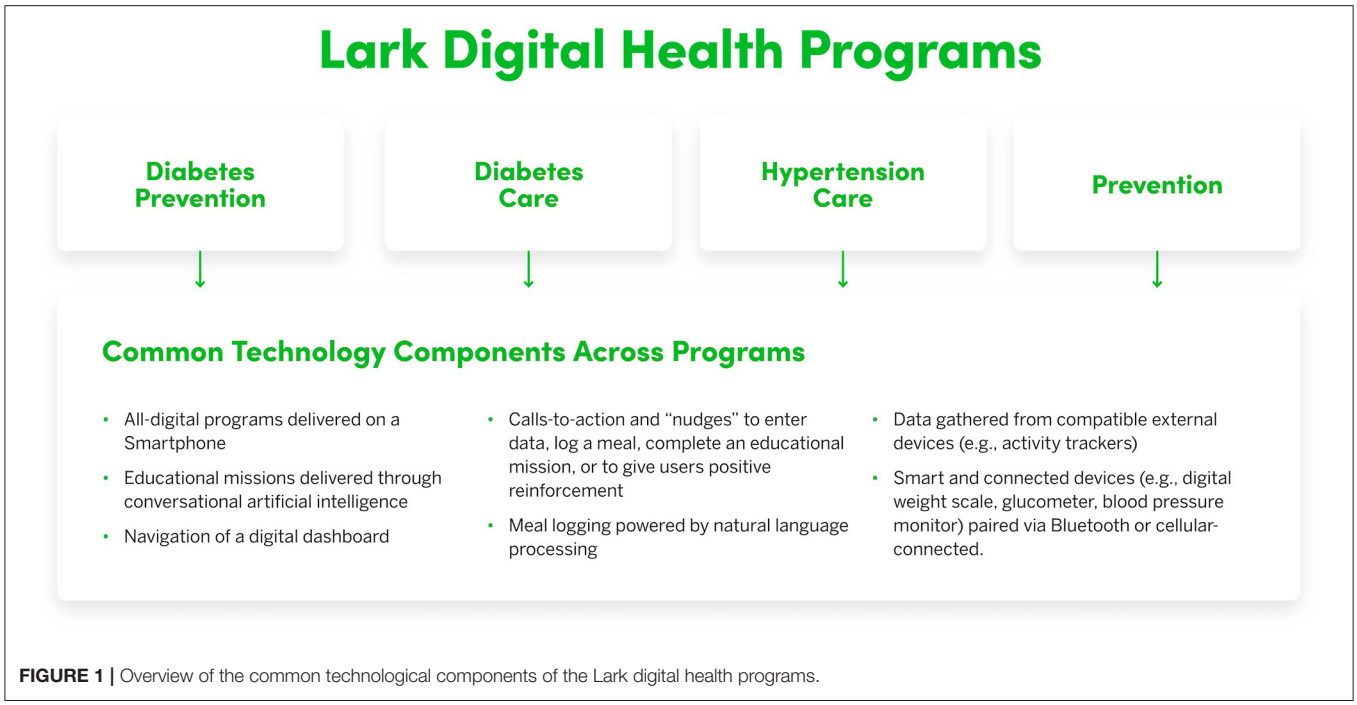


FIGURE 1 | Overview of the common technological components of the Lark digital health programs.

based on the program focus [e.g., Diabetes Prevention Program users must meet risk criteria established by the Centers for Disease Control and Prevention (25), Diabetes Program users must have a diagnosis of diabetes, and Hypertension Program users must have a diagnosis of hypertension]. A user's health plan confirms their eligibility for a particular program prior to enrollment. There are no specific eligibility requirements for the General Wellness Program, and these users are eligible to enroll if their insurance covers preventive wellness programs. Participants who opted in a program received a link *via* text message to download the Lark program to their smartphones. Some participants received a connected device (e.g., digital weight scale) as a part of their specific program.

Participant inclusion criteria were: (1) enrollment in a Lark Health program between January 1st, 2019 and July 28th, 2019; (2) aged 35 years and older; (3) those who received a connected device as a part of their program; and (4) those who completed at least one educational mission (i.e., an educational lesson that included a series of automated check-ins with the digital coach and coaching conversations around a topic related to the program focus). We selected the date-range criteria to reduce time-dependent variations in the content of coaching and types of participant-coach interactions offered within each program. We further focused our analyses by excluding young adults 18–34 years who would require separate considerations due to lifetime technology exposure, users who did not have a connected device since they were not participating in a full version of a program, and users who did not complete any educational missions since they did not demonstrate a minimum level of intent to participate in their program.

Digital Health Platform

The Lark digital health platform provides automated and personalized coaching using conversational AI. Each program has weekly "educational missions" consisting of daily check-ins and educational material around a weekly topic related to the focus of each program. The Lark programs differ in clinical focus and content (e.g., diabetes prevention, hypertension management) and employ a standard set of engagement methods that include automated coaching conversations, meal logging, and device measurements. Every Lark program includes AI coaching on lifestyle choices such as healthy eating, physical activity, sleep, and stress management. The AI coach employs elements of cognitive behavioral therapy to encourage users to adopt healthy behaviors and build self-management skills and knowledge to sustain these behaviors. Users receive regular "calls-to-action" and "nudges" that either encourage them to engage with the Lark platform through actions like having a coaching conversation or offer them positive reinforcement (e.g., great job on your walk today). Users have the option to set a weight-loss goal and receive personalized coaching. Lark responds immediately with personalized feedback when users log data such as weight or meals, or when they indicate they want to have a conversation. Lark also provides daily and weekly summaries of progress. The intuitive meal-logging system uses natural language processing to provide personalized coaching regarding meal content and quality. Lark can also gather data from external devices like activity trackers that are connected to Google Fit or iOS Health Kit. The AI coach is available for unlimited use 24 h a day if users want to check in to discuss challenges or progress. The main technological aspects of these programs are summarized in **Figure 1**.

Measures of Engagement With the Digital Health Platform

We defined three metrics to quantify the engagement of users with the digital health platform. *Coaching conversations* included interactions between the AI coach and user and included educational missions. *Meals logged* included data provided by users regarding food intake. *Device measurements* included measurements obtained from the smart and connected devices (i.e., digital weight scale, glucometer, or blood pressure monitor).

We considered the total number of coaching conversations, meals logged, and device measurements experienced over the first 12 months after the program start date. We did not separately assess engagement metrics per program due to uneven sample sizes, but we did separately analyze two program-specific groupings, (1) clinically oriented (diabetes prevention, diabetes care, and hypertension care) and (2) wellness (prevention).

Statistical Analyses

We conducted all statistical tests in Python version 3.7.3. We checked distributions for each variable and compared age groups on continuous measures with the Mann-Whitney *U*-test (*U*-statistic) using the "mannwhitneyu" function from the scipy.stats module in python due to non-normal data, and Chi-Square tests (χ^2 statistic) using the "chi2_contingency" function from scipy.stats for categorical data. Users self-reported their age, gender, weight, and height upon enrollment in the Lark digital platform. We calculated body mass index (BMI; kg/m^2) from height and weight. We reported medians with interquartile (IQ) ranges for demographics and user characteristics for: (1) all users, (2) users 65 years and older, and (3) users 35 to 64 years, as well as the distribution of users across program types (**Table 1**).

We compared age groups (i.e., older adults vs. younger adults) on engagement metrics (i.e., number of coaching conversations, meals logged, and device measurements). We also compared these engagement metrics between age groups for two program-specific groupings (i.e., clinically oriented vs. wellness) and within each age group between these program groupings. We reported both medians with IQ ranges and means with 95% CIs for all engagement metrics in **Tables 2**, **3**. We used an alpha $\leq$0.05 to evaluate significance for all tests.

RESULTS

Per our inclusion/exclusion criteria, the final sample size included in the analyses was 2,169 users. Older adults aged 65+ years comprised 14% of the sample, and the remaining 86% consisted of adults aged 35–64 years. Per design, we had complete separation between age groups (**Table 1**). Older users were more likely to be male than younger users (37 vs. 32%; $p < 0.01$) and had a lower body weight (87 kg vs. 93 kg; $p < 0.01$) and BMI (32 vs. 33 kg/m^2; $p < 0.01$) at program enrollment (**Table 1**). We did not observe a difference in the distribution of users across programs between the two age groups ($p = 0.49$).

Aggregate engagement across all users during the 12-month period included 1,623,178 coaching interactions, 588,436 meals logged, and 203,693 device measurements. We observed that older adults engaged with the Lark digital health platform to a greater degree than younger adults, evidenced by a significantly larger median number of coaching conversations ($U = 233{,}794$; $p \leq 0.01$), meals logged ($U = 212{,}673$; $p \leq 0.01$), and device measurements ($U = 238{,}056$; $p \leq 0.01$) across all programs (**Table 2**).

When we separately considered user engagement per thematic program grouping (i.e., clinically oriented vs. wellness), we again observed that older adults engaged with the Lark platform to a greater degree than younger adults. Compared to younger adults, older adults had a higher median number of coaching conversations ($U = 144{,}761$; $p \leq 0.01$), meals logged ($U = 138{,}824$; $p \leq 0.01$), and device measurements ($U = 140{,}316$; $p \leq 0.01$) for clinically oriented programs and a higher median number of coaching conversations ($U = 10{,}114$; $p \leq 0.01$), meals logged ($U = 7{,}909$; $p \leq 0.01$), and device measurements ($U = 12{,}590$; $p \leq 0.01$) for the wellness program (**Table 3**).

We further found that within each age group, older adults enrolled in the wellness program had a higher median number of coaching conversations ($U = 309{,}663$; $p = 0.03$) and meals logged ($U = 258{,}785$; $p \leq 0.01$) compared with older adults enrolled in the clinically oriented programs (**Table 3**). Older adults did not differ in device measurements ($p = 0.15$) between program-specific groupings. In contrast, younger adults enrolled in the wellness program had a lower median number of coaching conversations ($U = 6{,}599$; $p = 0.02$) and meal logging ($U = 6{,}109$; $p \leq 0.01$) compared to younger adults enrolled in the clinically oriented programs (**Table 3**). Younger adults also did not differ in device measurements ($p = 0.25$) between program-specific groupings.

DISCUSSION

The present study characterized the engagement of older adults aged 65 and older with a digital health platform compared to adults aged 35–64 years. Users of the Lark digital health platform engaged with multiple modes of technology over a 12-month period, including navigating a mobile application on a smartphone, engaging in conversational AI with a digital coach, receiving and responding to prompts to interact with the platform, logging meals, and monitoring progress *via* measurements of weight, glucose, and blood pressure collected *via* smart and connected devices. Contrary to our main hypothesis, we observed that older adults engaged more with these technologies than younger adults, evidenced by engagement in a larger number of coaching conversations and more meals logged and device measurements.

Older Adults Engaged With Fully Digital Health Programs

The higher engagement observed in older adults in this study is promising. Although the literature on the use of digital health technologies among older adults is sparse, some evidence has suggested they have lower levels of engagement than younger adults due to barriers to use (11, 12). For example, older adults experience declining physical and cognitive functioning (26),

TABLE 1 | Participant demographics and characteristics and distribution of users across programs.

		Full sample ($N = 2,169$)	35–64 years ($n = 1,868$)	65+ years ($n = 301$)	
		Median [IQ range]	**Median [IQ range]**	**Median [IQ Range]**	***U*-Stat.; *p*-val**
Age	[years]	53 [45, 60]	51 [43, 57]	68 [66, 71]	0; $p < 0.01$
Weight	[kg]	92 [79, 108]	93 [79, 109]	87 [75, 101]	199,447; $p < 0.01$
Height	[cm]	168 [163, 175]	168 [163, 175]	168 [160, 175]	255,768; $p = 0.01$
BMI	[kg/m^2]	32 [29, 36]	32 [29, 37]	31 [28, 35]	174,003; $p < 0.01$
		***N* [%]**	***n* [%]**	***n* [%]**	**χ^2 Stat.; *p*-val**
Gender	F	1,448 [67]	1,264 [68]	184 [61]	11; $p < 0.01$
	M	708 [33]	596 [32]	112 [37]	
	N/A	13 [0]	8 [0]	5 [2]	
Race	White	1,570 [72]	1,326 [71]	257 [85]	11; $p = 0.001$
	Not White	599 [28]	542 [29]	44 [15]	
Ethnicity	Hispanic or Latino	210 [10]	192 [10]	13 [4]	27; $p \leq 0.0001$
	Not Hispanic or Latino	1,959 [90]	1,676 [90]	288 [96]	
		***N* [%]**	***n* [%]**	***n* [%]**	**χ^2 Stat.; *p*-val**
Programs	Diabetes prevention	1,396 [64]	1,201 [64]	195 [65]	3; $p = 0.49$
	Diabetes care	86 [4]	69 [4]	17 [6]	
	Hypertension care	151 [7]	130 [7]	21 [7]	
	Prevention	536 [25]	468 [25]	68 [22]	

Data presented for the full sample and stratified by age into older (65+ years) and younger (35–64 years) groups. Statistical comparisons are between age groups for demographics, characteristics, and distribution of users across programs.

TABLE 2 | Engagement metrics of users across all programs.

Engagement metrics	Values	Full sample ($N = 2,169$)	35–64 years ($n = 1,868$)	65± years ($n = 301$)
Number of coaching conversations	Median [IQ range]	437 [281, 615]	428 [276, 598]	514 [312, 720]**
	Mean [95% CI]	486 [474, 499]	474 [461, 488]	561 [524, 597]
Number of meals logged	Median [IQ range]	96 [39, 220]	89 [38, 201]	174 [54, 398]**
	Mean [95% CI]	176 [167, 186]	161 [151, 170]	273 [240, 306]
Number of device measurements	Median [IQ range]	30 [10, 70]	28 [10, 67]	39 [15, 101]**
	Mean [95% CI]	71 [55, 87]	69 [51, 88]	82 [69, 94]

Statistical significance based on comparison of the medians between age groups and denoted by $^{*}p \leq 0.05$; $^{**}p \leq 0.01$ next to older adult medians.
Data presented for the full sample and stratified by age into older (65+ years) and younger (35–64 years) groups. Median [Interquartile (IQ) Range] and Mean [95% Confidence Interval (CI)] provided for all comparisons.

which may directly affect their ability to visually navigate a digital screen, remember how to interact with digital programs, or understand technological prompts or notifications. However, we observed that older adults engaged in more coaching conversations, logged more meals, and recorded more device measurements than younger adults. These interactions suggest that despite potential barriers, adults over 65 years of age were able to engage with an all-digital, app-based coaching platform. If we consider our results in the context of commonly cited barriers to technology use of older adults, our findings indicate that older adult users were able to optically interpret text, use touch-based interactions, navigate the in-app menu, take measurements with smart and connected digital devices, pair connected Bluetooth devices with their mobile phones, and maintain battery charge to support device use.

Trust is another commonly cited barrier to technology use of older adults, with an unwillingness to adopt technologies stemming from high perceptions of risk and desire for privacy (27). However, older adult users in this study shared personal details including their age, gender, weight, height, meal information, and health-related measurements. Potential trust-building factors that may have been uniquely appealing to older adults warrant further exploration. Research has shown that there are both enablers to trust (e.g., fair data access, ease of use, lack of judgment) and impediments (e.g., fear of data exploitation, insufficient training) that digital health services must consider when designing their platforms (28). Such elements are critical since not just adoption of, but also effective engagement with, digital health platforms is necessary to reap the greatest health benefits and sustain these benefits (29).

TABLE 3 | Engagement metrics of users over a 12-month period broken down by program-specific grouping into clinically oriented (diabetes prevention, diabetes care, and hypertension care) and wellness (prevention).

Clinically oriented programs				
Engagement metrics	**Values**	**Full sample (n = 1,633)**	**35–64 years (n = 1,400)**	**65± years (n = 233)**
Number of coaching conversations	Median [IQ Range]	437 [281, 624]	431 [278, 614]	485 [300, 706]**
	Mean [95% CI]	494 [479, 509]	486 [470, 503]	540 [500, 581]
Number of meals logged	Median [IQ Range]	106 [45, 224]	102 [45, 214]	141 [49, 368]**
	Mean [95% CI]	183 [172, 194]	171 [160, 182]	252 [215, 288]
Number of device measurements	Median [IQ Range]	30 [9, 71]	28 [9, 67]	38 [13, 112]**
	Mean [95% CI]	75 [54, 96]	73 [49, 97]	84 [69, 99]
Wellness program				
Engagement Metrics	**Values**	**Full Sample (n = 536)**	**35–64 years (n = 468)**	**65± years (n = 68)**
Number of coaching conversations	Median [IQ Range]	434 [284, 585]	423 [269, 561]†	584 [398, 767]**†
	Mean [95% CI]	463 [441, 485]	439 [417, 460]	630 [548, 712]
Number of meals logged	Median [IQ Range]	69 [27, 187]	62 [25, 152]‡	285 [114, 493]**†
	Mean [95% CI]	157 [139, 175]	130 [113, 147]	347 [278, 416]
Number of device measurements	Median [IQ Range]	30 [13, 68]	28 [12, 63]	40 [29, 92]**
	Mean [95% CI]	60 [53, 67]	58 [50, 66]	73 [53, 93]

Statistical comparisons based first on the medians between age groups and denoted by $^{*}p \leq 0.05$; $^{**}p \leq 0.01$ next to older adult medians for both program-specific groupings. Statistical comparisons also presented based on the medians between program-specific groupings within each age group and denoted by † $p \leq 0.05$; $^{\ddagger}p \leq 0.01$ next to each age-group's median under the wellness program results.

Data presented for the full sample and stratified by age into older (65+ years) and younger (35–64 years) groups. Median [Interquartile (IQ) Range] and Mean [95% Confidence Interval (CI)] provided for all comparisons.

Studies of digital health technology use have shown that one reason why older adults may engage less with these technologies is simply that they are less likely than younger age groups to be offered digital health access by their healthcare provider (30). In fact, of those with access, one study of Canadian older adults found that older adults sustained their use of health-related mobile apps for longer than the general population (31). Older adults may be characterized in their use of digital technologies along a spectrum from non-users to savvy users like the general population (32). Despite suggestions to the contrary, some research has shown that older adults are willing to engage with new technologies and demonstrate positive attitudes toward technology (33). Although we did not independently assess each potential barrier, our results also collectively suggest that older adults will engage with digital health technologies when provided the opportunity.

Facilitators and Patterns of Use of Fully Digital Health Programs by Older Adults

The engagement of older adults with digital health is important to the field of chronic disease management. Many chronic diseases are preventable or effectively managed through lifestyle changes (34, 35). However, direct contact with healthcare professionals that may offer conventional lifestyle behavior coaching is a challenge due to the shortage of practitioners that provide care (36), and the costs associated with regular human-provided care (37), resulting in unmet care needs of older adults. Telehealth initiatives have been successfully deployed to older adults for chronic disease management (38); however, classical models of telehealth still require costly touchpoints with human care providers to facilitate program engagement (39). The engagement of older adults in the fully digital programs assessed in this study demonstrates that older adults readily adopted programs that did not require any human touchpoints and that fully digital programs may therefore present a widely scalable and cost-effective alternative to traditional forms of telehealth.

Given that increased engagement with lifestyle interventions has been associated with improved health outcomes (40), our findings require further exploration of the underlying facilitators supporting the engagement of older adult users. There are other potential facilitators of engagement in digital health programs besides age, such as clinician referral, incentives (e.g., compensation) for participation, and disease diagnosis (41). Exploring interactions between age and other potential facilitators is an important future area of focus to determine how to best facilitate program engagement for various subgroups. We found that older and younger adults differed in their engagement patterns between wellness vs. clinically oriented programs. In younger adults, there were more coaching conversations and meals logged in the clinically oriented programs when compared to those enrolled in the wellness program. In contrast, older adults in the wellness program had more coaching conversations and meals logged when compared with older adults enrolled in the clinically oriented programs. It is possible that the digital coach presented older adults with an opportunity for social interactions that they desired (42, 43), particularly when the program was not focused on clinical issues. If older adults viewed the digital coach as a form of social support, this could explain the greater number of interactions of apparently healthy older adults enrolled in the wellness program, and this would

support the use of fully digital health platforms for not only disease management but also for prevention of chronic disease—a minority focus of currently available digital platforms (44). Prevention is a critical area of focus for digital health programs because these technologies have the potential to stabilize or reverse the declining health of older adults before they need clinical intervention. Future work is necessary to elucidate these findings.

The results of this study must be considered with respect to the median age of the sample (68 years). Although our hypothesis was that older adults over 65 years would have less engagement with digital programs due to real or perceived barriers to technology use, a potential counter-hypothesis could be that the "younger" end of the older-adult spectrum may include newly retired individuals who have more free time to engage with digital technologies than working-age adults. We may have observed different results had we included the "oldest old" (≥80 years) who are even more likely to experience barriers to technology use (45).

Strengths and Limitations

We did not directly assess health outcomes as they related to engagement metrics. Such an assessment would be complicated due to the different outcomes associated with each of the different programs and was beyond the scope of this study. However, the high level of engagement of older adults is a promising indicator of the potential for fully digital health interventions, since we know from the published literature that those who engage in lifestyle interventions to a greater degree are generally more successful (40, 46). The present study included only users of an existing commercial digital health product. However, a strength is that these participants represented real-world users of digital health programs rather than participants recruited as a part of a carefully controlled research study. Users had no contact with research staff; thus, their engagement with the digital platform can be attributed to their personal choices rather than instructions to behave in a particular manner. We observed less diversity in race and ethnicity in older adults than younger adults. Race and ethnicity have been found to be predictive of digital health and technology use, with minority populations less likely to engage (47). The fact that we had few older adult users of non-white and Hispanic/Latino origins may support these findings, and more work is necessary to improve inclusivity and help mitigate health disparities. Finally, our measures of engagement assessed the total number of engagement metrics rather than temporal patterns of user-coach interactions, which recent studies have suggested may be predictive of individual outcomes (48). A more detailed understanding of the ways in which older adult users interact with the digital platform will be key to optimizing the mechanisms of coaching delivery (e.g., content, timing, and frequency) and platform navigations.

CONCLUSIONS

The present study found that older adults had greater engagement in coaching conversations, meals logged, and device measurements than younger adults, suggesting that older adults were able to navigate a digital screen, interact with a fully automated digital coach, and take measurements with smart and connected digital devices. Health-related digital technologies and digital coaches may offer older adults a way to manage the large amount of information associated with lifestyle behavior changes, and further, provide 24-h continuous encouragement and support in sustaining these lifestyle changes. Our findings collectively suggest that older adults will engage with digital health technologies when provided the opportunity. These findings support the use of fully digital health programs to deliver behavior change interventions for older adults and provide a foundation for future studies to explore age-specific relationships of patterns of engagement and outcomes.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Advarra Institutional Review Board approval (Protocol #Pro00047181) for retrospective analyses of previously collected and de-identified data. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

SG, NS, and OB wrote and edited the manuscript. FS performed data analyses and edited the manuscript. JP provided clinical oversight and edited the manuscript. SK designed the study and contributed to writing and editing the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by Lark Technologies, Inc.

ACKNOWLEDGMENTS

We thank the team at Lark for their assistance in managing the data collection.

REFERENCES

1. Mathews SC, McShea MJ, Hanley CL, Ravitz A, Labrique AB, Cohen AB. Digital health: a path to validation. *NPJ Digit Med.* (2019) 2:1–9. doi: 10.1038/s41746-019-0111-3
2. Mann DM, Chen J, Chunara R, Testa PA, Nov O. COVID-19 transforms health care through telemedicine: evidence from the field.

J Am Med Informatics Assoc. (2020) 27:1132–5. doi: 10.1093/jamia/ocaa072
3. Gelburd R. The Coronavirus Pandemic and the Transformation of Telehealth. *US News.* Available online at: https://www.palmettocareconnections.org/news/the-coronavirus-pandemic-and-the-transformation-of-telehealth/ (accessed June 1, 2020).
4. Ramaswamy A, Yu M, Drangsholt S, Ng E, Culligan PJ, Schlegel PN, et al. Patient satisfaction with telemedicine during the COVID-19 pandemic: retrospective cohort study. *J Med Internet Res.* (2020) 22:e20786. doi: 10.2196/20786
5. US Census Bureau. *National Population Projections*. Population Estimates Program (2017). Available online at: https://www.census.gov/data/tables/2017/demo/popproj/2017-summary-tables.html (cited May 12, 2020).
6. Mitchell EM. *Concentration of Health Expenditures in the US Civilian Noninstitutionalized Population, 2014.* Rockville, MD: Agency for Healthcare Research and Quality (US) (2017).
7. Evangelista L, SR, Topol EJ. Digital health care for older adults. *Lancet.* (2019) 393:1493. doi: 10.1016/S0140-6736(19)30800-1
8. Wang J. Mobile and connected health technologies for older adults aging in place. *J Gerontol Nurs.* (2018) 44:3–5. doi: 10.3928/00989134-20180509-01
9. Medicare.gov: Telehealth. *Your Medicare Coverage.* (2020). Available online at: https://www.medicare.gov/coverage/telehealth (cited December 8, 2020).
10. Health and Human Services Administration. *Policy Changes During COVID-19*. TELEHEALTH.HHS.GOV (2020). Available online at: https://telehealth.hhs.gov/providers/policy-changes-during-the-covid-19-public-health-emergency/ (cited May 12, 2020).
11. Wildenbos GA, Jaspers MWM, Schijven MP, Dusseljee-Peute LW. Mobile health for older adult patients: using an aging barriers framework to classify usability problems. *Int J Med Inform.* (2019) 124:68–77. doi: 10.1016/j.ijmedinf.2019.01.006
12. Hall AK, Bernhardt JM, Dodd V, Vollrath MW. The digital health divide: evaluating online health information access and use among older adults. *Heal Educ Behav.* (2015) 42:202–9. doi: 10.1177/1090198114547815
13. Centers for Disease Control Prevention. *Percent of U.S. Adults 55 and Over with Chronic Conditions*. National Center for Health Statistics. (2015). Available online at: https://www.cdc.gov/nchs/health_policy/adult_chronic_conditions.htm (accessed December 20, 2020).
14. Bolnick HJ, Bui AL, Bulchis A, Chen C, Chapin A, Lomsadze L, et al. Health-care spending attributable to modifiable risk factors in the USA: an economic attribution analysis. *Lancet Public Heal.* (2020) 5:e525–e35. doi: 10.1016/S2468-2667(20)30203-6
15. Alzeidan R, Shata Z, Hassounah MM, Baghdadi LR, Hersi A, Fayed A, et al. Effectiveness of digital health using the transtheoretical model to prevent or delay type 2 diabetes in impaired glucose tolerance patients: protocol for a randomized control trial. *BMC Public Health.* (2019) 19:1550. doi: 10.1186/s12889-019-7921-8
16. Widmer RJ, Collins NM, Collins CS, West CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. *Mayo Clin Proc.* (2015) 90:469–80. doi: 10.1016/j.mayocp.2014.12.026
17. Labrique AB, Wadhwani C, Williams KA, Lamptey P, Hesp C, Luk R, et al. Best practices in scaling digital health in low and middle income countries. *Global Health.* (2018) 14:103. doi: 10.1186/s12992-018-0424-z
18. Centers for Disease Control and Prevention. *Using Telehealth to Expand Access to Essential Health Services during the COVID-19 Pandemic*. National Center for Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases (2020). Available from: https://www.cdc.gov/coronavirus/2019-ncov/hcp/telehealth.html (accessed December 10, 2020).
19. Vaportzis E, Giatsi Clausen M, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. *Front Psychol.* (2017) 8:1687. doi: 10.3389/fpsyg.2017.01687
20. Berkowsky RW, Sharit J, Czaja SJ. Factors predicting decisions about technology adoption among older adults. *Innov Aging.* (2017) 1:igy002. doi: 10.1093/geroni/igy002
21. Greenwood DA, Gee PM, Fatkin KJ, Peeples M. A systematic review of reviews evaluating technology-enabled diabetes self-management education and support. *J Diabetes Sci Technol.* (2017) 11:1015–27. doi: 10.1177/1932296817713506
22. Kim JY, Wineinger NE, Steinhubl SR. The influence of wireless self-monitoring program on the relationship between patient activation and health behaviors, medication adherence, and blood pressure levels in hypertensive patients: a substudy of a randomized controlled trial. *J Med Internet Res.* (2016) 18:e5429. doi: 10.2196/jmir.5429
23. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. *J Med Internet Res.* (2016) 18:e6. doi: 10.2196/jmir.4790
24. Alkire L, O'Connor GE, Myrden S, Köcher S. Patient experience in the digital age: an investigation into the effect of generational cohorts. *J Retail Consum Serv.* (2020) 57:102221. doi: 10.1016/j.jretconser.2020.102221
25. Centers for Disease Control and Prevention. *Diabetes Prevention Recognition Program Standards and Operating Procedures CDC Diabetes Prevention Recognition Program.* (2018). Available online at: www.cdc.gov/diabetes/prevention/recognition (accessed December 10, 2020).
26. Levine DM, Lipsitz SR, Linder JA. Changes in everyday and digital health technology use among seniors in declining health. *J Gerontol Ser A.* (2018) 73:552–9. doi: 10.1093/gerona/glx116
27. Fox G, Connolly R. Mobile health technology adoption across generations: narrowing the digital divide. *Inf Syst J.* (2018) 28:995–1019. doi: 10.1111/isj.12179
28. Adjekum A, Blasimme A, Vayena E. Elements of trust in digital health systems: scoping review. *J Med Internet Res.* (2018) 20:e11254. doi: 10.2196/11254
29. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med.* (2016) 51:833–42. doi: 10.1016/j.amepre.2016.06.015
30. Peacock S, Reddy A, Leveille SG, Walker J, Payne TH, Oster N V, et al. Patient portals and personal health information online: perception, access, and use by US adults. *J Am Med Informatics Assoc.* (2017) 24:e173–e7. doi: 10.1093/jamia/ocw095
31. Jaana M, Paré G. Comparison of mobile health technology use for self-tracking between older adults and the general adult population in Canada: cross-sectional survey. *JMIR mHealth uHealth.* (2020) 8:e24718. doi: 10.2196/24718
32. Quan-Haase A, Williams C, Kicevski M, Elueze I, Wellman B. Dividing the grey divide: deconstructing myths about older adults' online activities, skills, and attitudes. *Am Behav Sci.* (2018) 62:1207–28. doi: 10.1177/0002764218777572
33. Kuerbis A, Mulliken A, Muench F, Moore AA, Gardner D. Older adults and mobile technology: factors that enhance and inhibit utilization in the context of behavioral health. *Ment Health Addict Res.* (2017) 2:1:11. doi: 10.15761/MHAR.1000136
34. Adams ML, Grandpre J, Katz DL, Shenson D. The impact of key modifiable risk factors on leading chronic conditions. *Prev Med (Baltim).* (2019) 120:113–8. doi: 10.1016/j.ypmed.2019.01.006
35. Kushner RF, Sorensen KW. Lifestyle medicine: the future of chronic disease management. *Curr Opin Endocrinol Diabetes Obes.* (2013) 20:389–95. doi: 10.1097/01.med.0000433056.76699.5d
36. Markit IH. *The Complexities of Physician Supply and Demand: Projections From 2015 to 2030*. Association of American Medical Colleges (2017). Available online at: https://www.luminafoundation.org/wp-content/uploads/2017/08/aamc-projections-update-2017.pdf (accessed December 10, 2020).
37. Osborn R, Doty MM, Moulds D, Sarnak DO, Shah A. Older Americans were sicker and faced more financial barriers to health care than counterparts in other countries. *Health Aff.* (2017) 36:2123–32. doi: 10.1377/hlthaff.2017.1048
38. Guo Y, Albright D. The effectiveness of telehealth on self-management for older adults with a chronic condition: a comprehensive narrative review of the literature. *J Telemed Telecare.* (2018) 24:392–403. doi: 10.1177/1357633X17706285
39. Henderson C, Knapp M, Fernández J-L, Beecham J, Hirani SP, Cartwright M, et al. Cost effectiveness of telehealth for patients with long term conditions (Whole Systems Demonstrator telehealth questionnaire study): nested economic evaluation in a pragmatic, cluster randomised controlled trial. *BMJ.* (2013) 346:f1035. doi: 10.1136/bmj.f1035

40. Ely EK, Gruss SM, Luman ET, Gregg EW, Ali MK, Nhim K, et al. A national effort to prevent type 2 diabetes: participant-level evaluation of CDC's National Diabetes Prevention Program. *Diabetes Care.* (2017) 40:1331–41. doi: 10.2337/dc16-2099
41. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *npj Digit Med.* (2020) 3:1–10. doi: 10.1038/s41746-020-0224-8
42. Brandtzaeg PB, Følstad A. Why people use chatbots. In: Kompatsiaris I, Cave J, Satsiou A, Carle G, Passani A, Kontopoulas E, et al., editors. *Internet Science. INSCI 2017, Vol. 106373, Lecture Notes in Computer Science*. Cham: Springer (2017).
43. De Gennaro M, Krumhuber EG, Lucas G. Effectiveness of an empathic chatbot in combating adverse effects of social exclusion on mood. *Front Psychol.* (2020) 10:3061. doi: 10.3389/fpsyg.2019.03061
44. Cohen AB, Dorsey ER, Mathews SC, Bates DW, Safavi K. A digital health industry cohort across the health continuum. *NPJ Digit Med.* (2020) 3:1–10. doi: 10.1038/s41746-020-0276-9
45. Piper AM, Garcia RC, Brewer RN. Understanding the challenges and opportunities of smart mobile devices among the oldest old. *Int J Mob Hum Comput Interact.* (2016) 8:83–98. doi: 10.4018/IJMHCI.2016040105
46. Hibbard JH, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. *Health Aff.* (2013) 32:207–14. doi: 10.1377/hlthaff.2012.1061
47. Mitchell UA, Chebli PG, Ruggiero L, Muramatsu N. The digital divide in health-related technology use: The significance of race/ethnicity. *Gerontologist.* (2019) 59:6–14. doi: 10.1093/geront/gny138
48. Chien I, Enrique A, Palacios J, Regan T, Keegan D, Carter D, et al. A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions. *JAMA Netw open.* (2020) 3:e2010791. doi: 10.1001/jamanetworkopen.2020.10791

Conflict of Interest: SG, FS, NS, OB, JP, and SK receive a salary from Lark and some authors might have some stock options in the event that the company could go public at an unanticipated date in the future. The funder (Lark Technologies, Inc.) had the following involvement in the study: collection of data, analyses, and writing.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 07 February 2022
doi: 10.3389/fdgth.2022.814179

Dynamics of Loneliness Among Older Adults During the COVID-19 Pandemic: Pilot Study of Ecological Momentary Assessment With Network Analysis

Varsha D. Badal[1,2], *Ellen E. Lee*[1,2,3], *Rebecca Daly*[1,2], *Emma M. Parrish*[4], *Ho-Cheol Kim*[5], *Dilip V. Jeste*[1,2,6] *and Colin A. Depp*[1,2,7]*

[1] Department of Psychiatry, University of California, San Diego, San Diego, CA, United States, [2] Department of Psychiatry, Sam and Rose Stein Institute for Research on Aging, University of California, San Diego, San Diego, CA, United States, [3] Desert-Pacific Mental Illness Research Education and Clinical Center, Veterans Affairs San Diego Healthcare System, San Diego, CA, United States, [4] San Diego State University/University of California, San Diego Joint Doctoral Program in Clinical Psychology, San Diego, CA, United States, [5] AI and Cognitive Software, International Business Machines (IBM) Research-Almaden, San Jose, CA, United States, [6] Department of Neurosciences, University of California, San Diego, San Diego, CA, United States, [7] Veterans Affairs (VA) San Diego Healthcare System, La Jolla, CA, United States

OPEN ACCESS

Edited by:
Patricia A. Arean, University of Washington, United States

Reviewed by:
Clara Caldeira, Indiana University Bloomington, United States
Oswald David Kothgassner, Medical University of Vienna, Austria

***Correspondence:**
Colin A. Depp
cdepp@ucsd.edu

Specialty section:
This article was submitted to Human Factors and Digital Health, a section of the journal Frontiers in Digital Health

Received: *12 November 2021*
Accepted: *05 January 2022*
Published: *07 February 2022*

Citation:
Badal VD, Lee EE, Daly R, Parrish EM, Kim H-C, Jeste DV and Depp CA (2022) Dynamics of Loneliness Among Older Adults During the COVID-19 Pandemic: Pilot Study of Ecological Momentary Assessment With Network Analysis. Front. Digit. Health 4:814179. doi: 10.3389/fdgth.2022.814179

Objective: The COVID-19 pandemic has had potentially severe psychological implications for older adults, including those in retirement communities, due to restricted social interactions, but the day-to-day experience of loneliness has received limited study. We sought to investigate sequential association, if any, between loneliness, activity, and affect.

Methods: We used ecological momentary assessment (EMA) with dynamic network analysis to investigate the affective and behavioral concomitants of loneliness in 22 residents of an independent living sector of a continuing care retirement community (mean age 80.2; range 68–93 years).

Results: Participants completed mean 83.9% of EMA surveys (SD = 16.1%). EMA ratings of loneliness were moderately correlated with UCLA loneliness scale scores. Network models showed that loneliness was contemporaneously associated with negative affect (worried, anxious, restless, irritable). Negative (but not happy or positive) mood tended to be followed by loneliness and then by exercise or outdoor physical activity. Negative affect had significant and high inertia (stability).

Conclusions: The data suggest that EMA is feasible and acceptable to older adults. EMA-assessed loneliness was moderately associated with scale-assessed loneliness. Network models in these independent living older adults indicated strong links between negative affect and loneliness, but feelings of loneliness were followed by outdoor activity, suggesting adaptive behavior among relatively healthy adults.

Keywords: aging, dynamic networks, causal networks, positive affect, negative affect, social isolation, loneliness

INTRODUCTION

Social Isolation and Loneliness (SI/L) have assumed pandemic proportions over recent decades, in part driven by globalization and ultra-rapid rise in technology (1, 2). The situation has been exacerbated by the ongoing containment measures for the COVID-19 pandemic and mandated lockdowns. The impact could be greater in older adults due to their physical vulnerability (2, 3). However, studies have shown higher levels of resilience and wisdom in older than in younger adults during the pandemic (4). A common inference during COVID-19 pandemic and the ensuing strict isolation measures is that older adults in independent living conditions were likely to have encountered loneliness (5–8); During the pandemic period, it is unclear whether and how day-to-day or micro-level experiences of loneliness related to affect or behavior.

Chronic loneliness is a consistent set of beliefs regarding the lack of connections with others and yet state loneliness refers immediate experience of social disconnection. Under Cacioppo model, state loneliness is not necessarily negative but may motivate behavior such as outreach or seeking social interaction (9, 10). Loneliness and social isolation are weakly correlated (11–13). An individual's relationships such as friends and family may influence activity (14) and social isolation was associated with behavioral inactivity in general (15). The relationship between loneliness and social behavior is somewhat unclear in older adults. Loneliness was not found to be related to social activity among older adults in one study (16). Moreover, the COVID-19 pandemic placed additional restrictions on mobility which further may have altered social behavior. While much is known about chronic loneliness and long-term health effects, the impact of state loneliness on day-to-day behavior is less researched, particularly its dynamics among older adults during the pandemic.

Ecological momentary assessment (EMA) allows for relatively unobtrusive monitoring of affect as well as physical and social context variables, which when monitored repeatedly over time, can uncover dynamic relationships between variables (17). While traditional approaches are limited to discovering associations or correlations, EMA allows one not only to establish the said associations, but also time lags and leads which enable hypotheses for possible causality (18). A recent meta-analysis suggested 81.9% mobile EMA compliance in adults (19). However, to date, use of EMA in the "older-old" adults (persons over age 80) or to study loneliness is somewhat limited, with none focused on loneliness during the pandemic using network models. A broad search in PubMed on EMA in geriatric populations with manual screening of mean age around 70 produced a few results, focused on perception and usability (20–22), and diverse applications included adverse event monitoring (23), Multiple Sclerosis (24) and pain (25). Some EMA studies have included adults with mean ages ranging from 69 to 73, suggesting feasibility (26–28). An EMA study on loneliness in the older population (mean age 73.7) suggested men reported greater intensity of loneliness, and being outdoors lessened the feeling—the effects were weaker among women and non-Whites (29). Another EMA study, not limited to older adults, during COVID-19 lockdown found that a composite "negative-mood" score (comprising fatigue, anxiety, stress, depression and unhappiness) tended to accumulate over time, and the score was positively and significantly associated with COVID19-related worry, the perception of restrictions, and loneliness (30).

EMA studies evaluating lagged associations (e.g., mood associations with subsequent behavior or vice versa) typically evaluated fixed time lags and univariate relationships. However, between-people networks, constructed by combining data from several individuals, allow us to discover multiple contemporaneous and lagged associations representative of the group (17). To our knowledge, this is the first study to apply network models to EMA data to explore the loneliness experience of older adults (mean age 80+) residing in senior housing communities during the COVID-19 pandemic. Due to the older mean age of this sample (80 years) than in prior studies, we evaluated both the feasibility of EMA with respect to adherence and also convergence of EMA questions on loneliness with standard scale-based measures of loneliness. We then applied network models to evaluate sequential relationships and moment-to-moment interactions among emotions, and loneliness, and behavior.

We hypothesized that: (a) Older adults would evidence acceptable (e.g., >75% adherence to EMA procedures, (b) Loneliness as measured by EMA would be significantly associated with an in-lab scale-based measure of loneliness (UCLA Loneliness Scale), and (c) Network models applied to EMA data would reveal significant contemporaneous and lagged connections between momentary loneliness, affect, and social behavior.

METHODS

Participants

Participants were recruited from an ongoing longitudinal study of older adults aged 65 years and above living independently in a Continued Care Senior Housing Community (CCSHC) (31). Participants were contacted by study staff to assess level of interest. Eligibility requirements included current enrollment in the parent study and access to a smartphone capable of receiving daily text messages and surveys. Parent study exclusion criteria included people with dementia, major mental illness or other conditions that could interfere with study participation and those who are unable to read and write in English. The sample ($n = 22$) included 19 women and three men (**Table 1**). The EMA surveys were collected between 5/25/2020 and 8/16/2020.

The study protocol was approved by the UC San Diego Human Research Protections Program (HRPP) and all the participants provided a written informed consent prior to study participation.

Measures

Assessments included sociodemographic as well as clinical measures of depression (Patient Health Questionnaire, 9-item, or, PHQ-9) (33), anxiety (Brief Symptom Inventory—Anxiety subscale, or, BSI) (34), and UCLA Loneliness scale (Version 3) or UCLA-3 (35) which is a 20-item scale. The tests were

TABLE 1 | Socio-demographic and clinical factors ($N = 22$).

	Mean or %	SD	Min	Max
Socio-demographic				
Age (years)	80.24	7.13	68.2	93.4
Education (years)*	15.59	2.63	12.0	20.0
Race (% Caucasian)*	91%			
Marital Status (% married/co-habitating)	32%			
Loneliness and social support measures				
UCLA-3 1st Administration	35.86	7.92	24.0	49.0
UCLA-3 2nd Administration	29.87	5.74	23.0	44.0
UCLA-3 3rd Administration	33.33	9.55	24.0	56.0
UCLA Averaged over all available	34.77	7.86	24.5	49.3
Emotional Support* (ESS-E)	2.74	0.46	1.5	3.0
Instrumental Support* (ESS-I)	1.67	0.83	0.5	3.0
Negative social interactions* (ESS-NI)	0.33	0.43	0.0	1.5
Clinical measures				
Depression* (PHQ-9)	2.14	2.41	0.0	8.0
Anxiety* (BSIAS)	1.86	3.48	0.0	12.0
EMA measures				
Worried	1.344	0.61	1.0	5.0
Happy	4.040	0.94	1.0	5.0
Anxious	1.616	0.83	1.0	5.0
Restless	1.328	0.61	1.0	5.0
Irritable	1.200	0.51	1.0	5.0
Lonely	1.248	0.54	1.0	5.0
Exercise	1.995	0.92	1.0	5.0
Outdoor	1.733	1.12	1.0	5.0
Social interaction	2.208	1.21	1.0	5.0

BSIAS, Brief Symptom Inventory Anxiety Scale; ESS-E, Emotional Support Scale—Emotional Support score; ESS-I, Emotional Support Scale—Instrumental Support; ESS-NI, Emotional Support Scale—Negative Interaction Score (32); PHQ-9, Patient Health Questionnaire 9-item (33); UCLA-3, UCLA Loneliness Scale (Version 3).
**Baseline data.*

administered between 5/25/2020 and 8/16/2020. The scores on UCLA-3 loneliness scale can be interpreted as low (range: 20–34), moderate (range: 35–49), moderately high (range: 50–64), and high (range: 65–80) (36, 37). For descriptive purposes, we also administered the PHQ-9 scale for depression, wherein score ranges from mild (5–9), moderate (10–14), moderately severe (15–19) and severe depression ($\geq$20), respectively, along with the BSI anxiety subscale (34, 38) comprises six items, it is a self-report measure of anxiety that ranges from 0 to 24 with higher scores indicating a greater level of anxiety.

EMA Procedure

Participants were sent text notifications to their personal smartphones to complete the smartphone-based surveys three times daily for 7 days through the online-based survey platform, Alchemer. Each text notification contained a unique participant link to the study surveys. The daily survey notifications were sent at varying times each day, with a minimum 4-h increment between surveys. Participants received the surveys once in the morning, once in the afternoon, and once at night. Two participants opted out of the morning surveys and requested to receive afternoon and evening surveys only. Upon receiving the link, participants completed EMA questions assessing context, mood, and behaviors. Once the link was delivered, the morning and afternoon surveys stayed active for at least 3 h, until 1 h prior to the next scheduled survey being sent, at which point the survey was closed and no longer accessible. The evening surveys closed at 11:00 p.m. each night. Study surveys were linked to participant's smartphone number and were therefore opened only by the participant's device. Deidentification of participant's data was performed and the data was not stored locally on the devices. Survey data were sent to encrypted, HIPAA-compliant cloud storage in Amazon Web Services (AWS), and responses were recorded even if participants did not complete the entire survey. Real-time access to participant's data and daily progress was available through the AWS system. When three surveys in a row were missed by the participants, they were contacted by the research staff to address any technical difficulties or adherence issues.

Each survey was comprised of the 15 EMA prompts related to the previous 2 h; out of these, the responses to following nine prompts were used in the study:

(1) How worried were you generally? (2) how happy vs. sad were you? (3) how relaxed vs. anxious were you? (4) how fidgety or restless were you? (5) how irritable or easily angered have you been? (6) how lonely were you? (7) how many minutes did you exercise or move regularly? (8) how many minutes did you spend time outdoors? and (9) how many people did you spend time with? All responses were scored on 1–5 scale, interpreted from the lowest to the highest intensity based on the prompt context.

Statistical Analysis

Pearson's correlation was used to assess correlations between EMA variables and UCLA-3 measures of loneliness.

Network Analysis

Time-series for the EMA response variables were constructed by splicing together the data for each participant (in the same order across the variables). Tigramite, the python implementation of PCMCI (39) algorithm was used to construct the temporal networks with contemporaneous and lagged edges. Temporal lags up to six sampling intervals (2 days) were analyzed. The implementation is designed to handle some missing data when appropriately tagged. It generates error when an unacceptable amount of data is missing, however, we did not encounter that situation.

Unlike studies based upon effect sizes that draw direct benefit from large sample sizes, small sample correlation-based studies are susceptible to type-1 error, of identifying correlations when none exists in larger population. Since our sample was small ($n = 22$), our network models use PCMCI that incorporates Benjamini–Hochberg Method (40) (also called BH procedure) to limit false discovery rate.

RESULTS

Demographic and clinical details are presented in **Table 1**. The mean age of participants was 80.24(SD = 7.13) years, 32% were

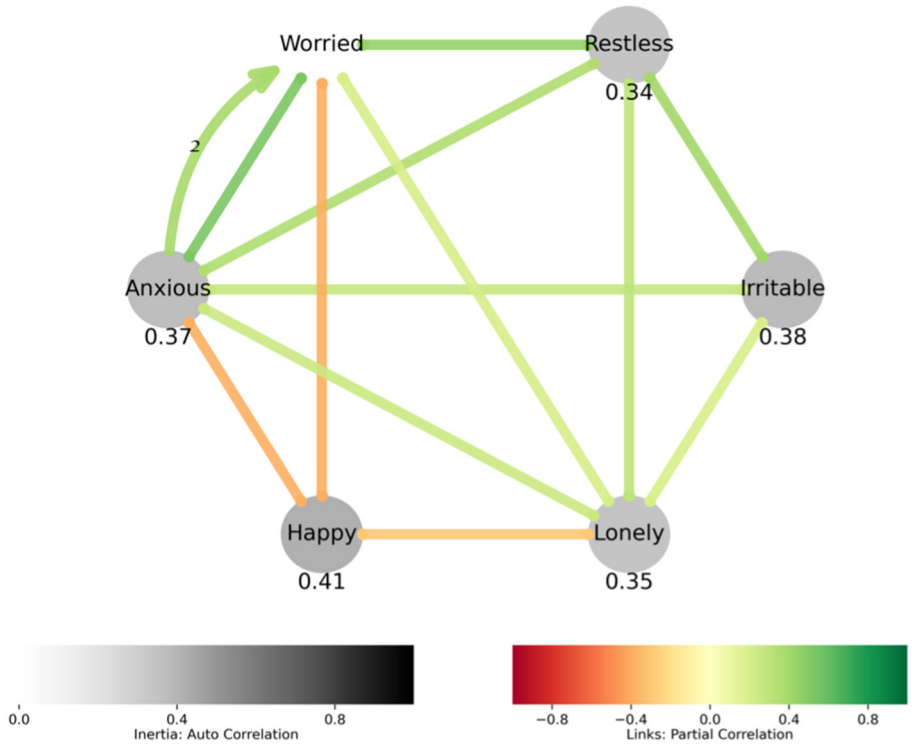


FIGURE 1 | Affect model describing how loneliness relates to affect: Undirected straight edges between variables indicate contemporaneous associations, and the directed labeled arcs represent lagged associations, with the label being the lag in multiples of sampling interval. The colors of the edges and the arcs represent negative (red) or positive (green) association. The variables with gray nodes indicate significant autocorrelation, or inertia, which can be interpreted as the values of these variables showing high resistance to change. If a variable measures polar quantity (happy-sad or relaxed-anxious), the variable is represented in the graph by the label to which higher values are assigned, and "how happy vs. sad were you" is simply "Happy." Positive associations are in shades of green, and negative in shades of red. Lagged links are curved, have arrowheads and display lag in multiples of 8 h. Negative emotional states are associated with loneliness. Inverse relationship between happy and loneliness is also expected. Anxiety and worry display a positive feedback loop.

married or cohabitating. Average loneliness score on UCLA-3 scale was 34.8 (SD = 7.86). Scores for depressive and anxious symptoms indicated minimal severity, well within the normal range. Participants had over 15 years of education on an average (**Table 1**) and resided in a single continuing care community that had spaces for both socializing and exercising. Of the 22 participants, one reported ethnicity as Asian and one as African American, the rest reported Caucasian.

Average adherence to the EMA surveys was 83.9% (SD = 16.1%) or an average of 17.0 (SD = 3.6) responses out of a total of 21 survey opportunities. Two participants opted out of the morning surveys and requested to only be sent surveys in the afternoon and evening, therefore receiving 14 survey opportunities each. Evening surveys had the highest adherence at 86.4% (SD = 20%), afternoon surveys had the second highest adherence at 84.4% (SD = 19.2%), and morning survey had the lowest adherence at 80.0% (SD = 18.2%). In addition to the high rate of surveys completed (84% of administered) all participants who were approached to participate in EMA surveys enrolled and completed the 7-day protocol. Notably, adherence was worse on the first few days and then improved (Spearman's $r = 0.33$, $p < 0.001$), thus, EMA surveys were not associated with fading or fatigue effects but rather non-adherence problems at the outset that resolved. There were however, two participants who consistently declined to respond to the morning survey but were allowed to continue in the study.

EMA loneliness was associated with UCLA-3 Loneliness Scale ($r = 0.375$). **Supplementary Table 1** shows the correlations for the EMA affective variables, with EMA loneliness correlated significantly with positive affect and fidgety/restlessness, but not other affective states.

The networks (**Figures 1–3**) show subsets of variables analyzed, and their lagged and contemporaneous associations. Network analysis of affective experience identified that loneliness was contemporaneously associated with feelings of restlessness, worry, irritability and anxiety and a lack of happiness (**Figure 1**).

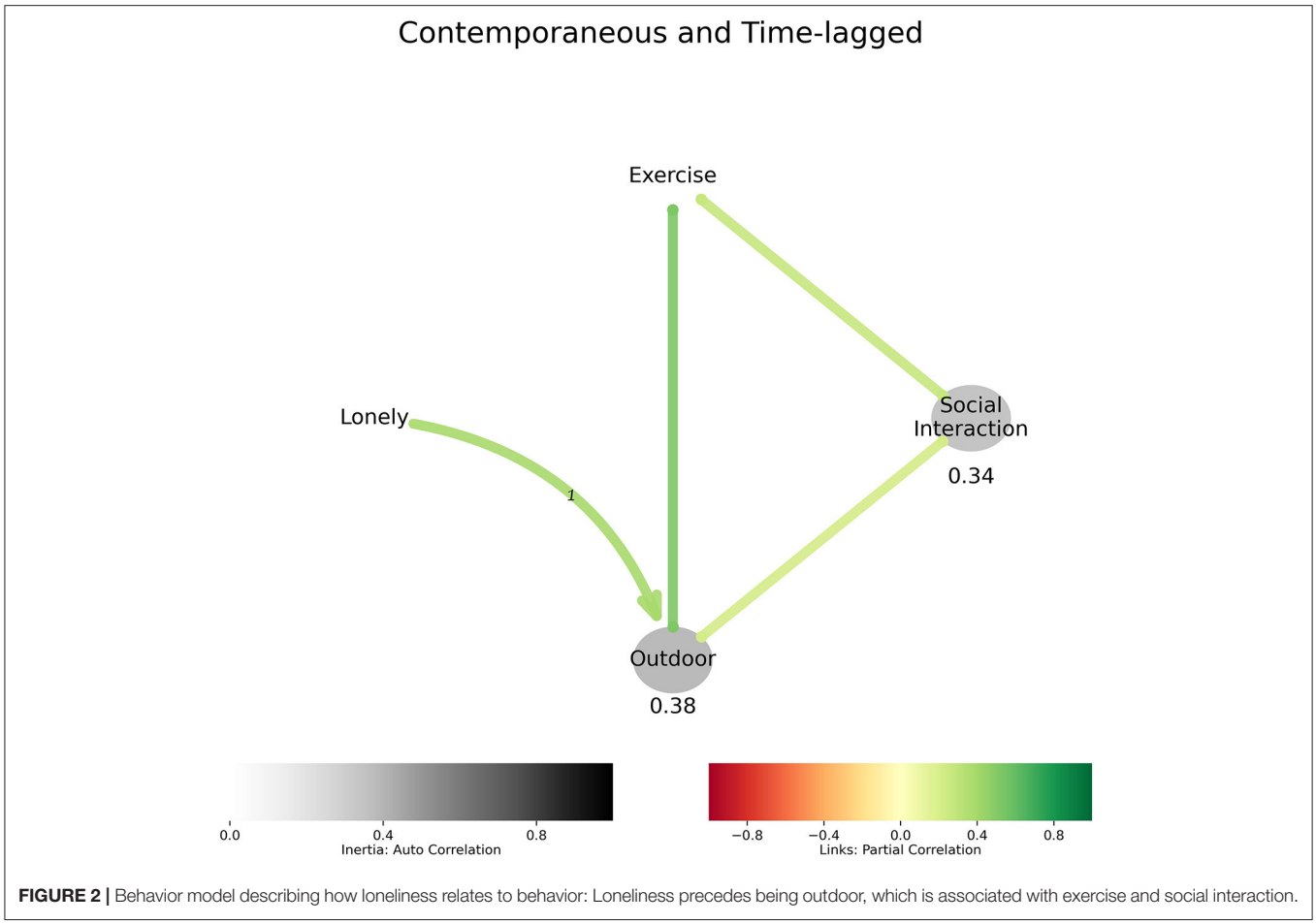


FIGURE 2 | Behavior model describing how loneliness relates to behavior: Loneliness precedes being outdoor, which is associated with exercise and social interaction.

A positive feedback loop between anxiety and worry suggests these experiences may converge to increase each other.

Figure 2 evaluated loneliness and resultant behaviors. Loneliness preceded being outdoors in the short-term and being outdoors was contemporaneous with exercise and social interaction in these older adults.

Figure 3 integrates affect and behavior models and shows that loneliness was strongly associated with negative feelings and a general lack of happiness. Being outdoors was associated with lower irritability. **Figure 3** also shows a relationship that seemed to exist between being lonely and being outdoors. Since the two did not exist contemporaneously, loneliness can be interpreted as being experienced when indoors. This was followed by an outdoor-seeking adaptive behavior that showed up in the next sampling (a lag of 1τ or, 8 h) when the participant was outdoors. The feeling of loneliness seemed to return soon after returning from outdoors (again, a lag of 1τ or, 8 h), and being outdoors was associated with exercise and social interaction.

DISCUSSION

We used EMA and dynamic network models to explore loneliness and its behavioral and affective concomitants in a sample of older adults. The primary findings from this study are three-fold; (1) EMA of loneliness and its concomitants was a feasible technique in older adults (mean age 80.2 years) with a sample adherence rate of 83.9%. (2) EMA of momentary loneliness was moderately associated with scale-assessed loneliness (UCLA-3 Loneliness Scale). (3) Network models displayed a variety of links between loneliness, affect and behavior. While loneliness was associated with negative emotions, our results suggest that loneliness was associated with short-term adaptive behavior, in particular spending time outdoors. This temporal finding is supported by another EMA study that being outdoors lessened the feeling of loneliness in the short term (29). These network models point to the need for future research to understand the behavioral sequelae of loneliness, delineating adaptive and maladaptive responses (and the influence of policies on those responses) to acute loneliness as they might contribute to or mitigate chronic loneliness.

The finding that loneliness is associated with negative emotions and diminished happiness is not surprising and is consistent with other studies (41–43), and during the lockdown in particular (30). A potentially novel finding through network models applied to EMA data is that at least some older people may have coped with momentary experiences of loneliness by actively seeking outdoor activity. There was a strong association

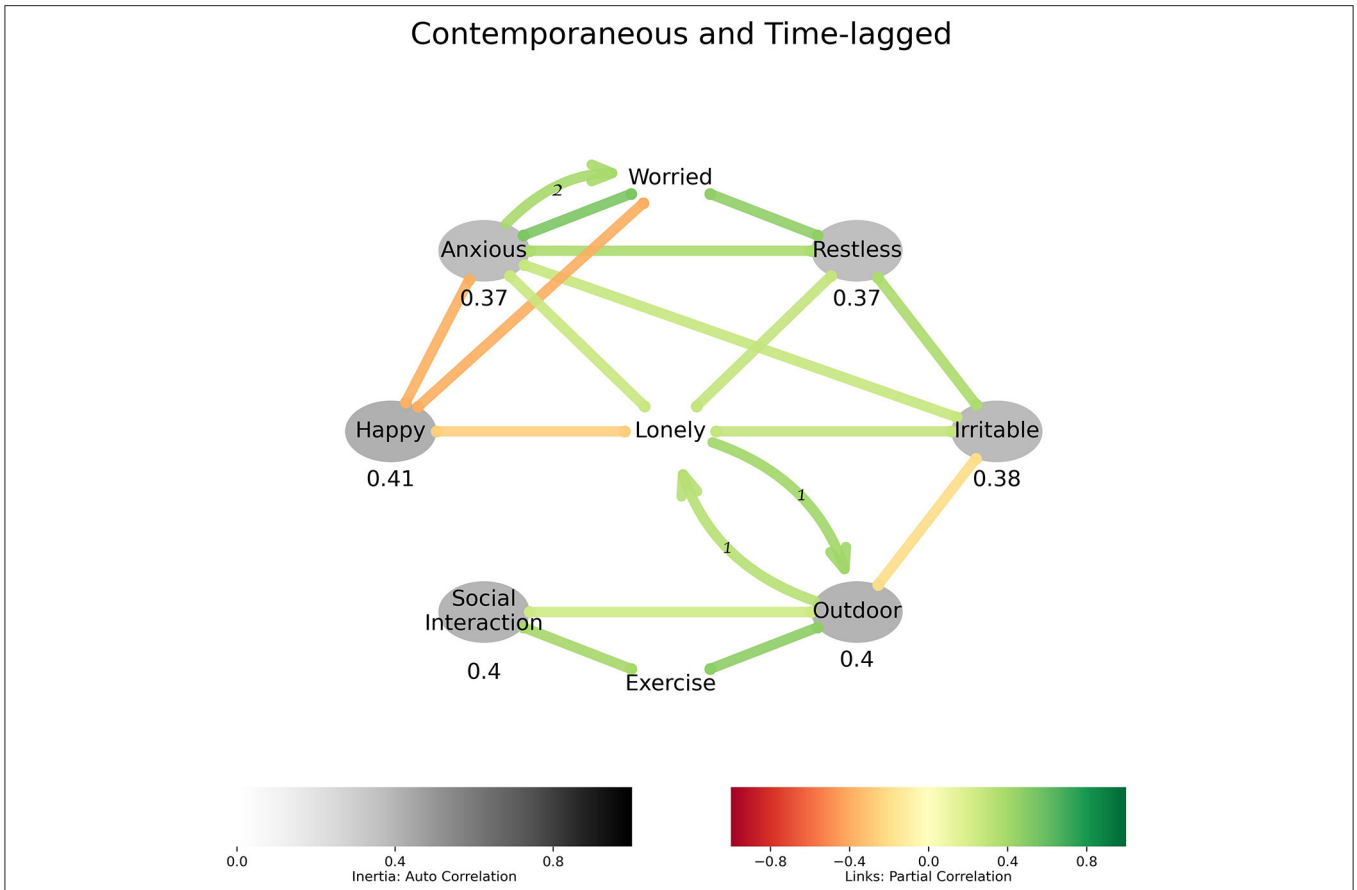


FIGURE 3 | Affect, loneliness, and behavior: Loneliness is predictive of being outdoors. Loneliness is not contemporaneous to being outdoors, it precedes it and returns soon after.

between being outdoor and exercise, and exercise and social interaction, but a weaker association between outdoor and social interaction in **Figure 3**. These findings are consistent with the literature indicating that a direct link between loneliness and social interaction behavior is weaker than might be expected (11–13). Since loneliness was not contemporaneous with being outdoors (and its correlates of activity and social interaction), it can be inferred to be associated with lower activity levels in the moment, and subsequent outdoor time. In that sense, acute loneliness, in this relatively healthy sample with a low level of distress, may have led to adaptive social behaviors. It has previously been suggested that loneliness serves a variety of adaptive functions (44). Previous literature also shows that coping mechanisms also differ by severity of depression among older adults, as self-distraction has been shown to be common among people with depression depressed group, while active coping was common among people without depression (45). Furthermore, our results are consistent with emotion/loneliness preceding activity, as in a different study, activity in-and-of itself had little effect on positive or negative affect (46). A study identified going outdoors as a coping strategy for social isolation during the pandemic among adults and included it in the survey (47), however no significant difference was observed in social isolation of those who did and did not seek outdoors. In an online study that included PHQ-9 questions and coping strategies, staying outdoors and looking outside were among the best predictors of lower levels of depressive symptoms associated with COVID-19 related isolation (48). Thus, how acute loneliness intersects with chronic loneliness is an important area for future research; EMA may be useful for contrasting loneliness at different time scales from day-to-day variations to more chronic experiences as well as for identifying which individuals would most likely benefit from specific types of interventions (e.g., those best suited for acute or chronic loneliness).

It was notable that the adaptive response to loneliness in this sample was to go outdoors. Variation in the extent of lockdowns or shelter-in-place guidelines observed during the pandemic may have influenced how people accessed outdoor activity and putatively coped with loneliness. Since this was a single-site study, it is not possible to evaluate variation by outdoor access. Nonetheless, technologically based alternative solutions to provide adaptive opportunities might be considered to help older adults cope with loneliness under circumstances where access to outdoor activity may be restricted.

This study has some limitations, and it should be considered as a preliminary work to test feasibility and explore relationships

among study variables for future replication. The sample size was small. The participants were drawn from a single site disallowing analysis of variation by level of restriction. There are also technical aspects of EMA study design that have a strong bearing on the findings, such as sampling interval and duration. Three samplings per day, as in our case, would imply that phenomena lasting less than the sampling interval (24/3 = 8 h) may not be captured in sufficient detail in our network models. More frequent sampling may reveal greater detail; however, it may also easily become intrusive and burdensome to older participants. It should be noted that objective measures of loneliness using UCLA-3 were available at three distinct checkpoints, whereas the subjective measures were a part of EMA sampling—this time gap may have attenuated the correlation between EMA and scale-based loneliness. In understanding the influence of loneliness on behavior, it is important to account for concurrent depressive symptoms. This sample had very low levels of depression on average, and so these results may not generalize to samples with greater variation in depressive symptoms. Lastly, the study was performed during the early period of COVID-19 pandemic (between 5/25/2020 and 8/16/2020) and before the FDA approval of first vaccine, the social-distancing rules may have altered the living conditions and limited the activities of the cohorts.

In conclusion, EMA-based network modeling appears to be a useful tool for assessing momentary loneliness in older adults. Given issues with early adherence that later resolved, follow-up with participants at the outset of EMA survey protocols may support adherence. Our study points to potentially important nuances to understanding the connection between acute loneliness and behavior, and how policy and environmental influences may impact response to short-term loneliness. Future study should examine how momentary loneliness, day-to-day behavior and affective experience converge to contribute to chronic loneliness, such as in a measurement burst design (49). This technique uses *bursts* of frequently repeated assessments in a short period of time, spanning a few days or weeks. Such burst measurements are repeated longitudinally over a longer interval (after a few months or a year), capturing not only individual differences, but also the short-term variability in measured variables and long-term trends, *vis a vis* chronic loneliness and its impact on health over the course.

DATA AVAILABILITY STATEMENT

De-identified data supporting the conclusions of this article will be made available by the authors to qualified investigators. Further queries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by UCSD Human Research Protections Program. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

CAD, VDB, and EEL contributed to conception and design of the study. RD organized the database and supported the data collection activity. VDB performed the network analysis and CAD oversaw the study. VDB wrote the first draft of the manuscript. CAD, VDB, H-CK, EMP, EEL, RD, and DVJ edited and contributed to the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work was supported by IBM Research AI through the AI Horizons Network. This study was supported, in part, by the National Institute of Mental Health [NIMH T32 Geriatric Mental Health Program MH019934 (DVJ), K23 grant-MH119375-01 to EEL, by the Stein Institute for Research on Aging (DVJ) and Veterans Affairs Healthcare System.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2022.814179/full#supplementary-material

REFERENCES

1. Jeste DV, Lee EE. loneliness vs. wisdom amid two concurrent pandemics of Loneliness. *Am J Geriatr Psychiatry*. (2020) 28:1245–7. doi: 10.1016/j.jagp.2020.08.017
2. Jeste DV, Lee EE, Cacioppo S. Battling the modern behavioral epidemic of loneliness: suggestions for research and interventions. *JAMA Psychiatry*. (2020) 77:553–4. doi: 10.1001/jamapsychiatry.2020.0027
3. Donovan NJ, Blazer D. Social isolation and loneliness in older adults: review and commentary of a National Academies Report. *Am J Geriatr Psychiatry*. (2020) 28:1233–44. doi: 10.1016/j.jagp.2020.08.005
4. Vahia IV, Jeste DV, Reynolds CF. Older adults and the mental health effects of COVID-19. *JAMA*. (2020) 324:2253–4. doi: 10.1001/jama.2020.21753
5. Mouratidis K. How COVID-19 reshaped quality of life in cities: a synthesis and implications for urban planning. *Land Use Policy*. (2021) 111:105772. doi: 10.1016/j.landusepol.2021.105772
6. Hwang T-J, Rabheru K, Peisah C, Reichman W, Ikeda M. Loneliness and social isolation during the COVID-19 pandemic. *Int Psychogeriatr*. (2020) 32:1217–20. doi: 10.1017/S1041610220000988
7. Van Tilburg TG, Steinmetz S, Stolte E, van der Roest H, de Vries DH. Loneliness and mental health during the COVID-19 pandemic: a study among Dutch older adults. *J Gerontol Ser B*. (2021) 76:e249–55. doi: 10.1093/geronb/gbaa111
8. Kotwal AA, Holt-Lunstad J, Newmark RL, Cenzer I, Smith AK, Covinsky KE, et al. Social isolation and loneliness among San Francisco Bay Area older adults during the COVID-19 shelter-in-place orders. *J Am Geriatr Soc*. (2021) 69:20–9. doi: 10.1111/jgs.16865
9. Cacioppo J, Patrick W. *Loneliness: Human Nature and the Need for Social Connection*. New York, NY: W.W. Norton and Company, Inc. (2008).
10. Cacioppo JT, Hawkley LC. Perceived social isolation and cognition. *Trends Cogn Sci*. (2009) 13:447–54. doi: 10.1016/j.tics.2009.06.005

11. Coyle CE, Dugan E. Social isolation, loneliness and health among older adults. *J Aging Health.* (2012) 24:1346–63. doi: 10.1177/0898264312460275
12. Perissinotto CM, Covinsky KE. *Living Alone, Socially Isolated or Lonely—What Are We Measuring?* Cham: Springer (2014). pp. 1429–31.
13. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. *Perspect Psychol Sci.* (2015) 10:227–37. doi: 10.1177/1745691614568352
14. Tucker JS, Health-related social control within older adults' relationships. *J Gerontol Ser B Psychol Sci Soc Sci.* (2002) 57:P387–395. doi: 10.1093/geronb/57.5.P387
15. Kobayashi LC, Steptoe A. Social isolation, loneliness, and health behaviors at older ages: longitudinal cohort study. *Ann Behav Med.* (2018) 52:582–93. doi: 10.1093/abm/kax033
16. Schrempft S, Jackowska M, Hamer M, Steptoe A. Associations between social isolation, loneliness, and objective physical activity in older men and women. *BMC Public Health.* (2019) 19:1–10. doi: 10.1186/s12889-019-6424-y
17. Badal VD, Parrish EM, Holden JL, Depp CA, Granholm E. Dynamic contextual influences on social motivation and behavior in schizophrenia: a case-control network analysis. *Npj Schizophrenia.* (2021) 7:62. doi: 10.1038/s41537-021-00189-6
18. Durlauf S, Blume L. *Macroeconometrics and Time Series Analysis*. Cham: Springer (2016).
19. Williams MT, Lewthwaite H, Fraysse F, Gajewska A, Ignatavicius J, Ferrar K. Compliance with mobile ecological momentary assessment of self-reported health-related behaviors and psychological constructs in adults: systematic review and meta-analysis. *J Med Internet Res.* (2021) 23:e17023. doi: 10.2196/17023
20. Manini TM, Mendoza T, Battula M, Davoudi A, Kheirkhahan M, Young ME, et al. Perception of older adults toward smartwatch technology for assessing pain and related patient-reported outcomes: pilot study. *JMIR mHealth uHealth.* (2019) 7:e10044. doi: 10.2196/10044
21. Rullier L, Atzeni T, Husky M, Bouisson J, Dartigues JF, Swendsen J, et al. Daily life functioning of community-dwelling elderly couples: an investigation of the feasibility and validity of Ecological Momentary Assessment. *Int J Methods Psychiatr Res.* (2014) 23:208–16. doi: 10.1002/mpr.1425
22. Laborde CR, Cenko E, Mardini MT, Nerella S, Kheirkhahan M, Ranka S, et al. Usability, and compliance with the use of smartwatches for ecological momentary assessment of knee osteoarthritis symptoms in older adults: usability study. *JMIR Aging.* (2021) 4:e24553. doi: 10.2196/24553
23. Mattos MK, Sereika SM, Beach SR, Kim H, Klunk WE, Knox M, et al. Research use of ecological momentary assessment for adverse event monitoring following amyloid-β results disclosure. *J Alzheimers Dis.* (2019) 71:1071–9. doi: 10.3233/JAD-190091
24. Kratz AL, Murphy SL, Braley TJ. Pain, fatigue, and cognitive symptoms are temporally associated within but not across days in multiple sclerosis. *Arch Phys Med Rehabil.* (2017) 98:2151–9. doi: 10.1016/j.apmr.2017.07.003
25. Mardini MT, Nerella S, Kheirkhahan M, Ranka S, Fillingim RB, Hu Y, et al. The temporal relationship between ecological pain and life-space mobility in older adults with knee osteoarthritis: a smartwatch-based demonstration study. *JMIR mHealth uHealth.* (2021) 9:e19609. doi: 10.2196/19609
26. Potts C, Bond R, Ryan A, Mulvenna M, McCauley C, Laird E, et al. Ecological momentary assessment within a digital health intervention for reminiscence in persons with dementia and caregivers: user engagement study. *JMIR mHealth uHealth.* (2020) 8:e17120. doi: 10.2196/17120
27. Maher JP, Dunton GF. Editor's Choice: dual-process model of older adults' sedentary behavior: an ecological momentary assessment study. *Psychol Health.* (2020) 35:519–37. doi: 10.1080/08870446.2019.1666984
28. Fritz H, Tarraf W, Saleh DJ, Cutchin MP. Using a smartphone-based ecological momentary assessment protocol with community dwelling older African Americans. *J Gerontol Ser B Psychol Sci Soc Sci.* (2017) 72:876–87. doi: 10.1093/geronb/gbw166
29. Compernolle EL, Finch LE, Hawkley LC, Cagney KA. Momentary loneliness among older adults: contextual differences and their moderation by gender and race/ethnicity. *Soc Sci Med.* (2021) 285:114307. doi: 10.1016/j.socscimed.2021.114307
30. Haucke M, Liu S, Heinzel S. The persistence of the impact of COVID-19–related distress, mood inertia, and loneliness on mental health during a postlockdown period in germany: an ecological momentary assessment study. *JMIR Mental Health.* (2021) 8:e29419. doi: 10.2196/29419
31. Daly JR, Depp C, Graham SA, Jeste DV, Kim H-C, Lee EE, et al. Health impacts of the stay-at-home order on community-dwelling older adults and how technologies may help: focus group study. *JMIR Aging.* (2021) 4:e25779. doi: 10.2196/25779
32. Seeman TE, Berkman LF, Blazer D, Rowe JW. Social ties and support and neuroendocrine function: the MacArthur studies of successful aging. *Ann Behav Med.* (1994) 16:95–106.
33. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* (2001) 16:606–13. doi: 10.1046/j.1525-1497.2001.016009606.x
34. Derogatis LR, Melisaratos N. The brief symptom inventory: an introductory report. *Psychol Med.* (1983) 13:595–605. doi: 10.1017/S0033291700048017
35. Russell DW, UCLA Loneliness Scale (Version 3): reliability, validity, factor structure. *J Pers Assess.* (1996) 66:20–40. doi: 10.1207/s15327752jpa6601_2
36. Smith S. *Measurement of Loneliness Among Clients Representing Four Stages of Cancer: An Exploratory Study*. Wright-Patterson AFB, OH: Air Force Institute of Technology (1985).
37. Deckx L, M. van den Akker, Buntinx F. Risk factors for loneliness in patients with cancer: a systematic literature review and meta-analysis. *Eur J Oncol Nurs.* (2014) 18:466–77. doi: 10.1016/j.ejon.2014.05.002
38. Derogatis LR. *BSI Brief Symptom Inventory. Administration, Scoring, and Procedures Manual*. New York, NY: Springer (1993).
39. Runge J, Nowack P, Kretschmer M, Flaxman S, Sejdinovic D. Detecting and quantifying causal associations in large nonlinear time series datasets. *Sci Adv.* (2019) 5:eaau4996. doi: 10.1126/sciadv.aau4996
40. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodological).* (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
41. Badal VD, Graham SA, Depp CA, Shinkawa K, Yamada Y, Palinkas LA, et al. Prediction of loneliness in older adults using natural language processing: exploring sex differences in speech. *Am J Geriatr Psychiatry.* (2020) 29:853–66. doi: 10.1016/j.jagp.2020.09.009
42. Weiss RS. *Loneliness: The Experience of Emotional and Social Isolation*. Cambridge, MA: The MIT Press (1973).
43. Fardghassemi S, Joffe H. Young adults' experience of loneliness in london's most deprived areas. *Front Psychol.* (2021) 12:660791. doi: 10.3389/fpsyg.2021.660791
44. Cacioppo JT, Cacioppo S, Boomsma DI. Evolutionary mechanisms for loneliness. *Cogn Emot.* (2014) 28:3–21. doi: 10.1080/02699931.2013.837379
45. Raut NB, Singh S, Subramanyam AA, Pinto C, Kamath RM, Shanker S. Study of loneliness, depression and coping mechanisms in elderly. *J Geriatr Mental Health.* (2014) 1:20. doi: 10.4103/2348-9995.141920
46. Hevel DJ, Dunton GF, Maher JP. Acute bidirectional relations between affect, physical feeling states, and activity-related behaviors among older adults: an ecological momentary assessment study. *Ann Behav Med.* (2021) 55:41–54. doi: 10.1093/abm/kaaa027
47. Clair R, Gordon M, Kroon M, Reilly C. The effects of social isolation on well-being and life satisfaction during pandemic. *Human Soc Sci Commun.* (2021) 8:1–6. doi: 10.1057/s41599-021-00710-3
48. Fullana MA, Hidalgo-Mazzei D, Vieta E, Radua J. Coping behaviors associated with decreased anxiety and depressive symptoms during the COVID-19 pandemic and lockdown. *J Affect Disord.* (2020) 275:80–1. doi: 10.1016/j.jad.2020.06.027
49. Stawski RS, MacDonald SW, Sliwinski MJ. Measurement burst design. In Whitbourne SK, editor. *The Encyclopedia of Adulthood and Aging*. Hoboken, NJ: John Wiley & Sons (2015). pp. 1–5.

Author Disclaimer: The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Conflict of Interest: H-CK is an employee of IBM.

The remaining 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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 10 February 2022
doi: 10.3389/fdgth.2022.840169

Generational Perspectives on Technology's Role in Mental Health Care: A Survey of Adults With Lived Mental Health Experience

Molly Woerner[1], *Nichole Sams*[1,2], *Cristian Rivera Nales*[1,3], *Tara Gorstein*[1], *Morgan Johnson*[2], *Brittany A. Mosser*[1,2] *and Patricia A. Areán*[1,2]*

[1] *Conducting Research to Enhance Assessment and Treatment Through Innovation in Mental Health (CREATIV) Lab, Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States,* [2] *Advanced Laboratories for Accelerating the Reach and Impact of Treatments for Youth and Adults With Mental Illness (ALACRITY Center), Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States,* [3] *Adolescent Adversity and Depression Intervention (AADI) Lab, Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States*

OPEN ACCESS

Edited by:
Amit Baumel, University of Haifa, Israel

Reviewed by:
Henry Onyeaka, Massachusetts General Hospital and Harvard Medical School, United States
Emily G. Lattie, Northwestern University, United States

***Correspondence:**
Patricia A. Areán
parean@uw.edu

Specialty section:
This article was submitted to Human Factors and Digital Health, a section of the journal Frontiers in Digital Health

Received: *20 December 2021*
Accepted: *18 January 2022*
Published: *10 February 2022*

Citation:
Woerner M, Sams N, Rivera Nales C, Gorstein T, Johnson M, Mosser BA and Areán PA (2022) Generational Perspectives on Technology's Role in Mental Health Care: A Survey of Adults With Lived Mental Health Experience. Front. Digit. Health 4:840169. doi: 10.3389/fdgth.2022.840169

Introduction: Personal technology (e.g., smartphones, wearable health devices) has been leveraged extensively for mental health purposes, with upwards of 20,000 mobile applications on the market today and has been considered an important implementation strategy to overcome barriers many people face in accessing mental health care. The main question yet to be addressed is the role consumers feel technology should play in their care. One underserved demographic often ignored in this discussion are people over the age of 60. The population of adults 60 and older is predicted to double by 2,050 signaling a need to address how older adults view technology for their mental health care.

Objective: The objective of this study is to better understand why digital mental health tools are not as broadly adopted as predicted, what role people with lived mental health experience feel technology should play in their care and how those results compare across age groups.

Method: In a mixed-methods approach, we analyzed results from a one-time cross-sectional survey that included 998 adults aged 18–83 with lived experience of mental health concerns recruited from Prolific, an online research platform. We surveyed participant's use of technology including their perspectives on using technology in conjunction with their mental health care. We asked participants about their previous use of digital mental health tools, their treatment preferences for mental health care, and the role technology should play in their mental health care.

Results: Across all age groups, respondents had favorable views of using digital mental health for managing mental health care. However, older adults rated their acceptability of digital mental health tools lower than middle-aged and younger adults. When asked what role technology should play in mental health care in an open-ended response, most participants responded that technology should play a complementary role in mental health care (723/954, 75.8%).

Conclusion: Digital mental health is seen as a valuable care management tool across all age groups, but preferences for its role in care remain largely administrative and supportive. Future development of digital mental health should reflect these preferences.

Keywords: older adults, digital mental health, lived experience, technology, mental health

INTRODUCTION

There is no question that access to mental health services in the US is difficult; 58% of people with mental illness never receive treatment (1). Decades of research has shown that poor access is due to several factors, from the stigma people experience when they ask for help to the fact that there are severe mental health provider shortage areas particularly in rural areas or urban poverty areas (2). Even in high provider density areas, older adults remain sorely underserved because too few mental health providers are credentialed in geriatric mental health (3), and few providers accept Medicaid or Medicare (4). For the US, this shortage will result in significant societal costs, particularly in the case of older adults; it is projected by the year 2040, the number of individuals 85 and older will increase by 129% (5–7). Approximately 20% of older adults will experience a mental health issue, with the most common diagnoses being Major Depressive Disorder, Bipolar Disorder and Schizophrenia (8). Additionally, 17.9% of all suicide deaths are of older adults (9). Access to mental health care in this demographic is important to address.

Since the advent of COVID-19, attention has turned to the potential opportunity digital mental health (apps, tele-health, and message-based care) has in overcoming access problems and mental health disparities. According to a recent Banbury Report on digital mental health (10) such technologies may address the mental health needs of US citizens; these tools are evidence-based and prolific, with over 20,000 apps and 40 companies offering these services (11). Further, digital mental health is cost-effective, scalable, lowers the cost of hospitalizations, lowers burden on providers, increases access to care in rural settings where broadband is available, results in lower wait times, and can personalize treatment (12–14).

Although digital mental health tools are an effective and efficient mental health service delivery method, very few people with mental health needs use these tools. Even during COVID-19, only 16% of essential workers and unemployed individuals in need of mental health care ever used a digital mental health tool, with 86% of those with mental health distress during COVID-19 indicating that they did not believe such tools would help (15). Other studies have found that large proportions of the general population have similar quality concerns in addition to security and safety concerns about digital mental health care (16). Older adults are less likely than younger adults to use digital mental health, although this trend may be changing due the increase in general technology use by older adults during COVID-19 (17). For older adults, utilization challenges may be driven by challenges they experience when using mobile technology, such as trouble understanding application interfaces, dexterity issues, preference for human interaction, and cognitive issues (18–20). Taken together these studies suggest that underutilization of digital mental health tools may be in large part due to its acceptability as a mental health care option, and, for older adults, usability challenges and preference for care delivery.

To date the field has not included representative user's perspectives on the role technology should play in mental health recovery, and relative preferences among digital options and in-person care, particularly among older populations. According to Human Centered Design Theory (21, 22), the needs and values of intended audiences or users, in this case people with lived experience in mental health, must be central to determining the role technology plays in solving a problem, its design, and its purpose. The purpose of this study is to better understand why digital mental health tools are not as broadly adopted as predicted, what role people with lived mental health experience feel technology should play in their care and in particular the accessibility challenges of digital mental health by age group. Our specific aims are: (1) to ask people with lived experience with mental health concerns the role technology should play in their care, (2) their preferences among different digital mental health care options, and (3) compare responses by age group.

METHODS

Study Design

This was a cross-sectional, remote survey of digital mental health tool use.

Participants and Procedures

One thousand thirty-two research participants were recruited using Prolific (www.prolific.co), an online research platform that allows researchers to screen, recruit, enroll, and pay participants. Participants in this study were 18 years or older, English speaking, living in the United States.

All participants endorsed having experienced a mental health condition or treatment on a pre-screening survey. The questions of this pre-screening survey included (1) "Have you ever received help in the past for stress or mental health issues? The help may have been psychotherapy, counseling, and/or medication for depression, anxiety, stress management or other mental health issue," (2) "Are you currently receiving help (psychotherapy, counseling, medication) for stress or mental health issues?," and (3) "Have you ever experienced a mental health condition, such as depression, anxiety, or psychosis?" Participants were excluded if they did not endorse "yes" to at least one of these three questions.

After recruited participants completed the pre-screening survey using Research Electronic Data Capture (REDCap) and met eligibility criteria, they were given a link to complete

the full survey on REDCap (23, 24). Recruitment was stratified by US census racial categories to include a racially representative sample. We oversampled under-represented minority populations in an attempt to obtain a sample consistent with US Census representation. We also oversampled older adults to ensure a large enough older adult sample to accurately report on age differences. Participants were paid $0.50 for completing the pre-screening survey and $5.00 for completing the full survey.

Ethics Statement

The study was approved by the University of Washington's Institutional Review Board (STUDY00014041; FWA #00006878).

Data Collection

Participants completed the survey on REDCap. The survey took an average 20 min to complete and included sections on demographics, health and wellbeing, use of technology, use of apps, social media and the internet, treatment preferences, and a "Build Your Own Mental Health Experience" section. See **Supplementary Material** for full survey questions. Although participants provided answers to several questions about comfort with technology, use of technology for physical health and preferences within mental health tools, we report here only on data concerning current use of mental health technology, preferences for different types of digital mental health technology, and participant's perspectives on the role that technology should play in mental health care.

Protection Against Malicious Actors

In interest of data quality, several procedures were enacted to prevent "bad actors" (individuals or entities who participate in bad faith to accumulate monetary incentives) from participating in the study and joining the final study sample. Prolific research platform enacts measures to vet participants such as verification of email address, phone number (ensuring it correlates with their country of residence), photo identification, and PayPal address (25). Technical measures include restricting signups based on IP address and ISP, requiring a unique non-VOIP phone number, and analysis of data to review unusual patterns (26). A PayPal or Circle account is required to be paid and both have procedures to prevent duplicate accounts. We included two attention checks in our survey that required participants to read questions fully and answer them both accurately to be included in the final sample. By including open-ended questions in our study, we were able to screen any non-coherent answers and address possible malicious actors like bots (27). Both specific platform quality checks and our internal quality checks assure our final sample is comprised of quality "good actors."

Demographics

Participants were asked to provide their age, zip code, description of community size, employment level, number of people living in home, if someone identifies as an underrepresented person, US Census racial categories, ethnicity, gender identity, and financial comfort. All participants were stratified in one of three age ranges: (1) younger adults (YA: 18–34 years, mean = 24.6 years), (2) middle aged adults (MA: 35–59 years, mean = 50.5 years), and (3) older adults (OA: 60+ years, mean = 66.0 years). Although young, middle, and older adulthood is not clearly defined by age we determined age ranges based on common developmental and social periods of adulthood including the transitionary periods of family roles, careers, and changing health (28). Younger adulthood comprises of pursing new careers, social relationships, and identity formation (29). Middle age presents new life circumstances such as new family demands, career seniority, and new health challenges (30). The older adult population is characterized as a period with changes in cognition, physical health, fewer work and family responsibilities, and increased focus on meaningful aging and experiences toward the end of life (31). We asked participants to describe their current health and general wellbeing to characterize the sample on current physical and emotional functioning. To assess current levels of emotional distress, we asked participants to complete the 9-item Patient Health Questionnaire (PHQ-9) (32) which asks about the frequency and severity of depressive symptoms over the previous 2 weeks, and the 7-item General Anxiety Disorder (GAD-7) (33) which asks about the frequency and severity of generalized anxiety symptoms over the previous 2 weeks.

Survey Questions

Use of Technology for Mental Health

We used questions from a previous survey of preferences for digital mental health among COVID-19 essential workers and those unemployed during the pandemic (15), modified slightly for this study. The survey asked participants if they have ever used technology to manage mental health problems and perspectives on different types of digital mental health tools, specifically apps, message-based care and telehealth. The survey also asked participant preferences for these different mental health technologies. See **Supplementary Material** for full survey.

Role of Technology in Mental Health

Participants were asked a single open-ended question about their thoughts about the role of technology in mental health care. Specifically, they were asked "In your own words, what role should technology serve in mental health care?"

Statistical Analysis Plan

The original sample included 1,032 participants. Data cleaning removed 34 individuals from the original sample because they either (1) had a duplicate Prolific ID ($n = 2$), (2) did not move past consent ($n = 1$), or (3) did not pass both attention checks ($n = 31$). The final sample included 998 participants.

For each aim, we analyzed differences across age groups. Age was categorized into three groups (1) younger adults (YA: 18–34 years), (2) middle aged adults (MA: 35–59 years), and (3) older adults (OA: 60+ years). To examine differences between age groups, we used chi-square tests for categorical variables, Kruskal–Wallis tests for ordinal variables, and one-way analysis of variance (ANOVA) for continuous variables. When statistically significant differences across groups were found for variables with multiple discrete categories, we conducted *post-hoc*

probes using standardized residuals to identify which categories were responsible for the significant difference.

To control the prevalence of false positives due to multiple testing, we applied the Benjamini-Hochberg (B-H) procedure with the false discovery rate set to 10% to 45 statistical tests (34, 35) All statistical analyses were performed with SAS version 9.4.

Qualitative Analysis

Responses to the question, "In your own words, what role should technology play in your mental health care?" was coded using thematic, content analysis (36). Codes and their definitions were developed by two independent reviewers (M.W. and N.S.) who identified emergent themes in the data. Themes were then codified and confirmed by both reviewers. In the event reviewers did not agree on themes, a third reviewer (P.A.A.) was available to resolve the discrepancy. However, in this project, no third review was needed. We tallied the frequency that a theme was endorsed by different participants to rank themes by prominence in the data. Representative quotes were selected for each theme and are reported in the results section.

RESULTS

Sample Characteristics

Table 1 shows frequencies, means, and standard deviations of participant demographics and clinical characteristics. Most of the sample identified as female (651/995, 65.4%), were employed (610/981, 62.2%) and did not self-identify as an underrepresented population based on skin color, heritage, socio-economic status, gender-identity, sexual orientation, or other identified characteristics (654/989, 66.1%). The sample was predominately white (750/994, 75.5%) followed by Black or African American (74/994, 7.4%), multi-racial (69/994, 6.9%), and Hispanic/Latinx (55/994, 5.5%).

Older adults were more likely to have received help in the past for stress or mental health issues ($\chi^2_2 = 11.3$, $p = 0.0035$) while no differences emerged for current mental health treatment use or past lived experience across groups.

Overall, the sample's average score on the PHQ-9 and GAD-7 indicated mild levels of depression (mean 7.7, SD 5.3) and anxiety (mean 6.6, SD 5.8) severity. Older adults demonstrated the lowest levels of depression (mean 5.6, SD 4.2) and anxiety (mean 4.2, SD 4.5) severity.

Survey Results

App, Message Based Care, and Telehealth Usage

Table 2 presents app, message-based care, and telehealth usage across age groups.

Mental Health Apps

When asked *Have you considered using an app for your mental health?* one in three participants (334/988, 33.8%) said they had considered using an app for this purpose. Compared to younger and middle-aged adults, significantly fewer older adults considered an app for their mental health ($\chi^2_2 = 91.6, p < 0.01$).

Among those who have considered using an app for their mental health, about a third (115/334, 34.4%) indicated that they have downloaded a mental health app, with no differences by age found.

Message-Based Care, With and Without Video Conferencing

Although older adults were similar to younger and middle-aged adults in their general view of mental health apps, they were less likely to rank them (1) as playing an important role in managing mental health ($\chi^2_2 = 11.1$, $p < 0.01$), (2) to manage mental health ($\chi^2_2 = 15.7$, $p < 0.01$), (3) as an effective intervention for managing mental health conditions ($\chi^2_2 = 14.6, p < 0.01$), and (4) were less willing to use these tools ($\chi^2_2 = 25.2, p < 0.01$).

When individuals were asked their preference between message-based care with or without video-conferencing, all age groups ranked supplementing message-based care with video-conferencing more favorably than message-based care without video-conferencing.

Tele-Mental Health

Older adults were less likely to positively rank telehealth care than younger and middle age adults in (1) playing an important role in managing mental health ($\chi^2_2 = 11.8$, $p < 0.01$), (2) used to manage mental health ($\chi^2_2 = 15.6$, $p < 0.01$), (3) effective intervention for managing mental health conditions ($\chi^2_2 = 13.7$, $p < 0.01$), and (4) willingness to use ($\chi^2_2 = 17.9$, $p < 0.01$). However, all age groups viewed this mode of care favorably.

Preferences for In-person, Tele-Mental Health, Message-Based Care and Mental Health Apps

Table 3 displays treatment preferences across age groups.

Participant's preferences for the mode of treatment delivery differed significantly by age ($\chi^2_6 = 25.4$, $p < 0.01$). While all three group's leading choice of treatment was in-person therapy, younger and older adults were more likely to rank this as a leading preference compared to middle aged adults, who were more likely to prefer tele-health and mobile mental health apps.

Concerns Over Mental Health Apps, Message-Based Care and Tele-Mental Health

All age groups indicated the greatest concern with message-based care and mobile mental health apps. Older adults in particular had greater concerns for mobile mental health apps ($\chi^2_2 = 13.5$, $p < 0.01$) and telehealth ($\chi^2_2 = 13.9$, $p < 0.01$), than younger and middle age adults.

The Role of Technology in Mental Health Care

Participants were asked to provide open-ended responses to the question *"In your own words, what role should technology play in your mental health care?"*

Nearly all ($N = 954$) participants responded to this question. Our qualitative analysis of responses found that a majority of the total sample (56%) felt that technology should play a complementary role to traditional mental health care, with 19.8% recommending technology play a major role in mental health care. This finding was shared across age groups, with one exception: older adults were more likely to indicate that technology should play no role in mental health care (10.8%), whereas only 4.9% of younger adults and 7.5% of middle-aged

TABLE 1 | Demographics stratified by age (categorized).

	Age categorized			
	Younger adults (18–34 years, mean = 24.6) (*n* = 290)	**Middle aged adults (35–59 years, mean = 50.5) (*n* = 406)**	**Older adults (60 + years, mean = 66.0) (*n* = 302)**	**Total (*N* = 998)**
Employment				
Unemployed	87 (30.5%)	75 (18.8%)	141 (47.3%)	303 (30.9%)
Unpaid work at home (e.g., primary unpaid caregiver of family member)	10 (3.5%)	33 (8.3%)	6 (2.0%)	49 (5.0%)
Unpaid work out of the home (e.g., volunteerism)	6 (2.1%)	3 (0.8%)	10 (3.4%)	19 (1.9%)
Part time paid work outside the house	63 (22.1%)	29 (7.3%)	21 (7.0%)	113 (11.5%)
Part time paid work at home	16 (5.6%)	42 (10.6%)	50 (16.8%)	108 (11.0%)
Full time paid work outside the house	80 (28.1%)	139 (34.9%)	40 (13.4%)	259 (26.4%)
Full time paid work at home	23 (8.1%)	77 (19.3%)	30 (10.1%)	130 (13.3%)
Missing	5	8	4	17
Household members				
1	69 (27.3%)	116 (34.8%)	152 (69.4%)	337 (41.9%)
2	77 (30.4%)	92 (27.6%)	49 (22.4%)	218 (27.1%)
3	66 (26.1%)	69 (20.7%)	12 (5.5%)	147 (18.3%)
4	24 (9.5%)	40 (12.0%)	5 (2.3%)	69 (8.6%)
5+	17 (6.7%)	16 (4.8%)	1 (0.5%)	34 (4.2%)
Missing	37	73	83	193
Identify as an underrepresented population based on skin color, heritage, socio-economic status, gender-identity, sexual orientation, or other aspects of your identity.				
Yes	129 (44.6%)	111 (27.7%)	38 (12.7%)	278 (28.1%)
No	140 (48.4%)	261 (65.1%)	253 (84.6%)	654 (66.1%)
Not sure	20 (6.9%)	29 (7.2%)	8 (2.7%)	57 (5.8%)
Missing	1	5	3	9
Race				
Multi-racial	29 (10.0%)	33 (8.1%)	7 (2.3%)	69 (6.9%)
American Indian or Alaska native or indigenous	0 (0.0%)	4 (1.0%)	1 (0.3%)	5 (0.5%)
Asian	16 (5.5%)	20 (4.9%)	1 (0.3%)	37 (3.7%)
Black or African American	28 (9.7%)	38 (9.4%)	8 (2.7%)	74 (7.4%)
Hispanic/Latinx	30 (10.4%)	22 (5.4%)	3 (1.0%)	55 (5.5%)
Middle Eastern or North African	3 (1.0%)	1 (0.2%)	0 (0.0%)	4 (0.4%)
White	183 (63.3%)	287 (70.9%)	280 (93.3%)	750 (75.5%)
Missing	1	1	2	4
Ethnicity				
Hispanic/Latinx	47 (16.6%)	52 (12.9%)	4 (1.3%)	103 (10.4%)
Non-Hispanic/Latinx	236 (83.4%)	351 (87.1%)	296 (98.7%)	883 (89.6%)
Missing	7	3	2	12
Gender identity				
Female	203 (70.0%)	256 (63.4%)	192 (63.8%)	651 (65.4%)
Male	66 (22.8%)	141 (34.9%)	108 (35.9%)	315 (31.7%)
Transgender, Non-binary, or Gender-nonconforming	21 (7.2%)	7 (1.7%)	1 (0.3%)	29 (2.9%)
Missing	0	2	1	3
Financial stability, *n* (%)				
Can't make ends meet	35 (12.3%)	62 (15.5%)	35 (11.7%)	132 (13.4%)
Have just enough to get by	162 (56.8%)	164 (41.0%)	137 (45.7%)	463 (47.0%)
Are comfortable	88 (30.9%)	174 (43.5%)	128 (42.7%)	390 (39.6%)

(Continued)

TABLE 1 | Continued

	Age categorized			Total ($N = 998$)
	Younger adults (18–34 years, mean = 24.6) ($n = 290$)	**Middle aged adults (35–59 years, mean = 50.5) ($n = 406$)**	**Older adults (60 + years, mean = 66.0) ($n = 302$)**	
Missing	5	6	2	13
Home location, *n* (%)				
Large city	92 (31.9%)	92 (22.7%)	50 (16.6%)	234 (23.5%)
Suburb near a large city	96 (33.3%)	150 (36.9%)	106 (35.1%)	352 (35.3%)
Small city or town	72 (25.0%)	109 (26.8%)	86 (28.5%)	267 (26.8%)
Rural area	28 (9.7%)	55 (13.5%)	60 (19.9%)	143 (14.4%)
Missing	2	0	0	2
Clinical characteristics				
PHQ-9 total score				
N	290	406	302	998
Mean (SD)	10.1 (5.5)	7.5 (5.3)	5.6 (4.2)	7.7 (5.3)
GAD-7 total score				
N	290	406	302	998
Mean (SD)	9.6 (5.9)	6.2 (5.6)	4.2 (4.5)	6.6 (5.8)

adults felt technology should play no role in mental health care. See **Table 4**.

Five major themes emerged regarding the specific role that technology should play in mental health care: paired with a mental health professional, for communication purposes, symptom monitoring, improving access to care, and specific recommendations for technology in mental health. Nearly a third of respondents (29.3%) indicated that technology should be paired with a mental health professional, with 33.1% of older adults endorsing this, compared to 25.5% of younger adults and 29.1% of middle-aged adults endorsing this theme. One older adult participant responded, "I think technology can be very useful with my mental health care, however, I still want that human interaction. I feel the information gathered would be a great tool for a health professional in aiding in treatment." Similarly, technology as a means of addressing barriers to access (e.g., expanded hours, ready documentation about illness, increased quality of care) was endorsed by 31.1% of the sample, although fewer older adults (25.7%) endorsed this theme than the other age groups (see **Table 4**). One participant wrote, "Make it easier to access care immediately or as needed. Remove barriers or reduce such as cost, time, distance." Approximately 15% of the total sample felt technology should be used for communication and administrative purposes, with more older adults endorsing this theme (18.9%) than other age groups. One older adult responded, "For me, scheduling appointments, alerting for appointments, initial intake data to get set up with a counselor or clinician." Symptom monitoring was only endorsed in 9.3% of the sample, with minimal differences between age groups. A participant responded, "I think technology can serve as a watchdog and an alert for changes in mental health or like a warning signal if difficulties arise that I may not be able to handle." See **Table 4**.

Finally, 26% of the sample who answered this question provided detailed recommendations for technology use. These recommendations were to use technology to promote and reinforce healthy behaviors, utilize specific applications to help mental health, and for disorders that would specifically benefit from technology, such anxiety. One participant suggested, "I am impressed by Woebot and hope that with AI [artificial intelligence] even better mental health apps become available because I would use them every day."

DISCUSSION

To our knowledge, this is the first study to directly ask potential end-users of digital mental health tools the role they feel technology should serve in mental health care. This is also one of the first studies to intentionally include the voice of older adult populations in the discussion of such technology in their mental health care, and to compare older adult responses to those of younger adults. A major finding of this study is that with a few exceptions, younger, middle-age and older adults have similar perspectives on the use of digital mental health tools, and in some circumstances, older and younger adults are more aligned in their view of such tools than would be expected, given the vastly different presence of technology in their lives. Overall, all generational groups felt digital mental health tools should serve a supportive role to traditionally delivered mental health care, with technology being used primarily for extended access to clinicians, a means of communication, and, to a lesser degree, symptom monitoring. There were a few generational differences, with older adults rating message-based care and telehealth as less acceptable than their younger counterparts, and older adults were less likely to consider using an app to manage their mental health care than the younger and

TABLE 2 | App, message based care, and telehealth usage stratified by age categorized.

	Age Categorized				
	Younger Adults (18–34 years) (*n* = 290)	**Middle Aged Adults (35–59 years) (*n* = 406)**	**Older Adults (60+ years) (*n* = 302)**	**Total (*N* = 998)**	***P*-value**
App usage					
Have you considered using an app for your mental health?					**<0.01[1]**
No	131 (46.0%)	274 (67.8%)	249 (83.3%)	654 (66.2%)	
Yes	154 (54.0%)	130 (32.2%)	50 (16.7%)	334 (33.8%)	
Missing	5	2	3	10	
Have you downloaded a mental health app?[a]					0.27[1]
No	94 (61.0%)	90 (69.2%)	35 (70.0%)	219 (65.6%)	
Yes	60 (39.0%)	40 (30.8%)	15 (30.0%)	115 (34.4%)	
Message based care usage					
Message-based care can play an important role in managing my mental health.					**<0.01[2]**
N	288	405	301	994	
Mean (SD)	4.3 (1.2)	4.2 (1.2)	4.0 (1.3)	4.2 (1.2)	
Median	5.0	4.0	4.0	4.0	
I would use message-based care to manage my mental health.					**<0.01[2]**
N	284	405	301	990	
Mean (SD)	4.2 (1.3)	4.1 (1.3)	3.8 (1.4)	4.0 (1.4)	
Median	4.5	4.0	4.0	4.0	
Message-based care can be an effective intervention for managing mental health conditions.					**<0.01[2]**
N	284	403	298	985	
Mean (SD)	4.4 (1.1)	4.2 (1.2)	4.0 (1.2)	4.2 (1.2)	
Median	4.0	4.0	4.0	4.0	
If science found that message-based care is effective for managing mental health conditions, I would use it.					**<0.01[2]**
N	284	404	299	987	
Mean (SD)	4.6 (1.2)	4.4 (1.3)	4.1 (1.3)	4.4 (1.3)	
Median	5.0	5.0	4.0	5.0	
I believe using only message-based care, without video-conferencing, can be effective in managing my mental health.					0.79[2]
N	282	400	299	981	
Mean (SD)	3.3 (1.3)	3.3 (1.4)	3.3 (1.4)	3.3 (1.4)	
Median	3.0	4.0	3.0	3.0	
I would prefer to use message-based care (without video-conferencing) rather than video-conferencing to manage my mental health.					0.46[2]
N	286	402	298	986	
Mean (SD)	3.2 (1.6)	3.3 (1.5)	3.2 (1.6)	3.2 (1.6)	
Median	3.0	3.0	3.0	3.0	
I would prefer to use message-based care with video-conferencing rather than message-based care without video-conferencing to manage my mental health.					**<0.01[2]**
N	284	401	300	985	
Mean (SD)	4.1 (1.5)	3.7 (1.5)	3.7 (1.5)	3.8 (1.5)	
Median	5.0	4.0	4.0	4.0	

(Continued)

TABLE 2 | Continued

	Age Categorized				
	Younger Adults (18–34 years) (n = 290)	**Middle Aged Adults (35–59 years) (n = 406)**	**Older Adults (60+ years) (n = 302)**	**Total (N = 998)**	***P*-value**
Telehealth usage					
Tele-health can play an important role in managing my mental health.					**<0.01[2]**
N	290	406	302	998	
Mean (SD)	4.7 (1.1)	4.7 (1.1)	4.5 (1.1)	4.6 (1.1)	
Median	5.0	5.0	5.0	5.0	
I would use tele-health to manage my mental health.					**<0.01[2]**
N	288	405	302	995	
Mean (SD)	4.5 (1.3)	4.5 (1.2)	4.2 (1.3)	4.4 (1.3)	
Median	5.0	5.0	4.0	5.0	
Tele-health can be an effective intervention for managing mental health conditions.					**<0.01[2]**
N	285	404	299	988	
Mean (SD)	4.7 (1.0)	4.6 (1.1)	4.3 (1.2)	4.6 (1.1)	
Median	5.0	5.0	5.0	5.0	
If science found that tele-health is effective for managing mental health conditions, I would use it.					**<0.01[2]**
N	284	405	301	990	
Mean (SD)	4.7 (1.2)	4.7 (1.2)	4.3 (1.3)	4.6 (1.2)	
Median	5.0	5.0	5.0	5.0	

[1]Chi-Square p-value; [2]Kruskal-Wallis p-value. ***Bold*** *indicates P-value < 0.05 and less than Benjamini-Hochberg critical value, considered to be statistically significant. [a]Among those who have considered using an app for their mental health. Note: scale range 1–6; higher score reflect agreement.*

middle-aged respondents. These results are consistent with the overall lower use of technology by older adults (37–40). We found that across all groups the leading preference for mental health treatment was in-person therapy which is consistent with prior research on DMH preferences of adults under 68 (41), although middle-aged adults were more likely to endorse using such tools compared to younger and older adults. In sum, our results suggest that few people in need turn to digital mental health tools because the tools available do not align with how people feel these tools should be used in mental health care.

Generational Lifestyles, Preferences, and Needs

Variation in responses among the three age groups in this study may be attributable to generational differences in life circumstances and work/social demands. Middle-aged adults are more likely to be faced with high productivity demands from work and home, given this is a stage in life where people are managing both young and old family members with growing families while juggling the demands of a career, financial pressures, and increasing health problems (30). This age group's greater acceptability of digital mental health-based solutions may reflect their need for efficient care that allows them to prioritize other life matters. Of particular interest is the finding that the majority (93.7%) of younger adults endorsed a preference for in-person therapy over technology-based care, more so than both middle aged and older adults. Although younger adult's technology use is high, our study suggests that younger adults see technology serving a specific role in their lives that does not include managing mental health issues.

For older adults, their barriers to technology, little understanding of DMH, and low-tech skills may explain part of why acceptability of DMH is lower than middle-aged adults and they prefer one-to-one therapy (19, 39). Additionally, previous research shows privacy concerns are significant hinderances to using technology (39, 42).

Suggestions for Successful Implementation of Digital Mental Health Across the Generations

Our results suggest that digital mental health interventions should be part of a multimodal package of care that combines aspects of in-person treatment with technology (43). Accounting for patient choice in their mental health care is associated with positive outcomes in therapeutic interventions and lower drop-out rates (44). Some older adults raised concerns about the "automation" of decisions and courses of action

TABLE 3 | Treatment preference stratified by age (categorized).

	Age Categorized				
	Younger Adults (18-34 years) (*n* = 290)	**Middle Aged Adults (35-59 years) (*n* = 406)**	**Older Adults (60+ years) (*n* = 302)**	**Total (*N* = 998)**	***P*-value**
Have you ever received help in the past for stress or mental health issues? The help may have been psychotherapy, counseling, and/or medication for depression, anxiety, stress management or other mental health issue					**<0.01[1]**
No	93 (32.4%)	104 (25.7%)	61 (20.3%)	258 (26.0%)	
Yes	194 (67.6%)	301 (74.3%)	240 (79.7%)	735 (74.0%)	
Missing	3	1	1	5	
Are you currently receiving help (psychotherapy, counseling, medication) for stress or mental health issue					0.85[1]
No	184 (63.9%)	257 (63.5%)	197 (65.4%)	638 (64.2%)	
Yes	104 (36.1%)	148 (36.5%)	104 (34.6%)	356 (35.8%)	
Missing	2	1	1	4	
Have you ever experienced a mental health condition, such as depression, anxiety, or psychosis?					0.17[1]
No	39 (13.8%)	73 (18.3%)	58 (19.3%)	170 (17.3%)	
Yes	244 (86.2%)	327 (81.8%)	242 (80.7%)	813 (82.7%)	
Missing	7	6	2	15	
Suppose all types of counseling described above are equally effective, which would you be most likely to choose					**<0.01[1]**
One-to-one in-person therapy	169 (59.1%)	183 (45.9%)	168 (56.6%)	520 (53.0%)	
Tele-health	54 (18.9%)	123 (31.6%)	75 (25.3%)	255 (26.0%)	
Message-based Care	39 (13.6%)	44 (11.0%)	38 (12.8%)	121 (12.3%)	
Mobile Mental Health App	24 (8.4%)	46 (11.5%)	16 (5.4%)	86 (8.8%)	
Missing	4	7	5	16	
Do you think you would have any concerns about these options?	115 (39.7%)	180 (44.3%)	141 (46.7%)	436 (43.7%)	0.21[1]
Message-based Care					
Mobile Mental Health Apps	90 (31.0%)	155 (38.2%)	138 (45.7%)	383 (38.4%)	**<0.01[1]**
Tele-health	93 (32.1%)	105 (25.9%)	62 (20.5%)	260 (26.1%)	**<0.01[1]**
Prefer not to answer	93 (32.1%)	105 (25.9%)	62 (20.5%)	260 (26.1%)	**<0.01[1]**
One-to-one in-person therapy / counseling from a licensed clinician	34 (11.7%)	48 (11.8%)	41 (13.6%)	123 (12.3%)	0.73[1]

[1] *Chi-Square p-value;* ***Bold*** *indicates P-value < 0.05 and less than Benjamini-Hochberg critical value, considered to be statistically significant.*

being implemented in care when utilizing digital mental health tools, thus inclusion of patient choice and autonomy is critical (18) Participants in our study saw successful implementation of digital mental health as augmentative to their current treatment. These insights reinforce the importance of digital mental health as one component of a mental health intervention, and to the importance of vetting digital mental health tools with intended user groups, in order to create tools that will have the greatest likelihood of use (10).

LIMITATIONS

This study has a few limitations worth mentioning. First, while we stratified recruitment to reflect US census reports of racial demographic breakdown, ultimately, sample is not entirely representative of the demographic composition for the US population. Underrepresented racial and ethnic groups in this survey include African American, Asian, and American Indian or Alaska Native or Indigenous. The older adult sample was particularly over-representative of the White racial group

TABLE 4 | Role size and themes mentioned in qualitative responses.

Role Size	YA 8–34		MAA 35–59		OA 60+		Whole group	
	N	%	*N*	%	*N*	%	*N*	%
None	14	4.9	29	7.5	31	10.8	74	7.8
Small or limited scope	43	15.2	54	14.0	60	21.0	157	16.5
General or Complementary	159	56.2	224	58.2	151	52.8	534	56.0
Large	67	23.7	78	20.3	44	15.4	189	19.8
Themes mentioned								
Paired with a Provider	74	25.5	118	29.1	100	33.1	292	29.3
For Comm and Admin	45	15.5	50	12.3	57	18.9	153	15.2
Monitoring	25	8.6	42	10.3	27	8.9	94	9.3
Access to care	101	32.5	130	41.8	80	25.7	311	31.1
Specific recommendation	70	24.1	110	27.1	79	26.2	259	26.0

relative to the US census. Additionally, these findings cannot be generalized to non-English speaking individuals. Second, our sample was recruited from an online research community, which is naturally more experienced and inclined to participate in online, digital research. Thus, the perspectives reported here are from populations who may be more comfortable using technology and may also be more aware of the limitations and problems related to digital technology. Thirdly, the participant's experience with mental health services and previous treatment exceeds that of the general population which means our results on preference for treatment are not entirely representative (45). Finally, in many respects, participants were asked to reflect on mental health tools they may not have encountered before (e.g., message-based care) and were thus likely providing initial reactions to the role such tools should play in mental health care. Future research should conduct more in-depth user testing of existing tools to determine if the perspectives reported here are still valid after exposure to using these tools.

CONCLUSION

Digital mental health interventions provide both older adults and younger generations fruitful ways to prevent, assess, treat, and manage mental health conditions. In this study we ascertained important generational views on how adults view technology in conjunction with their mental healthcare. While most adults were supportive of effective digital mental health technologies, our results indicate that there is no "one-size-fits-all" for mental health care and older adults were less accepting of digital mental health than other generations. Adults see value in these technologies through administrative functions such as scheduling, treatment monitoring and the ability to message their provider. The unique preferences and lifestyles of different generations demonstrate the need for variance in mental health care offerings. As the older adult population increases, researchers and systems of care have the responsibility to address patient preferences for care and recognize the barriers and challenges to the use of digital mental health tools.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Washington's Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

PA and BM contributed to conception and design of the study. CR, NS, and TG organized the database. MJ performed the statistical analysis. NS, MW, and TG performed the qualitative coding and analysis. MW, MJ, NS, PA, and CR wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This project was funded by the University of Washington Department of Psychiatry and Behavioral Sciences. This study was supported in part by the Institute of Translational Health Science (ITHS) grant support (UL1 TR002319 NCATS/NIH).

ACKNOWLEDGMENTS

We acknowledge the University of Washington CREATIV Lab's support of this study. This work was also supported by the National Institute of Mental Health (Grant Numbers. R44 MH124334 and P50 MH115837).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2022.840169/full#supplementary-material

REFERENCES

1. Reinert M, Nguyen T, Fritze D. *State of Mental Health in America - 2020*. Alexandria (2019). Available online at: https://mhanational.org/sites/default/files/State%20of%20Mental%20Health%20in%20America%20-%202020_0.pdf
2. Mojtabai R. U.S. health care reform and enduring barriers to mental health care among low-income adults with psychological distress. *Psychiatr Serv.* (2021) 72:338–42. doi: 10.1176/appi.ps.202000194
3. Anderson C. *Geriatric Mental Health Workforce Faces A Growing Shortage*. Healthcare Finance (2012). Available online at: https://www.healthcarefinancenews.com/news/geriatric-mental-health-workforce-faces-growing-shortage#::text=Currently%20there%20are%20about%201%2C700,65%20and%20older%20b
4. Levisohn A, Higgins E. States *Address Provider Shortages to Meet the Health Care Demands of the Pandemic*. National Academy for State Health Policy (2020). Available online at: https://www.nashp.org/states-address-provider-shortages-to-meet-the-health-care-demands-of-the-pandemic/
5. *Administration on Aging A for CL. A Profile of Older Americans: 2017*. Washington, DC (2018). Available online at: https://acl.gov/sites/default/files/Aging%20and%20Disability%20in%20America/2017OlderAmericansProfile.pdf
6. SAMHSA. *Older Adults Living with Serious Mental Illness. The State of the Behavioral Health Workforce*. (2019). Available online at: https://store.samhsa.gov/sites/default/files/d7/priv/pep19-olderadults-smi.pdf
7. U.S. Census Bureau. *The 2017 National Populations Projections*. (2017). Available online at: https://www.census.gov/programs-surveys/popproj.html
8. Blazer DG. Depression in late life: review and commentary. *J Gerontol A Biol Sci Med Sci*. (2003) 58:249–65. doi: 10.1093/gerona/58.3.M249
9. SAMHSA. *Promoting Emotional Health and Preventing Suicide: A Toolkit for Senior Centers*. Rockville (2015). Available online at: https://store.samhsa.gov/sites/default/files/d7/priv/sma15-4416.pdf
10. Mohr DC, Azocar F, Bertagnolli A, Choudhury T, Chrisp P, Frank R, et al. Banbury forum consensus statement on the path forward for digital mental health treatment. *Psychiatric Services*. (2021) 72:677–83. doi: 10.1176/appi.ps.202000561
11. Whitcomb I. *Mental Wellness Apps Are Basically The Wild West of Therapy. Popular Science*. (2021). Available online at: https://www.popsci.com/science/mental-health-apps-safety/#::text=Mental%20health%20apps%20have%20been,to%20the%20American%20Psychological%20Association.
12. Graham SA, Stein N, Shemaj F, Branch OH, Paruthi J, Kanick SC. Older adults engage with personalized digital coaching programs at rates that exceed those of younger adults. *Front Digit Health*. (2021) 3:642818. doi: 10.3389/fdgth.2021.642818
13. Seifert A, Reinwand DA, Schlomann A. Designing and using digital mental health interventions for older adults: being aware of digital inequality. *Front Psychiatry*. (2019) 10:568–568. doi: 10.3389/fpsyt.2019.00568
14. Fortuna KL, Torous J, Depp CA, Jimenez DE, Areán PA, Walker R, et al. A future research agenda for digital geriatric mental healthcare. *Am J Geriatric Psychiatry*. (2019) 27:1277–85. doi: 10.1016/j.jagp.2019.05.013
15. Mata-Greve F, Johnson M, Pullmann MD, Friedman EC, Griffith Fillipo I, Comtois KA, Arean P. Mental health and the perceived usability of digital mental health tools among essential workers and people unemployed due to COVID-19: cross-sectional survey study. *JMIR Mental Health*. (2021) 8:e28360. doi: 10.2196/preprints.28360
16. Renn BN, Hoeft TJ, Lee HS, Bauer AM, Areán PA. Preference for in-person psychotherapy versus digital psychotherapy options for depression: survey of adults in the US. *NPJ Digit Med*. (2019) 2:6. doi: 10.1038/s41746-019-0077-1
17. Haase KR, Cosco T, Kervin L, Riadi I, O'Connell ME. Older adults' experiences with using technology for socialization during the COVID-19 pandemic: cross-sectional survey study. *JMIR Aging*. (2021) 4:e28010. doi: 10.2196/28010
18. Andrews JA, Brown LJE, Hawley MS, Astell AJ. Older adults' perspectives on using digital technology to maintain good mental health: interactive group study. *J Med Internet Res*. (2019) 21:e11694. doi: 10.2196/11694
19. Pywell J, Vijaykumar S, Dodd A, Coventry L. Barriers to older adults' uptake of mobile-based mental health interventions. *Digit Health*. (2020) 6:1–15. doi: 10.1177/2055207620905422
20. Wildenbos GA, Jaspers MWM, Schijven MP, Dusseljee-Peute LW. Mobile health for older adult patients: Using an aging barriers framework to classify usability problems. *Int J Med Inform*. (2019) 124:68–77. doi: 10.1016/j.ijmedinf.2019.01.006
21. Lyon AR, Brewer SK, Areán PA. Leveraging human-centered design to implement modern psychological science: return on an early investment. *Am Psychol*. (2020) 75:1067–79. doi: 10.1037/amp0000652
22. Lyon AR, Munson SA, Renn BN, Atkins DC, Pullmann MD, Friedman E, et al. Use of human-centered design to improve implementation of evidence-based psychotherapies in low-resource communities: protocol for studies applying a framework to assess usability. *JMIR Res Protoc*. (2019) 8:e14990. doi: 10.2196/14990
23. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. *J Biomed Inform*. (2019) 95:1–10. doi: 10.1016/j.jbi.2019.103208
24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. (2009) 42:377–81. doi: 10.1016/j.jbi.2008.08.010
25. Croissant C. *Why Participants Get Banned*. Prolific (2021). Available online at: https://blog.prolific.co/why-participants-get-banned/
26. Bradley P. *Bots and Data Quality On Crowdsourcing Platforms*. Prolific (2018). Available online at: https://blog.prolific.co/bots-and-data-quality-on-crowdsourcing-platforms/
27. Yarrish C, Groshon L, Mitchell J, Appelbaum A, Klock S, Winternitz T, et al. Finding the signal in the noise: minimizing responses from bots and inattentive humans in online research. *Behav Ther*. (2019) 42:235.
28. Toothman EL, Barrett AE. Mapping midlife: an examination of social factors shaping conceptions of the timing of middle age. *Adv Life Course Res*. (2011) 16:99–111. doi: 10.1016/j.alcr.2011.08.003
29. Levinson DJ. A conception of adult development. *Am Psychol*. (1986) 41:3–13. doi: 10.1037/0003-066X.41.1.3
30. Lachman ME, Teshale S, Agrigoroaei S. Midlife as a pivotal period in the life course: Balancing growth and decline at the crossroads of youth and old age. *Int J Behav Dev*. (2015) 39:20–31. doi: 10.1177/0165025414533223
31. Russo-Netzer P, Littman-Ovadia H. "Something to live for": experiences, resources, and personal strengths in late adulthood. *Front Psychol*. (2019) 10:2452. doi: 10.3389/fpsyg.2019.02452
32. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9 validity of a brief depression severity measure. *J Gen Intern Med*. (2001) 16:606–13. doi: 10.1046/j.1525-1497.2001.016009606.x
33. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder. *Arch Intern Med*. (2006) 166:1092. doi: 10.1001/archinte.166.10.1092
34. Benjamini Y, Hochberg Y. On the adaptive control of the false discovery rate in multiple testing with independent statistics. *J Educ Behav Stat*. (2000) 25:60–83. doi: 10.2307/1165312
35. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Series B*. (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
36. Neuendorf KA. *The Content Analysis Guidebook. 2nd ed*. Thousand Oaks: SAGE Publications (2016). p. 1–426.
37. Perrin A, Atske S. 7% of Americans don't use the internet. who are they? Washington DC (2021). Available online at: https://www.pewresearch.org/fact-tank/2021/04/02/7-of-americans-dont-use-the-internet-who-are-they/ (accessed November 3, 2021).
38. Alexopoulos GS. Digital mental health for older adults. *Am J Geriatr Psychiatry*. (2020) 28:191–3. doi: 10.1016/j.jagp.2019.06.013
39. Wang S, Bolling K, Mao W, Reichstadt J, Jeste D, Kim HC, Nebeker C. Technology to support aging in place: older adults' perspectives. *Healthcare*. (2019) 7:60. doi: 10.3390/healthcare7020060
40. Gordon NP, Crouch E. Digital information technology use and patient preferences for internet-based health education modalities: cross-Sectional survey study of middle-aged and older adults with chronic health conditions. *JMIR Aging*. (2019) 2:e12243. doi: 10.2196/12243

41. March S, Day J, Ritchie G, Rowe A, Gough J, Hall T, et al. Attitudes toward e-mental health services in a community sample of adults: Online survey. *J Med Int Res.* (2018) 20:e59. doi: 10.2196/jmir.9109
42. Elueze I, Quan-Haase A. Privacy attitudes and concerns in the digital lives of older adults: westin's privacy attitude typology revisited. *Am Behav Scientist.* (2018) 62:1372–91. doi: 10.1177/0002764218787026
43. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. *J Affect Disord.* (2020) 263:413–9. doi: 10.1016/j.jad.2019.11.167
44. Swift JK, Callahan JL, Cooper M, Parkin SR. The impact of accommodating client preference in psychotherapy: a meta-analysis. *J Clin Psychol.* (2018) 74:1924–37. doi: 10.1002/jclp.22680
45. Conner KO, Carr Copeland V, Grote NK, Koeske G, Rosen D, Reynolds Iii CF, et al. Mental health treatment seeking among older adults with depression: the impact of stigma and race. *Am J Geriatr Psychiatry.* (2010) 18:531–43. doi: 10.1097/JGP.0b013e3181cc0366

Conflict 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.

frontiers | Frontiers in Digital Health

ORIGINAL RESEARCH
published: 19 May 2022
doi: 10.3389/fdgth.2022.886783

Engagement in Digital Health App-Based Prevention Programs Is Associated With Weight Loss Among Adults Age 65+

Lisa A. Auster-Gussman, Kimberly G. Lockwood, Sarah A. Graham, Viveka Pitter and OraLee H. Branch**

Lark Health, Mountain View, CA, United States

Background: Digital health programs have been shown to be feasible and effective for the prevention of chronic diseases such as diabetes. Contrary to expectations, findings also suggest that older adults have higher levels of engagement with digital health programs than younger adults. However, there is a paucity of research examining outcomes among older adults in digital health programs and whether higher engagement is related to better outcomes.

Methods: We examined weight loss outcomes for 538 users aged 65 and older participating in one of two app-based prevention programs called the Diabetes Prevention Program and the Prevention Program, respectively. Both programs were available on a single artificial intelligence (AI)-powered digital health platform and shared a common goal of weight loss. We also examined the relationship between key engagement metrics (i.e., conversing with the AI-powered coach, weigh-ins, and initiating educational lessons early in the program) and weight loss outcomes.

Results: The average weight loss of all enrollees having a weight measurement after after the 9th week was 4.51%, and the average weight loss of the Diabetes Prevention Program enrollees meeting a minimum engagement level was 8.56%. Greater weight loss was associated with a greater number of days with AI-powered coaching conversations ($p = 0.03$), more weigh-ins ($p = 0.00$), and early educational lesson initiation ($p = 0.02$).

Conclusions: Digital health programs powered by AI offer a promising solution for health management among older adults. The results show positive health outcomes using app-based prevention programs, and all three engagement metrics were independently associated with weight loss.

Keywords: older adults, mHealth, weight loss, prevention, engagement, digital health, diabetes

OPEN ACCESS

Edited by:
Carlo Massaroni, Campus Bio-Medico University, Italy

Reviewed by:
Jeff Kullgren, University of Michigan, United States
Ivan Miguel Pires, Universidade da Beira Interior, Portugal

****Correspondence:***
Lisa A. Auster-Gussman
lisa.austergussman@lark.com
OraLee H. Branch
oralee.branch@lark.com

Specialty section:
This article was submitted to Connected Health, a section of the journal Frontiers in Digital Health

Received: *28 February 2022*
Accepted: *19 April 2022*
Published: *19 May 2022*

Citation:
Auster-Gussman LA, Lockwood KG, Graham SA, Pitter V and Branch OH (2022) Engagement in Digital Health App-Based Prevention Programs Is Associated With Weight Loss Among Adults Age 65+. Front. Digit. Health 4:886783. doi: 10.3389/fdgth.2022.886783

INTRODUCTION

The population of older adults, 65 years and older, in the United States is increasing rapidly (1). As a result, a top healthcare priority is the expansion and improvement of programs for prevention of chronic diseases, which are especially prevalent among older adults (1). Scalable solutions for disease prevention are essential given that there are over 54 million older adults in the United States as of 2020 and both their numbers and risk for chronic diseases, such as diabetes, hypertension, and obesity, are increasing (1). Digital health programs are a scalable way to facilitate the prevention of chronic diseases

across age groups, and emerging research suggests that this includes older adults (2).

Although past researchers have deemed age to be the "largest barrier to digital health adoption" (3), recent evidence suggests older adults are not only willing to engage with digital health offerings but do so at similar or even higher rates than younger adults. For example, older adults report willingness to engage with smartphone-based technologies for pain management (4). Moreover, recent research indicates that older adults show higher engagement with digital health programs than younger adults, contrary to conventional wisdom that older adults face too many barriers to engage with digital health programs at the same level as younger adults (5). The results from Graham and colleagues (5), comparing engagement in digital health programs of adults 35 to 64 years to those over 65, demonstrated that adults over 65 engaged in significantly more coaching conversations, logged more meals, and provided more connected device measurements than younger adults. Furthermore, research on engagement in a digital health program among a Medicare population showed that 92% of participants completed at least nine out of 16 program lessons and engaged in 19 out of 31 opportunities for weekly program engagement (6). Among 140 participants aged 50 to 80 years, 65% showed long-term engagement, operationalized as completing at least one task activity per month over 4 months, and there were no differences when stratified by age (i.e., 50 to 64, and 65 to 80) (7). These studies provide powerful evidence that older adults are both willing and able to engage with digital health programs.

Although evidence is mounting on engagement among older adults in fully digital health programs, little research has been conducted on clinical outcomes among older adults in these programs. The published studies in this domain are primarily feasibility and acceptability trials, trials with very small samples, include adults under 65 years of age, or examine programs that are not fully digital. Qualitative acceptability research suggests that older adults believe a digital health coach may help them improve health behaviors, such as increase their physical activity, but this has not been tested (8). One study measured feasibility and acceptability of a fully digital health program among older adults and reported an average weight loss of 3.44 pounds over 4 months, showing promise for the effectiveness of fully digital health programs for this population, but only included 19 individuals (7). A meta-analysis of six trials of adults with a mean age of 68 years old indicated that programs with smartphone-based intervention components helped these individuals decrease sedentary time, increase physical activity, and increase fitness; however, these studies included adults as young as 55 years of age (9). Finally, a study using a Medicare population with a mean age of 68 years old showed weight loss at 12 months, but this study included a combination of human and digital health coaching (6). In sum, this growing literature supports the fact that older adults can both engage with and benefit from digital health offerings, but whether older adults can lose weight in a fully digital program has not been tested.

As research on engagement has mounted, questions related to the importance of the timing of engagement have emerged, with a focus on early engagement in digital health programs as a predictor of longer-term positive health outcomes. These studies were not focused on older adults, but they demonstrate the importance of early engagement, and they beg the question of whether this would be similar in the older population. For example, early program engagement predicted weight loss at 1 year in a combined behavioral and pharmacotherapy weight loss program (10). Specifically, each additional day of meal logging during the 1st 3 weeks of the program was related to a 7% increase in the odds of attaining at least 5% weight loss at 1 year (10). Research on weight loss among emerging adults in a combined web-based and in-person behavioral weight loss program showed that engagement during the initial 4 weeks of treatment, operationalized as attendance at an initial in-person session and at least weekly weight reporting, was associated with increased weight loss (11). A variety of studies have also reported the importance of early engagers in weight loss trials; specifically, those who have the best weight loss in the first month also have better long-term weight loss (12). Taken together, these studies indicate that early program engagement plays a key role in predicting later clinical program outcomes.

There is a paucity of evidence focused on outcomes specifically among older adults and on the relationship between engagement and outcomes among this population in fully digital prevention programs. Therefore, this study examined weight loss as the primary outcome of two fully digital preventive health programs because weight loss is a common metric used to assess success in such programs and because even small amounts of weight loss are related to improvements in other clinically relevant metrics such as hemoglobin A1c, triglycerides, systolic blood pressure, and LDL cholestorol (13). The primary purpose of this research was to examine weight loss among adults 65 years and older enrolled in one of two prevention programs on a single digital health platform. Specifically, we examined weight loss among this population as well as the relationship between engagement and weight loss, with a focus on the impact of early engagement. The primary hypothesis was that higher engagement would be associated with greater weight loss among older adults.

METHODS

Study Design

This was a longitudinal, observational study of users enrolled in an AI-powered digital chronic disease prevention program available via a smartphone app called Lark. We examined weight loss as well as the relationship between engagement and weight loss. The study received exemption status from Advarra Institutional Review Board (Protocol #Pro00047181) for retrospective analyses of previously collected and de-identified data.

Participants and Recruitment

Participants were 538 users of an AI-based chronic disease prevention coaching app who joined one of two prevention programs [see Graham and colleagues (5) for details] offered through the platform (i.e., Diabetes Prevention Program or Prevention Program), both of which had a primary outcome of weight loss. All users had private insurance and gained access

to the app at no cost via partnerships between the app and their insurance provider. Those who were eligible and signed up received a link via text message to download the Lark app to their smartphones. Briefly, the Diabetes Prevention Program followed the National Diabetes Prevention Program (NDPP) guidelines and the established Prevent T2 curriculum for delaying or preventing progression to type 2 diabetes (14). Diabetes Prevention Program users must meet risk criteria established by the Centers for Disease Control and Prevention (14).

The Prevention program targets individuals who may not meet strict criteria for participation in the NDPP. The Prevention program emphasizes taking small steps that lead to significant and lasting behavior change in the areas of nutrition, physical activity, weight loss, sleep, and stress reduction. Any user with insurance coverage for Lark can sign up for the Prevention Program. Most participants in both programs (97.6% overall) had access to a connected digital body weight scale, provided through participation in the program, that automatically transmitted their weigh-ins to the digital platform.

Inclusion Criteria

Inclusion criteria for the analytic sample were: (1) enrollment in either the Diabetes Prevention or Prevention program on or after January 1, 2019, (2) aged 65 years and older, (3) a starting BMI $\geq$25, (4) in the program for at least 3 months, (5) had a weight after 9 weeks, (6) had full demographic information, and (7) had weight loss set as a goal in their program. Exclusion criteria were: (1) those who had earlier versions of the app (i.e., before January 1, 2019) and (2) those who had been in the programs <3 months. We also excluded users who were normal weight or underweight because the primary outcome was weight loss.

Program Flow

Users who qualified for the app, downloaded it, and signed up for a given program provided all measures and outcomes via the app-based digital health platform. The AI-based platform includes two prevention programs: Diabetes Prevention Program and Prevention Program. These programs include a core set of features plus condition-specific content. All programs are delivered via an iPhone or Android smartphone. They each include a series of educational lessons delivered via conversational AI as well as calls-to-action and nudges which provide positive reinforcement and prompt users to enter health data, log a meal, complete a lesson, or weigh themselves. The AI coach encourages user behavior change using cognitive behavioral therapy techniques and is available 24 hours per day for users to check in and discuss challenges or progress.

Engagement Measures

We examined three engagement variables as predictors of weight loss and measured these engagement variables and their relation to weight loss. All predictors occurred during the first 9 weeks of the program because the focus was on modeling early engagement as a predictor of later clinical outcomes. The three engagement metrics each represented different types of engagement: conversations, early lesson initiation, and weigh-ins. *Conversations* was the percentage of days in the first 9 weeks that a user engaged in a two-way conversation within the app and represented active app usage and contact with the coach. *Early lesson initiation* was a binary variable indicating whether a user completed >2 lessons during the first 3 weeks of the program and represented engagement with educational content beyond simple contact with the coach. The *early lesson initiation* variable captured those with high early use, as a single lesson took seven days of app usage to complete. Completion of two lessons in the first 2 weeks indicated 100% daily usage, so we expanded the window to include the third week of the program. *Weigh-ins* was the number of weigh-ins during the first 9 weeks of the program and represented a real-world behavior prompted by the coach.

Outcome Measures

The primary outcome was *percent weight loss*, calculated as (first weight-nadir weight)/first weight. First weight was the average of weights recorded on the first day a user provided weights. Nadir weight was the minimum recorded weight occurring after >9 weeks in the program, as past research suggests that the first several weeks in a program are critical to the prediction of long-term outcomes (12). As a quality control, the digital platform flags abnormal weigh-ins for review by identifying a weight loss rate of >7lbs/week and removes outliers unless confirmed by the member to be accurate. The duration of active participation in a digital health program may vary across members; thus, the use of nadir weight and time-to-nadir enabled a larger number of members to be assessed for weight loss outcomes. The mean time-to-nadir was 138 days after program enrollment (SD = 71 days), suggesting that the nadir weight occurred on average after approximately 4.5 months of program participation.

Statistical Analyses

We conducted all statistical analyses in R version 1.4.1717 (15). Users self-reported their age, gender, race, and height upon enrollment. We calculated BMI (kg/m^2) from height and starting weight and obtained engagement variables directly from users' interactions with the digital platform. We used linear regression to examine the association between engagement and percent weight loss (dependent variable) and log-transformed variables with non-normal distributions as necessary. We combined the Diabetes Prevention and Prevention programs for the regression since the programs shared a common goal (weight loss) and combining them increased the sample size for analysis. We also included program type as an independent variable in the model. We checked for multicollinearity of the independent variables in the model using variance inflation factors. All other analyses were simple descriptive statistics for each included variable. The *a priori* alpha level was $p < 0.05$.

RESULTS

Participants

Our final sample included 538 users (see **Figure 1** for inclusion flow chart) with most individuals enrolled in the Diabetes Prevention Program (N = 489). Mean age was 67.47 years

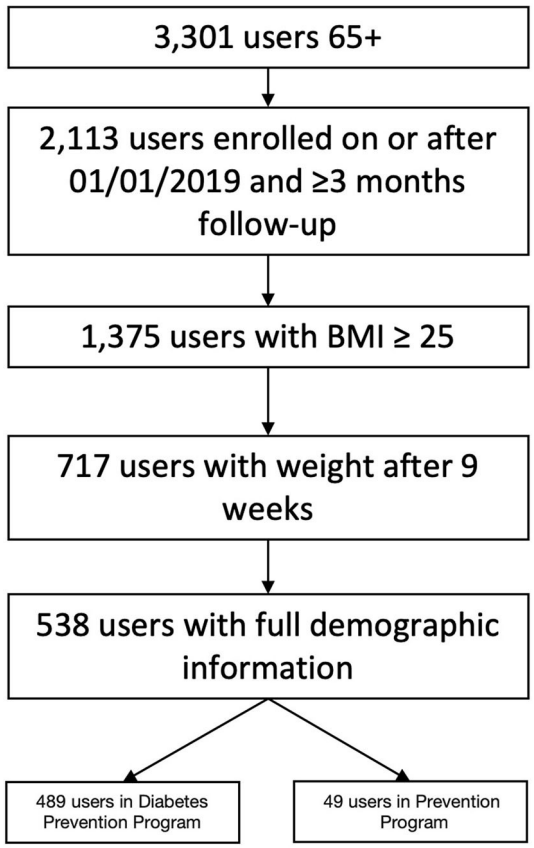


FIGURE 1 | Users based on inclusion and exclusion criteria.

TABLE 1 | Participants, engagement, and outcome metrics.

	Mean (SE/CI/IQR) or % (*n*)
Participants	
Age (years)	67.47 (0.14)
	Median 66.00 [25% = 65.00, 75% = 69.00]
Starting BMI (kg/m^2)	32.61 (5.59)
	Median 31.32 [25% = 28.28, 75% = 35.64]
% Female	61% (*N* = 329)
% Non-white	16% (*N* = 84)
Engagement features (9 weeks)	
Mean weigh-ins	21.76 (0.80)
Mean conversations	145.58 (4.59)
% Early mission initiation	45% (*N* = 241)
Weight loss	
Overall (*N* = 538)	4.51% (95% CI [4.14, 4.87])
Diabetes prevention program (*N* = 489)	4.57% (95% CI [4.19, 4.94])
CDC qualifiers (*N* = 60)	8.56% (95% CI [7.03, 10.08])
Prevention program (*N* = 49)	3.91% (95% CI [2.52, 5.31])
Mean nadir weight day	137.60 (3.06)

N = 538 unless noted otherwise; "CDC Qualifiers" indicates the subset of users in the Diabetes Prevention Program that met minimum lesson completion standards set forth by the CDC. This means completion of at least 3 lessons in the first 6 months and at least 1 lesson from months 9–12.

(SD = 3.28). More than half of users were female (61%), and 16% were non-white. Mean Body Mass Index (BMI) was 32.61 kg/m^2 (SD = 5.59) (see **Table 1** for complete demographics and characteristics).

Percent Weight Loss and Engagement

Mean overall percent weight loss was 4.51% (95% CI = 4.14, 4.87). Further details and program-specific means are in **Table 1**. An examination of engagement revealed that the mean number of *conversations* during the first 9 weeks was 145.58 (SE = 4.59). The percentage of users with *early lesson initiation* was 45%. The mean number of *weigh-ins* during the first 9 weeks was 21.76 (SE = 0.80). The mean nadir weight day was 137.60 (SE = 3.06) days after program enrollment, suggesting that nadir weight occurred on average after approximately 4.5 months, but with variation in the time of peak loss across users.

Relationship Between Engagement and Percent Weight Loss

The final regression model predicting percent weight loss included the effect of conversations, early lesson initiation, and weigh-ins. Control variables included user demographics as well as starting BMI, time-to-nadir weight, and program type (see **Table 2** for full results; all independent variables are included in **Table 2**). The model had good overall fit [$R^2 = 0.3$, $F_{(9,528)} = 24.4$, $p < 0.001$]. We observed that all three engagement variables were significantly related to percent weight loss, such that a higher number of *conversations*, more *weigh-ins*, and *early lesson initiation* were related to greater weight loss. In addition, all variance inflation factor (VIF) values were between 1.03 and 1.64, indicating no issues with multicollinearity, and bivariate correlations between the engagement variables ranged from $r = -0.006$ to $r = 0.32$.

DISCUSSION

The results of this study support the primary hypothesis that increased engagement was related to a greater percent weight loss among older adults. The relationship between program engagement and clinical outcomes suggests that digital health programs are an effective way to promote weight loss and, potentially, other related health outcomes among older adults. This research adds to the growing body of literature on older adults and digital health, further suggesting that greater engagement is associated with positive health outcomes among older adults.

Contribution to the Literature

Contrary to past conjecture, this study supports findings from recent research that suggests older adults are able and willing to engage in digital health apps and do engage in digital health.[5] Past

TABLE 2 | Results of the regression modeling the effect of engagement on weight loss percent.

	Unstandardized B	SE	95% CI		*p*
			LL	UL	
Constant	**−8.69**	**3.63**	**−15.81**	**−1.57**	**0.02**
Control variables					
Sex	0.32	0.33	−0.33	0.98	0.33
Race	−0.34	0.44	−1.20	0.52	0.44
Age	0.05	0.05	−0.05	0.14	0.35
Body mass index	0.05	0.03	−0.01	0.10	0.12
Nadir weight day	**0.03**	**0.00**	**0.02**	**0.03**	**0.00**
Program type (Ref. DPP)					
Prevention	−0.11	0.55	−1.21	0.99	0.84
Engagement					
Early lesson initiation	**0.89**	**0.39**	**0.14**	**1.65**	**0.02**
Conversations	**0.57**	**0.25**	**0.07**	**1.07**	**0.03**
Weigh-ins	**0.65**	**0.19**	**0.28**	**1.01**	**0.00**

$p < 0.05$ shown in bold. DPP, Diabetes Prevention Program. All independent variables used in analyses are included in table.

research has found that adults aged 65 and older can be successful in hybrid programs (6), but this research is the first to provide evidence of success in fully digital programs on a relatively large scale. The observed average weight loss of 4.5% in this study supports a previous small-sample feasibility study of a fully digital mobile app-based program that showed an average weight loss of 3.44 pounds over 4 months, indicating that older adults can successfully lose weight using fully digital programs (7). Results from research examining the association between engagement, age, and weight loss are instructive, although they include adults across the age span from 18 to 85 years of age (16). Specifically, researchers found that the association between engagement and weight loss was stronger for younger people compared to older adults (16). Although this may be true, our findings suggest that engagement is an important predictor of outcomes among older adults; thus, there is still merit in encouraging higher engagement among older adults.

Since we found that engagement is critical for outcomes, considering how different types of engagement relate to outcomes is also important. Past research has suggested the importance of measuring engagement at different levels of analysis, namely, "Big E" and "Little e" engagement, the former referring to health behavior engagement (e.g., weigh-ins) and the latter to app engagement, which is further broken down into app-use engagement (e.g., conversations) and behavior change content engagement (e.g., educational lessons) (17). The authors suggested the importance of examining engagement at multiple levels to better understand how app-use engagement is related to outcomes (e.g., the relation between conversations and weight loss) and the ways in which the outcome is explained by health behavior engagement, which is related to but separate from app-use engagement (e.g., the unique variance contributed by weigh-ins above and beyond conversations).

Assessing engagement at multiple levels also helps researchers understand which aspects of engagement are most important for influencing a given outcome. For example, if we had found that conversations, but not early lesson completion, was significant this might indicate the primary importance of two-way interactions with coaches regardless of participation in educational content. However, in line with the suggestions of past researchers, the results of this study demonstrated that engagement at each level independently contributed to weight loss; *conversations,* which showed active app usage and contact with the coach, *early lesson initiation,* which showed engagement with educational content beyond contact with the coach, and *weigh-ins,* which showed real-world behavior prompted by the coach. Our results indicate that it is not just one type of engagement that predicts weight loss, but that greater engagement at each of these three levels uniquely contributes to this important indicator of improved health. This finding is similar to results from a study with younger adults, which showed that a cluster of different engagement variables (e.g., attendance at an initial session in person, weight reporting) predicted weight loss (11). Thus, both older and younger adults appear to benefit from multiple modes of engagement, and our results further suggest that such early multifaceted engagement predicts improved weight loss outcomes in older adults.

Strengths and Limitations

There were several limitations in this study. All data were retrospective and collected via the digital health app platform. There were fewer users in the Prevention Program compared to the Diabetes Prevention Program, but the regression coefficient for program was not significant indicating that weight loss did not differ by program. The study only included users who had full demographic data available, as well as a BMI $\geq$25 and a weight loss goal, which led to a decrease in the available sample size. However, including complete data on all variables was an important first step in revealing predictors of weight loss among older adult users of a fully digital program. We aimed to capture engagement at three different levels, namely app-use engagement, behavior change content engagement, and health behavior (i.e., weigh-in) engagement. These measures provided important engagement insights but were still relatively rudimentary; future research could explore more nuanced types of engagement (e.g., patterns of engagement over time) or ways of operationalizing engagement that better capture unique user behavior patterns or trends. In line with research suggesting that older adult internet usage is increasing (18), these findings suggest that the digital divide may not be as wide as it once was. However, we only examined older adults who were willing and able to sign up for a digital app. Future research could examine whether there are age-specific factors in digital health design or content that lead to greater willingness of older adults to use fully digital health programs or facilitate even greater engagement given that encouraging and maintaining high rates of engagement is a key challenge in digital health. The primary strength of this study was that all data came from real-world users of the app rather than participants recruited specifically for a research study. Future users of the app would have the same user

experience as this sample since there were no study components that occurred outside of the app interface.

CONCLUSIONS

The findings of this study suggest that older adults lose weight while using preventive health programs on a fully digital platform, and that increased engagement is related to increased weight loss. As the need for scalable solutions for older adults rapidly increases over the coming years, digital health programs may offer a solution to the rapidly increasing needs of individuals and communities at risk for, and living with, chronic diseases. Scalable, fully digital health solutions were previously thought to be efficacious only among younger adults. However, the findings presented here reveal that this assumption is, at the very least, worthy of being tested in research. These findings are a first step in developing a body of literature supporting digital health solutions for all ages, including older adults.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Advarra (Protocol #Pro00047181) Institutional Review Board. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

VP conducted data cleaning, organization, and analysis for this manuscript. LA-G conducted data analysis. KL, SG, and OB assisted in analysis, writing, and editing. All authors contributed to the article and approved the submitted version.

REFERENCES

1. National Center for Chronic Disease Prevention and Health Promotion. *Chronic disease* [Fact sheets]. Centers for Disease Control and Prevention. (2021). Available online at: https://www.cdc.gov/chronicdisease/resources/publications/fact-sheets.htm (accessed August 31, 2021).
2. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. *BMC Health Serv Res.* (2016) 16:290. doi: 10.1186/s12913-016-1522-3
3. Evangelista L, Steinhubl SR, Topol EJ. Digital health care for older adults. *Lancet.* (2019) 393:1493. doi: 10.1016/S0140-6736(19)30800-1
4. Bhattarai P, Phillips JL. The role of digital health technologies in management of pain in older people: An integrative review. *Arch Gerontol Geriatr.* (2017) 68:14–24. doi: 10.1016/j.archger.2016.08.008
5. Graham SA, Stein N, Shemaj F, Branch OH, Paruthi J, Kanick SC. Older adults engage with personalized digital coaching programs at rates that exceed those of younger adults. *Front Digit Health.* (2021) 3:642818. doi: 10.3389/fdgth.2021.642818
6. Castro-Sweet CM, Chiguluri V, Gumpina R, Abbott P, Madero EN, Payne M, et al. Outcomes of a digital health program with human coaching for diabetes risk reduction in a medicare population. *J Aging Health.* (2018) 30:692–710. doi: 10.1177/0898264316688791
7. Tam E, Boas PKV, Ruaro F, Flesch J, Wu J, Thomas A, et al. Feasibility and adoption of a focused digital wellness program in older adults. *Geriatrics.* (2021) 6:54. doi: 10.3390/geriatrics6020054
8. Santini S, Galassi F, Kropf J, Stara V, A digital coach promoting healthy aging among older adults in transition to retirement: results from a qualitative study in Italy. *Sustainability.* (2020) 12:7400. doi: 10.3390/su12187400
9. Yerrakalva D, Yerrakalva D, Hajna S, Griffin S. Effects of mobile health app interventions on sedentary time, physical activity, and fitness in older adults: systematic review and meta-analysis. *J Med Internet Res.* (2019) 21:e14343. doi: 10.2196/14343
10. Fabricatore AN, Wadden TA, Moore RH, Butryn ML, Heymsfield SB, Nguyen AM. Predictors of attrition and weight loss success: results from a randomized controlled trial. *Behav Res Ther.* (2009) 47:685–91. doi: 10.1016/j.brat.2009.05.004
11. Gokee-LaRose J, Fava JL, Lanoye A, Caccavale LJ. Early engagement is associated with better weight loss in emerging adults. *Am J Health Behav.* (2019) 43:795–801. doi: 10.5993/AJHB.43.4.12
12. Unick JL, Leahey T, Kent K, Wing RR. Examination of whether early weight loss predicts 1-year weight loss among those enrolled in an internet-based weight loss program. *Int J Obes.* (2015) 39:1558–60. doi: 10.1038/ijo.2015.89
13. Ryan DH, Yockey SR. Weight loss and improvement in comorbidity: differences at 5%, 10%, 15%, and over. *Curr Obes Rep.* (2017) 6:187–94. doi: 10.1007/s13679-017-0262-y
14. Centers for Disease Control and Prevention. *Diabetes Prevention Recognition Program Standards and Operating Procedures*. Available online at: https://www.cdc.gov/diabetes/prevention/recognition/ (accessed April 30, 2021).
15. R Core Team. *R: A language and environment for statistical computing (Version 1.4.1717)*. R Foundation for Statistical Computing, Vienna, Austria. (2020). Available online at: https://www.R-project.org/
16. DeLuca L, Toro-Ramos T, Michaelides A, Seng E, Swencionis C. Relationship between age and weight loss in Noom: Quasi-experimental study. *JMIR Diabetes*. (2020) 5:e18363. doi: 10.2196/18363
17. Cole-Lewis H, Ezeanochie N, Turgiss J. Understanding health behavior technology engagement: Pathway to measuring digital behavior change interventions. *JMIR Form Rese.* (2019) 3:e14052. doi: 10.2196/14052
18. Hong YA, Cho J. Has the digital health divide widened? Trends of health-related internet use among older adults from 2003 to 2011. *J Gerontol B Psychol Sci Soc Sci.* (2016) 72:856–63. doi: 10.1093/geronb/gbw100

Conflict of Interest: All authors are employed by Lark Health.

frontiers | Frontiers in Digital Health

ORIGINAL RESEARCH
published: 10 June 2022
doi: 10.3389/fdgth.2022.884208

Wearable Use in an Observational Study Among Older Adults: Adherence, Feasibility, and Effects of Clinicodemographic Factors

Emily W. Paolillo[1]*, *Shannon Y. Lee*[1], *Anna VandeBunte*[1,2], *Nina Djukic*[1], *Corrina Fonseca*[1], *Joel H. Kramer*[1] and *Kaitlin B. Casaletto*[1]

[1] Memory and Aging Center, Department of Neurology, Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, United States, [2] Department of Psychology, Palo Alto University, Palo Alto, CA, United States

Introduction: Wearables have great potential to improve monitoring and delivery of physical activity interventions to older adults with downstream benefits to multisystem health and longevity; however, benefits obtained from wearables depend on their uptake and usage. Few studies have examined person-specific factors that relate to wearable adherence. We characterized adherence to using a wearable activity tracker for 30 days and examined associations between adherence and demographics, cognitive functioning, brain volumes, and technology familiarity among community-dwelling older adults.

Methods: Participants were 175 older adults enrolled in the UCSF Longitudinal Brain Aging Study who were asked to wear a Fitbit™ Flex 2 during waking hours for 30 days. Sixty two of these participants were also asked to sync their devices to the Fitbit smartphone app daily to collect minute-level data. We calculated adherence to wearing the Fitbit daily (i.e., proportion of days with valid activity data) and adherence to daily device syncing (i.e., proportion of days with minute-level activity data). Participants also completed a brain MRI and in-person cognitive testing measuring memory, executive functioning, and processing speed. Spearman correlations, Wilcoxon rank sum tests, and logistic regression tested relationships between wearable adherence and clinicodemographic factors.

Results: Participants wore the Fitbits for an average of 95% of study days and were 85% adherent to the daily syncing protocol. Greater adherence to wearing the device was related to female sex. Greater adherence to daily device syncing was related to better memory, independent of demographic factors. Wearable adherence was not significantly related to age, education, executive functioning, processing speed, brain gray matter volumes, or self-reported familiarity with technology. Participants reported little-to-no difficulty using the wearable and all reported willingness to participate in another wearable study in the future.

Conclusions: Older adults have overall high adherence to wearable use in the current study protocol. Person-specific factors, however, may represent potential barriers to equitable uptake of wearables for physical activity among older adults,

OPEN ACCESS

Edited by:
Raeanne Moore, University of California, San Diego, United States

Reviewed by:
Jorge Calvillo-Arbizu, Sevilla University, Spain
Nele A. J. De Witte, Thomas More University of Applied Sciences, Belgium

***Correspondence:**
Emily W. Paolillo
emily.paolillo@ucsf.edu

Specialty section:
This article was submitted to Human Factors and Digital Health, a section of the journal Frontiers in Digital Health

Received: *25 February 2022*
Accepted: *20 May 2022*
Published: *10 June 2022*

Citation:
Paolillo EW, Lee SY, VandeBunte A, Djukic N, Fonseca C, Kramer JH and Casaletto KB (2022) Wearable Use in an Observational Study Among Older Adults: Adherence, Feasibility, and Effects of Clinicodemographic Factors. Front. Digit. Health 4:884208. doi: 10.3389/fdgth.2022.884208

including demographics and cognitive functioning. Future studies and clinical providers utilizing wearable activity trackers with older adults may benefit from implementation of reminders (e.g., texts, calls) for device use, particularly among men and individuals with memory impairment.

Keywords: digital health, Fitbit, aging, memory, physical activity, wearable adherence

INTRODUCTION

The population of older adults is growing worldwide (1). In conjunction, the prevalence of multimorbid geriatric health conditions is increasing faster than the rate at which effective healthcare resources for older adults are being implemented (2–4). There is a pressing need to identify targets for preventative medicine approaches to maintaining optimal health among this growing population. Physical activity is one modifiable behavioral factor that has been strongly and consistently linked to better health across many domains, including cardiovascular health, physical frailty, mental health, and cognitive functioning (5, 6). In addition to the public health and economic benefits of reducing the burden of multimorbidity among older adults (7, 8), there are also clear individual benefits including improvement in quality of life and prolonged functional independence (9). Still, physical activity interventions in clinical settings remain underused (10) and even when patients are advised or encouraged to increase their physical activity by their healthcare providers (11, 12), there may be little long-term follow up.

Wearable devices are a potentially feasible, accessible, and effective way to bridge the gap between research and implementation with regard to physical activity as a preventative health measure. In fact, wrist-worn wearables for tracking physical activity are gaining popularity, even among older adults. Recent estimates suggest that 17% of U.S. adults aged 50 or older already use activity watches/wearable trackers regularly (13). Smart activity watches passively track and transmit objective activity information to an accessible yet secure cloud-based storage system, circumventing prior approaches reliant on self-report, which are often biased by recall errors, social desirability effects, or state-dependent bias (14, 15). In addition to collecting real-time objective data, wearables are also capable of delivering real-time individualized interventions to increase physical activity, including prompts to move when the device detects lack of activity (16). Many observational studies and interventions have already been conducted in several pediatric and adult populations (17, 18); however, less is known about the best practices for using wearables in research studies or clinical interventions with older adults.

Several studies support the validity of wearables for measuring physical activity in older adults and willingness to use these devices. A wide range of devices have been validated as accurate measures of activity, including both research-grade and commercially available wrist-worn devices (19, 20). Although no studies to date have examined factors that relate to wearable adherence among older adults in the context of a study or intervention, prior work has examined factors that relate to naturalistic wearable use. For example, Kononova et al. (21) examined factors that facilitate real-world wearable use among older adults, with long-term users being strongly motivated by social support and collaboration, while short-term users seemed most focused on competitive desires to increase physical activity. Older adult perceptions and real-world uses of activity trackers have also been well-characterized, with studies showing overall high levels of acceptability (22); however, acceptability and subsequent use is still highly dependent on a number of factors including cost, privacy, personal motivation, understanding device purpose, and ease of use (22–26). While studies thus far have demonstrated that many older adults are able to engage with wearable devices, there appear to be many device-specific qualities and subjective perceptions about wearables that affect their naturalistic uptake. Furthermore, cognitive changes, including declines in processing speed, memory, and executive functioning occur with age, highlighting the need to consider how cognitive and brain health relate to wearable adherence (e.g., forgetting to wear the device due to memory problems) in this population. Given the need for better implementation strategies for preventative healthcare among the growing population of older adults, it is imperative to examine person-specific factors that might be barriers to wearable use in the context of a study or intervention.

Thus, the primary aims of this study are to: (1) characterize engagement with wearables for physical activity among older adults using data from an observational exercise study; (2) examine associations between adherence to wearable usage and demographics, cognitive functioning, brain volumes, and self-reported familiarity with technology; and (3) characterize feedback from a post-study questionnaire. We hypothesized that better adherence to wearable usage will be related to younger age, better cognitive functioning, larger brain volumes, and greater familiarity with technology.

MATERIALS AND METHODS

Participants

This study cohort included 175 English-speaking older adults aged 55 years and older who were recruited from the UCSF Longitudinal Brain Aging Study at the UCSF Memory and Aging Center. This parent study totals 408 English-speaking participants (56% Female; age_{mean}= 76.5 years; $education_{mean}$ = 17.4 years; 83% White, 3% Black/African American, 9% Asian/Pacific Islander, and 5% Other or Unknown Race). Inclusion criteria for the Longitudinal Brain Aging study enrollment consisted of being age 55 and older and having no history or current evidence of the following conditions: clinically significant

stroke, acquired brain injuries, DSM-5 major psychiatric disorders, Multiple Sclerosis, Parkinson's Disease, major memory concerns or related diagnoses, active substance abuse, Diabetes Mellitus, Hepatitis C, Epilepsy, Blindness, Deafness, HIV, and Syphilis. The observational Fitbit study from which current study data were derived followed this same guidance, with no additional exclusion criteria. This study was approved by the UCSF Institutional Review Board. All subjects provided written, informed consent to voluntary research participation.

Procedure

Participants were scheduled for a Longitudinal Brain Aging baseline or annual follow-up research visit, which took place in-person at the UCSF Memory and Aging Center. Participants represent community-dwelling functionally intact older adults living in the Bay Area. These comprehensive visits included cognitive testing, neuroimaging, and questionnaire completion. Preceding or following their standard visit, all subjects were also invited to participate in the observational Fitbit study on an opt-in basis. This study was described as an optional add-on to the primary longitudinal program and aimed to investigate the link between lifestyle factors (i.e., physical activity) and brain health. At the Fitbit study appointment, participants were asked to wear an actigraphy watch (Fitbit^{TM} Flex 2 model) for 30 continuous days on the non-dominant wrist during all waking hours, including both active and sedentary time. They were instructed to charge the device every night and resume wearing it the following morning. A subset of 62 participants who owned smartphones agreed to download the mobile Fitbit app and sync their Fitbit device to the app once per day. The other 113 participants were not expected to complete a daily sync; instead, all Fitbit data was synced to the app by a research coordinator after study completion. Study FAQ sheets were provided to each participant and contained trouble-shooting and syncing details. Activity logs were also distributed as means for participants to record any deviations from the study protocol, such as forgetting to wear, sync, or charge the device on a given day. Research coordinators emphasized the observational nature of the study, and they encouraged participants to go about their daily activities as they usually would. To further minimize self-monitoring effects on behavior, Fitbit activity feedback was reduced as much as possible. Feedback was inherently limited by the minimalist design of Flex 2 model, which does not feature a visual screen display. In addition, all in-app activity tracking tiles were removed and all exercise-related Fitbit and mobile device goals and notifications were disabled. After 30 days of daily use, participants were contacted to return their Fitbit by mail using a provided, prepaid envelope. Interested participants were able to request post-completion summaries of their physical activity metrics. Collected physical activity data was then linked to all relevant standard visit measures captured on the same visit day or within 500 days of the Fitbit study start date. Thirty eight participants who completed Fitbit did not have cognitive testing or neuroimaging completed at their standard visit within this timeframe.

Fitbit Data Collection

Fitbit accounts were individually created for each participant through the mobile app, either on the subject's smartphone or on a research iPad to accommodate any subjects not using a personal cellular device for the study. Each participant was assigned a unique, de-identified username for app sign up, and each Fitbit was directly paired to the participant's respective app account via Bluetooth connection. Daily in-app Fitbit syncing required basic WIFI connection. Performing manual, daily syncs was optimal, as it allowed for minute-level presentation of physical activity data and in-depth analysis of step cadence for this subset of participants. For participants who did not sync every day, Fitbit stores daily aggregate metrics (e.g., daily total step counts and mileage). Upon device return, all Fitbits were charged and synced a final time to capture any aggregate-level data that was not previously uploaded to the app. All Fitbit accounts were then connected to Fitabase, a platform specifically tailored for wearable research data management. All de-identified participant Fitbit data were then exported from Fitabase, cleaned, and analyzed in R.

Measures

Study Adherence

We measured study adherence in two ways. First, we measured each participant's daily adherence wearing the device, which was calculated as the proportion of study days with >100 steps recorded. This step count cutoff was used to identify days when participants likely did not wear the device for any part of the day, following previous study approaches (27). Among the subset of participants who were asked to sync their device to the smartphone app daily, we also calculated the proportion of study days for which any minute-level data was collected, as an indicator of successful syncing events.

Cognitive Functioning

Participants completed a brief neuropsychological battery in person at their parent-study visit. Tests assessed three cognitive domains: memory, executive functioning, and processing speed. Sample based *z*-scores were calculated for individual tests and then averaged within each domain to create a composite *z*-score. The memory composite included the CVLT-II (total immediate recall, long delay free recall, and recognition discriminability) and Benson Figure Recall. The executive functioning composite included a modified version of the Trail Making Test requiring participants to serially alternate between numbers and days of the week (total time to complete), a Stroop interference task (number of correct items in 60 s), phonemic fluency (number of D words in 60 s), design fluency (D-KEFS Condition (1), and digit span backward (longest span). The processing speed composite included computerized visuospatial processing speed (reaction time) tasks previously described elsewhere (28). Higher scores indicate better performance for the memory and executive functioning domains, whereas lower scores indicate better performance for the processing speed domain (i.e., faster times).

Neuroimaging

Participants also completed magnetic resonance imaging (MRI) using a Siemens Prisma 3T scanner. Whole brain T1-weighted images were acquired sagittally using magnetization prepared rapid gradient-echo sequence (TR/TE/TI = 2,300/2.9/900 ms, $\alpha = 9°$) with field of view of 160 × 240 × 256 mm and isotropic voxel resolution of 1 mm3. All T1-weighted images were inspected visually for quality before processing and images with excessive motion or artifact were excluded. The N3 algorithm was used to correct for magnetic field bias (29). SPM12's unified segmentation procedure was used for tissue segmentation (30). Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (DARTEL) was used to create a study-specific template for warping individual participant T1-weighted images (31). Images were normalized and modulated within the study-specific template space using non-linear and rigid-body registration. Smoothing was performed using an 8-mm full width half maximum Gaussian kernel. Linear and non-linear transformations between DARTEL's space and International Consortium of Brain Mapping (ICBM) space were applied to facilitate registration with a brain parcellation atlas. Quantification of volumes was performed by transforming a standard parcellation atlas into ICBM space and summing all gray matter within parcellated regions of interest (32). Total intracranial volume (TIV) was calculated as the sum of gray matter, white matter, and cerebrospinal fluid. This study examined total gray matter volume and medial temporal lobe volume (i.e., bilateral entorhinal, parahippocampal, plus hippocampal volume) with TIV regressed out.

Technology Familiarity and Feedback Questionnaires

Questionnaires were available through the UCSF Qualtrics Web Survey platform and completed at the end of the visit on a research iPad or at home using distributed email survey links. The Technology Familiarity Questionnaire asked questions about participants' prior experiences using computers and technological devices, including if participants: (1) have ever used a wearable tracking device (i.e., Fitbit, Jawbone, Apple Watch), (2) own a "smartphone" (i.e., iPhone, Android), (3) experience difficulty when using computers, and (4) experience anxiety when using a computer, tablet, or smartphone. Questions 1 and 2 offered binary response options "Yes" or "No". Questions 3 and 4 response options were based on a Likert scale from 1 (least affected by difficulty) to 5 (most affected by difficulty). After completion of the Fitbit study, participants completed the Post-Study Feedback Questionnaire, which asked questions about experience using the Fitbit for the duration of their participation. Participants were asked to rate overall: (1) satisfaction with participating, (2) degree of Fitbit interference in day-to-day life, (3) degree of Fitbit comfort over the course of 30 days, (4) degree of difficulty maintaining the Fitbit's charge and using the Fitbit wristband, and (5) degree of change to day-to-day activities caused by wearing the Fitbit. The response options followed a Likert scale from 1 (e.g., not at all interfering, not at all difficult, no change to day-to-day activities) to 5 (e.g., very satisfied, very comfortable to wear). The survey also asked whether subjects would participate in a future wearable devices study (Yes/No).

Statistical Analyses

Descriptive statistics were used to characterize adherence to wearing the Fitbit daily, adherence to syncing the Fitbit daily, and responses to the post-study feedback questionnaire. To examine bivariate associations between adherence and demographic and clinical factors, Spearman correlations and Wilcoxon rank sum tests were used for continuous and categorical variables, respectively. These non-parametric statistical tests were used due to the skewed distribution of adherence rates. For any statistically significant (alpha = 0.05) bivariate relationship with clinical factors, follow up analyses were conducted to covary for demographics. Specifically, due to issues with skew, adherence was dichotomized (<90% adherence vs. ≧90% adherence) and logistic regression was used to examine the specified clinical factor as a predictor of adherence, covarying for age, sex, and education. All analyses were conducted using R, version 4.0.5.

RESULTS

Participant demographic and clinical characteristics are displayed in **Table 1**. Participants were 74 years old on average, majority female with more than college education on average, and mostly non-Hispanic White. This sample was also fairly active, with about 7,000 steps taken per day on average. Among the subset of participants with cognitive data ($n = 137$), a majority of participants were cognitively normal per consensus review. A majority of participants (70%) reported having prior experience with wearables. Participants also reported little-to-no difficulty or anxiety from using technology on average. Of note, the differences in tech-difficulty and tech-anxiety ratings between participants who were and were not asked to sync their device daily were not statistically significant ($p > 0.05$).

TABLE 1 | Demographic and clinical characteristics ($N = 175$).

	Mean (*SD*) or *N* (%)
Age	73.65 (8.68)
Sex (female)	101 (58%)
Years of education	17.62 (1.93)
Race/Ethnicity	
White	151 (86%)
Black/African American	3 (2%)
Asian	20 (11%)
Other	1 (1%)
Average daily steps	6,964 (3,760)
Cognitive status (cognitively normal)[a]	133 (97%)
Memory z-score[a]	0.03 (0.79)
Executive functioning z-score[a]	0.20 (0.66)
Processing speed score[a]	2.59 (1.58)
Smartphone ownership (yes)[b]	103 (85%)
Prior wearable experience (yes)[b]	85 (70%)
Difficulty with technology[b]	0.55 (0.72) [range = 1–4]
Anxiety from technology[b]	0.13 (0.36) [range = 1–3]

[a]$N = 137$. [b]$N = 121$.

Among all 175 participants enrolled in the study, there was a high rate of daily adherence, with participants wearing the Fitbit on an average of 89% of study days (range = 0–100%; IQR = 91–100%). Four participants had 0% adherence. Study records indicate that they reported not wearing the Fitbit for various reasons (e.g., lost the device, interfered with their own personal activity watch). Among the subset of 62 participants who were asked to sync the Fitbit device to the smartphone app on a daily basis, adherence to the protocol was still high. On average, participants were about 85% adherent to the daily syncing protocol (range = 3–100%; IQR = 82–100%).

Among the entire sample of 175 participants, adherence to wearing the Fitbit was not significantly related to age (Spearman's rho = −0.100; $p = 0.186$) or years of education (Spearman's rho = −0.110; $p = 0.146$); however, adherence was related to sex (Wilcoxon rank sum = 3111.5; $p = 0.044$) such that women (mean adherence = 94%) were more adherent than men (mean adherence = 83%). Greater adherence to wearing the Fitbit showed a small effect with better memory performances, but did not reach statistical significance (Spearman's rho = 0.145; $p = 0.090$). Adherence to wearing the Fitbit was not strongly related to executive functioning (Spearman's rho = 0.033; $p = 0.701$) or processing speed (Spearman's rho = −0.125; $p = 0.167$). Adherence was not significantly related to TIV-adjusted brain volumes (total gray matter: Spearman's rho = 0.048, $p = 0.653$; medial temporal lobe: Spearman's rho = 0.086; $p = 0.421$). Adherence was also not strongly related to reported difficulty with technology (Spearman's rho = 0.087; $p = 0.343$), technology-related anxiety (Spearman's rho = −0.055; $p = 0.548$), smartphone ownership (Wilcoxon rank sum = 1014.5; $p = 0.486$), or prior experience using wearables (Wilcoxon rank sum = 1507.5; $p = 0.891$).

Among the 62 participants with daily syncing, adherence to daily syncing was not significantly related to age (Spearman's rho = 0.162; $p = 0.210$), sex (Wilcoxon rank sum = 481; $p = 0.939$), or years of education (Spearman's rho = 0.086; $p = 0.506$). Greater adherence to daily syncing was significantly related to better memory (Spearman's rho = 0.356; $p = 0.019$). Follow-up logistic regression showed that memory remained a significant predictor of daily syncing adherence (OR = 2.69, 95%CI = 1.06–7.74, $p = 0.046$) even after covarying for age, sex, and education (**Figure 1**). Adherence to syncing the Fitbit was not statistically related to executive functioning (Spearman's rho = −0.019; $p = 0.903$) or processing speed (Spearman's rho = −0.076; $p = 0.599$). Adherence to syncing was not strongly related to brain volumes (total gray matter: Spearman's rho = 0.171, $p = 0.349$; medial temporal lobe: Spearman's rho = −0.022; $p = 0.904$). Adherence to syncing the Fitbit was also not strongly related to reported difficulty with technology (Spearman's rho = 0.040; $p = 0.805$), technology-related anxiety (Spearman's rho = 0.060; $p = 0.714$), or prior experience using wearables (Wilcoxon rank sum = 1507.5; $p = 0.891$).

Finally, responses on the study feedback questionnaire were generally positive (**Figure 2**). Both participants who were and were not asked to sync their devices daily reported high satisfaction with their participation in the study (No daily syncing: mean = 4.42/5, $SD = 0.78$, range = 3–5; Daily syncing: mean = 4.48/5, $SD = 0.69$, range = 3–5) and high comfortability with wearing the Fitbit daily (No daily syncing: mean = 4.63/5, $SD = 0.59$, range = 3–5; Daily syncing: mean = 4.66/5, $SD = 0.55$, range = 3–5). On average, participants also reported that the Fitbit contributed little-to-no interference in their day-to-day life (No daily syncing: mean = 1.26/5, $SD = 0.58$, range = 1–4; Daily syncing: mean = 1.17/5, $SD = 0.38$, range = 1–2), low difficulty with charging the device and using the wristbands (No daily syncing: mean = 1.40/5, $SD = 0.56$, range = 1–3; Daily syncing: mean = 1.34/5, $SD = 0.55$, range = 1–3), and little-to-no change in their daily activities as a result of wearing the Fitbit (No daily syncing: mean = 1.05/5, $SD = 0.29$, range = 1–3; Daily syncing: mean = 1.10/5, $SD = 0.31$, range = 1–2). Additionally, all participants (100%) indicated that they would be willing to participate in another study using wearable devices in the future.

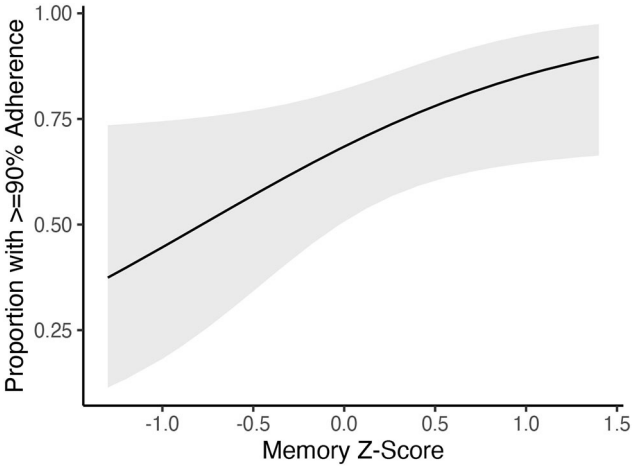


FIGURE 1 | Better memory performance is associated with a greater likelihood of being at least 90% adherent to syncing the Fitbit daily.

DISCUSSION

Given the increasing need to utilize effective behavioral interventions for prolonged health span among older adults, it is important to characterize the feasibility of wearable activity trackers and predictors of wearable use in a study context to inform future protocols and implementation procedures for clinical use. Adherence to wearable use is an important metric that can contribute to interpretation of observational and interventional exercise study effects, yet are often unreported (33). Our findings strongly support the use of wearable activity trackers for studies with older adults, including high overall adherence and satisfaction using the Fitbit in a way that was consistent with our study protocol. Results also suggest that sex and memory functioning may be important predictors of wearable adherence. The latter may be particularly relevant when individuals are required to manually sync devices with a smartphone app on a regular basis. Importantly, all participants

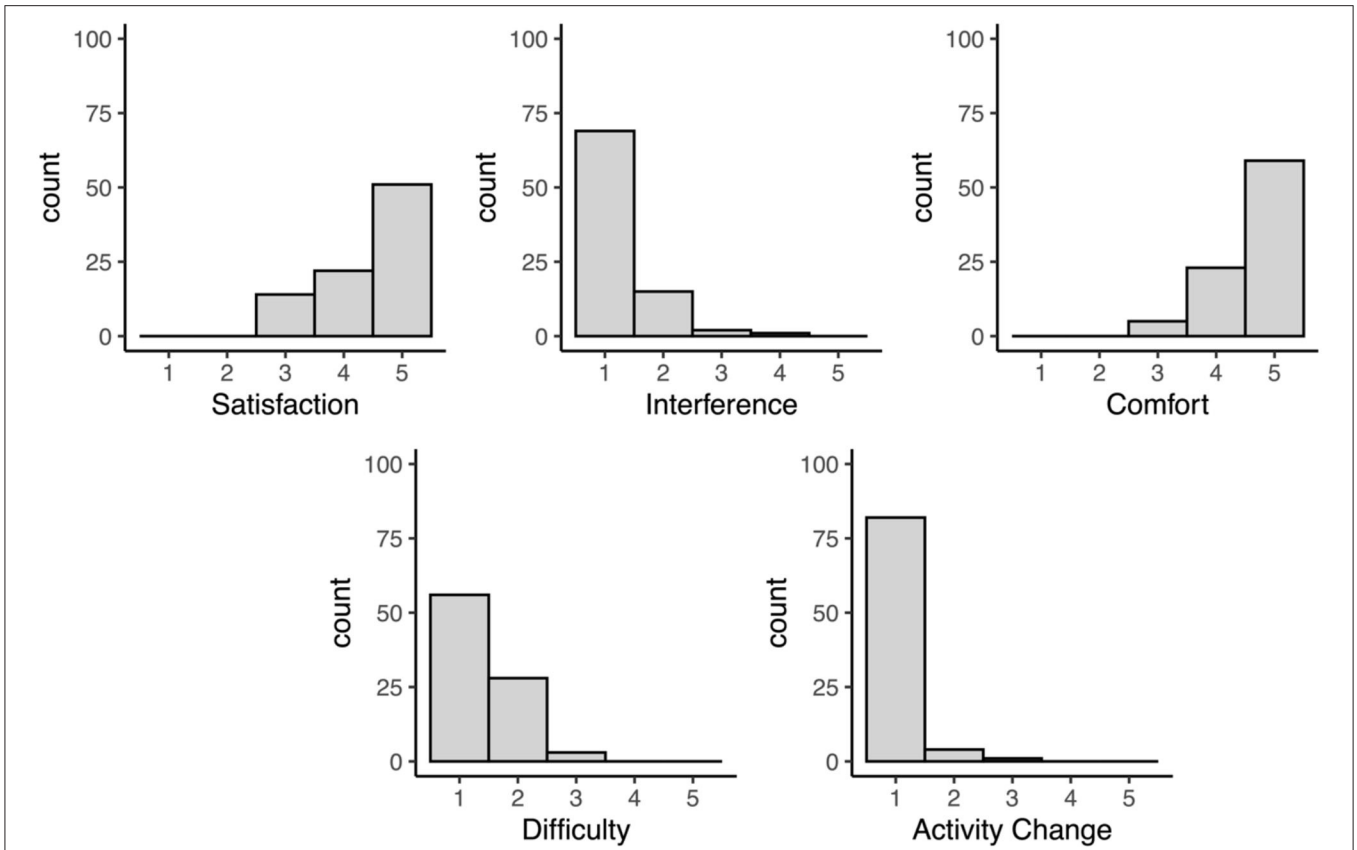


FIGURE 2 | Distributions of responses to the study feedback questionnaire. All items are rated on a scale from 1 (not at all) to 5 (very much).

indicated that they would be willing to use a wearable activity tracker again in another study the future.

Our findings showing high adherence rates for daily wearable use are generally consistent with previous studies in older adults, which have reported daily wearable use on as many as 98% of study days on average (34). Interestingly, we also found that female sex was associated with higher adherence to wearing the device daily. Although few previous studies have examined factors that predict wearable use among older adults, this sex-specific finding is somewhat consistent with at least one other report to our knowledge. Li and colleagues (35) found that women were more likely to be long-term (>6 months) wearable users in naturalistic everyday life than men. Such demographic factors are important to consider to ensure equity in the uptake of beneficial interventions using wearables.

We also identified a novel association between worse memory functioning and poorer adherence to device syncing in a cohort of otherwise functionally intact older adults. No prior studies to our knowledge have examined adherence to daily syncing, which allows for higher resolution data (e.g., steps per minute) to be collected. There was also some cognitive specificity, such that no strong associations between wearable adherence and executive functioning or processing speed were detected. The associations with memory are particularly notable given the high functioning status of our participants and raises potential concerns for wearable use among clinical, cognitively impaired samples. Our results suggest that studies using wearables in cognitively impaired populations may consider implementing daily reminders for use and manual syncing to minimize missing data. Other studies appear to have implemented successful reminder strategies among older adult participant samples, including phone calls or text messages. For example, one physical activity intervention study among older adults with cognitive impairment utilized reminders based on real-time data collection such that reminder calls were provided when no data was transferred to the cloud-based system for 3 consecutive days (36). Conversely, lower wearable adherence may even be used as a digital biomarker of memory status. Further research is needed to extend the work on passively-collected digital biomarkers of cognitive and everyday functioning in aging populations (37).

Finally, results from our study feedback questionnaire are very consistent with previous literature on acceptability of wearable use among older adults. Numerous studies have shown that older adults have high levels of acceptance and willingness to use wearable activity trackers (24, 38, 39). This is not unexpected given that a majority of U.S. older adults now own smartphones (40) and there is a slow but steady rise in uptake of digital health technology in general among this older population (41). This is promising for the integration of digital health technologies into research and clinical settings for improving our monitoring of

modifiable lifestyle factors for maintaining optimal health into older adulthood.

LIMITATIONS

This study was not without limitations. Our study involved a relatively small sample size. While there was a total of 175 subjects that participated in the study, only 62 of these subjects agreed to the continuous, manual syncing process by which they generated minute-level data and could be evaluated for daily syncing adherence. Our study also faced limitations of selection bias. Demographically, our study sample was largely limited to community-dwelling, cognitively healthy, mostly White, and highly educated older adults in the Bay Area. This study sample is demographically reflective of the broader UCSF Longitudinal Brain Aging Program from which participants were recruited and may not be generalizable to older adults in other geographic regions. This should be taken into account when considering the strong technological access and familiarity experiences dominantly reported by participants, and the overwhelmingly positive feedback reported about participating in the wearable study. Given that our study was completely observational, optional, and did not provide compensation, we expect that many of our older adult study volunteers may share a strong motivation to participate in research or other extracurricular activities and a particular interest in physical health and exercise. These characteristics reflect that this group may be more physically active and motivated to use wearables on average than the wider U.S. older adult population.

This possible selection bias also influences the generalizability of our study's findings. Further investigation, starting with the expansion of our wearables study to a broader range of older adults, is needed to better characterize and understand these feasibility, adherence, and memory-based relationships in the context of US older adults across different lifestyles and cognitive domains. The utility and feasibility of wearables in clinical older adult populations cannot be generalized by this study alone. Alongside many others, this study's sample highlights the crucial need for recruiting and including more diverse participant representation in our research across racial, socioeconomic, education, and cognitive diagnosis groups.

CONCLUSIONS

Based on our comprehensive evaluation of study-specific wearable adherence, self-reported feedback, and capture of objective physical activity measures, wearables appear to be feasible and acceptable among community-dwelling older adults; however, consistent with previous studies, there are person-specific factors that likely affect regular daily use. Our findings support the continued use of wearable devices in studies with older adult populations to reliably track physical activity. Notably, adherence to wearable use should be monitored and reminders (e.g., texts, calls) may be particularly helpful in older adults at risk for memory difficulties. Technological upgrades to wearable devices now allow for automatic, Bluetooth-based data collection capabilities (i.e., without the need for manual syncing by the participant). Ideally, studies can utilize these newer wearable models and have devices seamlessly sync to a smartphone or other cloud-based system, which would eliminate participant syncing inconsistencies or errors almost entirely. These recommendations would help to streamline the data collection process and facilitate frequent, consistent device management on the participant side.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by UCSF Human Research Protection Program (HRPP). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

EP contributed to study design, data analysis, interpretation, and manuscript writing and revision. SL contributed to data collection, study design, data analysis, interpretation, and manuscript writing. AV contributed to data interpretation and revision of manuscript for intellectual content. ND and CF contributed to data collection and revision of manuscript for intellectual content. JK contributed to study design, interpretation, and revision of manuscript for intellectual content. KC contributed to study design, data analysis, interpretation, and revision of manuscript for intellectual content. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by NIH-NIA grants K23AG058752 (PI: KC), R01AG072475 (PI: KC), 1R01AG032289 (PI: JK), R01AG048234 (PI: JK), and UCSF ADRC P30 AG062422 (PI: Miller). This work was also supported by Larry L. Hillblom Network Grant (PI: JK; 2014-A-004-NET) and the Alzheimer's Association (PI: KC; AARG-20-683875).

REFERENCES

1. United Nations Department of Economic and Social Affairs. *World Population Ageing 2019*. (ST/ESA/SER.A/444). (2020).
2. Banerjee S. Multimorbidity–older adults need health care that can count past one. *Lancet*. (2015) 385:587–9. doi: 10.1016/S0140-6736(14)61596-8
3. Craig LS, Cunningham-Myrie CA, Hotchkiss DR, Hernandez JH, Gustat J, Theall KP. Social determinants of multimorbidity in Jamaica: application of

latent class analysis in a cross-sectional study. *BMC Public Health.* (2021) 21:1197. doi: 10.1186/s12889-021-11225-6
4. Onder G, Palmer K, Navickas R, Jureviciene E, Mammarella F, Strandzheva M, et al. Time to face the challenge of multimorbidity. A European perspective from the joint action on chronic diseases and promoting healthy ageing across the life cycle (JA-CHRODIS). *Eur J Intern Med.* (2015) 26:157–9. doi: 10.1016/j.ejim.2015.02.020
5. Penedo FJ, Dahn JR. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Curr Opin Psychiatry.* (2005) 18:189–93. doi: 10.1097/00001504-200503000-00013
6. Warburton DE, Bredin SS. Health benefits of physical activity: a systematic review of current systematic reviews. *Curr Opin Cardiol.* (2017) 32:541–56. doi: 10.1097/HCO.0000000000000437
7. Glynn LG, Valderas JM, Healy P, Burke E, Newell J, Gillespie P, et al. The prevalence of multimorbidity in primary care and its effect on health care utilization and cost. *Fam Pract.* (2011) 28:516–23. doi: 10.1093/fampra/cmr013
8. Picco L, Achilla E, Abdin E, Chong SA, Vaingankar JA, McCrone P, et al. Economic burden of multimorbidity among older adults: impact on healthcare and societal costs. *BMC Health Serv Res.* (2016) 16:173. doi: 10.1186/s12913-016-1421-7
9. Makovski TT, Schmitz S, Zeegers MP, Stranges S, van den Akker M. Multimorbidity and quality of life: systematic literature review and meta-analysis. *Ageing Res Rev.* (2019) 53:100903. doi: 10.1016/j.arr.2019.04.005
10. Koorts H, Eakin E, Estabrooks P, Timperio A, Salmon J, Bauman A. Implementation and scale up of population physical activity interventions for clinical and community settings: the PRACTIS guide. *Int J Behav Nutr Phys Act.* (2018) 15:51. doi: 10.1186/s12966-018-0678-0
11. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br J Sports Med.* (2020) 54:1451–62. doi: 10.1136/bjsports-2020-102955
12. Lobelo F, Rohm Young D, Sallis R, Garber MD, Billinger SA, Duperly J, et al. Routine assessment and promotion of physical activity in healthcare settings: a scientific statement from the American Heart Association. *Circulation.* (2018) 137:e495–e522. doi: 10.1161/CIR.0000000000000559
13. Pew Research Center. *About One-in-Five Americans Use a Smart Watch or Fitness Tracker.* (2020). Available online at: https://www.pewresearch.org/fact-tank/2020/01/09/about-one-in-five-americans-use-a-smart-watch-or-fitness-tracker/.
14. Brenner PS, DeLamater JD. Social desirability bias in self-reports of physical activity: is an exercise identity the culprit? *Soc Indic Res.* (2014) 117:489–504. doi: 10.1007/s11205-013-0359-y
15. Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act.* (2008) 5:56. doi: 10.1186/1479-5868-5-56
16. Brickwood KJ, Watson G, O'Brien J, Williams AD. Consumer-based wearable activity trackers increase physical activity participation: systematic review and meta-analysis. *JMIR Mhealth Uhealth.* (2019) 7:e11819. doi: 10.2196/11819
17. Böhm B, Karwiese SD, Böhm H, Oberhoffer R. Effects of mobile health including wearable activity trackers to increase physical activity outcomes among healthy children and adolescents: systematic review. *JMIR mHealth uHealth.* (2019) 7:e8298. doi: 10.2196/mhealth.8298
18. Coughlin SS, Stewart J. Use of consumer wearable devices to promote physical activity: a review of health intervention studies. *J Environ Health Sci.* (2016) 2(6). doi: 10.15436/2378-6841.16.1123
19. Paul SS, Tiedemann A, Hassett LM, Ramsay E, Kirkham C, Chagpar S, et al. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. *BMJ Open Sport Exerc Med.* (2015) 1:e000013. doi: 10.1136/bmjsem-2015-000013
20. Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review. *Maturitas.* (2018) 112:85–93. doi: 10.1016/j.maturitas.2018.03.016
21. Kononova A, Li L, Kamp K, Bowen M, Rikard RV, Cotten S, et al. The use of wearable activity trackers among older adults: focus group study of tracker perceptions, motivators, and barriers in the maintenance stage of behavior change. *JMIR Mhealth Uhealth.* (2019) 7:e9832. doi: 10.2196/mhealth.9832
22. Moore K, O'Shea E, Kenny L, Barton J, Tedesco S, Sica M, et al. Older adults' experiences with using wearable devices: qualitative systematic review and meta-synthesis. *JMIR Mhealth Uhealth.* (2021) 9:e23832. doi: 10.2196/23832
23. Preusse KC, Mitzner TL, Fausset CB, Rogers WA. Older adults' acceptance of activity trackers. *J Appl Gerontol.* (2017) 36:127–55. doi: 10.1177/0733464815624151
24. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User acceptance of wrist-worn activity trackers among community-dwelling older adults: mixed method study. *JMIR mHealth uHealth.* (2017) 5:e8211. doi: 10.2196/mhealth.8211
25. Steinert A, Haesner M, Steinhagen-Thiessen E. Activity-tracking devices for older adults: comparison and preferences. *Universal Access Inf Soc.* (2018) 17:411–9. doi: 10.1007/s10209-017-0539-7
26. Ehn M, Eriksson LC, Akerberg N, Johansson AC. Activity monitors as support for older persons' physical activity in daily life: qualitative study of the users' experiences. *JMIR Mhealth Uhealth.* (2018) 6:e34. doi: 10.2196/mhealth.8345
27. Block VJ, Bove R, Zhao C, Garcha P, Graves J, Romeo AR, et al. Association of continuous assessment of step count by remote monitoring with disability progression among adults with multiple sclerosis. *JAMA Netw Open.* (2019) 2:e190570. doi: 10.1001/jamanetworkopen.2019.0570
28. Kerchner GA, Racine CA, Hale S, Wilheim R, Laluz V, Miller BL, et al. Cognitive processing speed in older adults: relationship with white matter integrity. *PLoS ONE.* (2012) 7:e50425. doi: 10.1371/journal.pone.0050425
29. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging.* (1998) 17:87–97. doi: 10.1109/42.668698
30. Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, Nichols TE. Statistical parametric mapping: the analysis of functional brain images: Elsevier; (2011).
31. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage.* (2007) 38:95–113. doi: 10.1016/j.neuroimage.2007.07.007
32. Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage.* (2006) 31:968–80. doi: 10.1016/j.neuroimage.2006.01.021
33. Liu JY, Kor PP, Chan CP, Kwan RY, Sze-Ki D. The effectiveness of a wearable activity tracker (WAT)-based intervention to improve physical activity levels in sedentary older adults: a systematic review and meta-analysis. *Arch Gerontol Geriatr.* (2020) 91:104211. doi: 10.1016/j.archger.2020.104211
34. St Fleur RG, St George SM, Leite R, Kobayashi M, Agosto Y, Jake-Schoffman DE. Use of Fitbit devices in physical activity intervention studies across the life course: narrative review. *JMIR Mhealth Uhealth.* (2021) 9:e23411. doi: 10.2196/23411
35. Li L, Peng W, Kononova A, Bowen M, Cotten SR. Factors associated with older adults' long-term use of wearable activity trackers. *Telemed J E Health.* (2020) 26:769–75. doi: 10.1089/tmj.2019.0052
36. Vidoni ED, Watts AS, Burns JM, Greer CS, Graves RS, Van Sciver A, et al. Feasibility of a memory clinic-based physical activity prescription program. *J Alzheimers Dis.* (2016) 53:161–70. doi: 10.3233/JAD-160158
37. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/wearable devices opportunity. *NPJ Digit Med.* (2019) 2:9. doi: 10.1038/s41746-019-0084-2
38. Lyons EJ, Swartz MC, Lewis ZH, Martinez E, Jennings K. Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: a randomized controlled pilot trial. *JMIR mHealth uHealth.* (2017) 5:e6967. doi: 10.2196/mhealth.6967
39. O'Brien T, Troutman-Jordan M, Hathaway D, Armstrong S, Moore M. Acceptability of wristband activity trackers among community dwelling older adults. *Geriatr Nurs.* (2015) 36:S21–5. doi: 10.1016/j.gerinurse.2015.02.019
40. Pew Research Center. *Mobile Fact Sheet.* (2021). Available online at: https://www.pewresearch.org/internet/fact-sheet/mobile/.

41. Levine DM, Lipsitz SR, Linder JA. Trends in seniors' use of digital health technology in the United States, 2011-2014. *JAMA*. (2016) 316:538–40. doi: 10.1001/jama.2016.9124

Conflict 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.

The handling editor RM declared a past collaboration with the EP.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 05 August 2022
DOI 10.3389/fdgth.2022.946685

OPEN ACCESS

EDITED BY
Amit Baumel,
University of Haifa, Israel

REVIEWED BY
Nelson Roque,
University of Central Florida, United States
Richard Lipton,
Albert Einstein College of Medicine,
United States

*CORRESPONDENCE
Raeanne C. Moore
r6moore@health.ucsd.edu

SPECIALTY SECTION
This article was submitted to Human Factors and Digital Health, a section of the journal Frontiers in Digital Health

RECEIVED 17 May 2022
ACCEPTED 20 July 2022
PUBLISHED 05 August 2022

CITATION
Moore RC, Ackerman RA, Russell MT, Campbell LM, Depp CA, Harvey PD and Pinkham Amy E. (2022) Feasibility and validity of ecological momentary cognitive testing among older adults with mild cognitive impairment. Front. Digit. Health 4:946685. doi: 10.3389/fdgth.2022.946685

Feasibility and validity of ecological momentary cognitive testing among older adults with mild cognitive impairment

Raeanne C. Moore[1]*, Robert A. Ackerman[2], Madisen T. Russell[2], Laura M. Campbell[3], Colin A. Depp[1,4], Philip D. Harvey[5,6] and Amy E. Pinkham[2]

[1]Department of Psychiatry, University of California San Diego, La Jolla, CA, United States, [2]Department of Psychology, School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, TX, United States, [3]San Diego State University/University of California San Diego Joint Doctoral Program in Clinical Psychology, San Diego, CA, United States, [4]Veterans Affairs San Diego Healthcare System, San Diego, CA, United States, [5]Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, Miami, FL, [6]Research Service, Bruce W. Carter VA Medical Center, Miami, FL, United States

It is critical to intervene early in the mild cognitive impairment (MCI) stage of the Alzheimer's disease trajectory, but traditional cognitive testing methods are costly, burdensome, and difficult to access. We examined adherence and validity data to a 30-day self-administered ecological momentary cognitive testing protocol among a sample of older adults with MCI and cognitively normal controls to evaluate feasibility, tolerability, and initial validity in comparison to standard neuropsychological tests. Participants included 48 participants with MCI (Mean age = 72 years, SD = 7 years) and 46 demographically-matched cognitively normal (NC) control participants (Mean age = 70 years, SD = 7 years). Participants completed traditional neuropsychological testing to determine MCI status, followed by 30 days of remote ecological momentary cognitive testing. Ecological momentary assessment (EMA) surveys were administered 3 times per day for 30 days (possible total = 90), and mobile cognitive tests were administered every other day (for a total of 15 administrations). Mobile cognitive tests included the Variable Difficulty List Memory Test (VLMT; measure of learning and memory), Memory Matrix (measure of visual working memory), and the Color Trick Test (measure of executive function). EMA and mobile cognitive test adherence, fatigue effects, mobile cognitive test performance and group differences, and psychometrics (reliability, convergent validity, ceiling effects, and practice effects) were examined. Overall mean-level adherence to the mobile cognitive tests was 85% and did not differ by MCI status. The reliability of stable between-person individual differences for the VLMT and Memory Matrix were very high. Moreover, although the reliability of within-person change for Memory Matrix was adequate, the corresponding reliability for VLMT was somewhat low. Averaged performance on the mobile cognitive tests was correlated with lab-based tests measuring the same construct. Participants with MCI performed worse than NCs on the VLMT and Color Trick Test, and there was no evidence of fatigue effects for these two tests. These findings support the feasibility and potential for ecological momentary cognitive testing to support clinical trials and for measuring cognitive changes over time in persons with increased risk for Alzheimer's disease such as those with MCI.

KEYWORDS

ecological momentary assessment, ambulatory assessment, smartphones, Alzheimer's disease, adherence, psychometrics

1 Introduction

Research that examines cognitive functioning has traditionally taken place in a lab with paper and pencil neuropsychological testing; however, there are barriers with this method, including high cost, time burden, and access to testing locations which are limited by transportation and uneven distribution in rural or remote areas. As a result, neurocognitive testing is infrequently repeated, if at all. Ecological momentary cognitive tests (EMCTs), which are brief and repeatable cognitive assessments that are self-administered *via* smartphone in participants' own environments, may be a valuable complement to traditional neuropsychological testing that can help overcome some of these barriers (1–4).

There are several advantages to EMCTs that may make them well suited for use in clinical trials. Cognition can fluctuate from day to day, which makes it is difficult to determine what should be considered a real change on neuropsychological testing from one time point to another. This is particularly problematic when trying to examine improvement over time (e.g., recovery from stroke) or cognitive decline as seen in Alzheimer's disease and related dementias. Alzheimer's disease is the most common cause of dementia in older adults (5) and places significant financial and emotional burden on affected families, not to mention the financial impact on healthcare systems. Therefore, it is no surprise that there are currently hundreds of ongoing clinical trials aimed at prevention of and intervention in Alzheimer's disease and related dementias (6).

To date, pharmacological interventions have been slow to show reductions in cognitive decline, and no treatments have been able to reverse cognitive decline despite some evidence for slowing disease progression; however, many of these studies use less-than-optimal cognitive outcome measures. For example, the Alzheimer's Disease Assessment Scale–Cognitive Subscale (ADAS-Cog) has been shown to have significant ceiling effects in those with normal cognition and mild cognitive impairment (MCI) and there are concerns about its ability to detect cognitive changes early in the disease course (7–9). Given that EMCTs can be given over multiple days, EMCTs may be a cost effective and time efficient method to establish a more accurate baseline for cognitive functioning and to detect person-specific changes more sensitively over time. Such procedures could also allow for dynamic titration of difficulty in order to more effectively probe variation in performance.

EMCTs can also be paired with other technologies such as ecological momentary assessment (EMA) or wearable devices (e.g., actigraphy to objectively assess physical activity and sleep). Therefore, observational studies or interventional studies can examine how mood, activities, sleep, and other fluctuating daily-life factors associate with cognition over time without relying on retrospective recall, which is particularly relevant to persons with memory impairments (e.g., 10, 11). Utilizing EMCTs to examine cognition in a person's everyday life with different contextual variables could lead to person-specific intervention strategies (4).

Additionally, the use of EMCT may reduce the number of in-person visits, which could reduce the burdens of time and transportation, particularly for participants that live in rural areas and older adults with mobility limitations. The tradeoff is that technology familiarity may impact one's ability to engage in EMCTs and is something to be mindful of in this group. However, a study conducted in 2021 by the Pew Research Center found that 83% of those aged 50–64 own a smartphone and 61% of adults aged 65 + own a smartphone, indicating that the majority of older adults are already engaged with smartphone technology (12). To date, there have been a handful of studies by other groups utilizing smartphone-based mobile cognitive testing among cognitively normal older adults (e.g., 10, 13, 14) and older adults with MCI (e.g., 15, 16), all of which have demonstrated feasibility, good adherence, and promising initial psychometric properties for use of these tests in this population.

Despite the clear appeal of EMCT in aging research, there are some current limitations. For example, a recent systematic search and evaluation found that the majority of currently-available commercial-grade app-based tools to assess cognition lack validity data for their assessments (17). This is concerning, as an absence of validity data in these tools could lead to unreliable information about possible cognitive impairment. Therefore, we present adherence and validity data in a group of older adults with and without MCI for three NeuroUX EMCTs assessing the domains of memory and executive functioning: 1) Variable Difficulty List Memory Test (VLMT), which is a verbal list-learning test in which we administered 6-word, 12-word, and 18-word versions; 2) Memory Matrix, a visual working memory task; and 3) Color Trick Test, an executive functioning task examining inhibition using a Stroop-Type paradigm. The aims of the study were to examine the 1) adherence to the 30-day EMCT protocol, 2) fatigue effects, 3) EMCT task performance and group differences, and 4) EMCT psychometrics, including reliability, convergent validity (compared to traditional neuropsychological tests), ceiling effects, and practice effects.

2 Materials and methods

2.1 Participants

Participants were English-proficient individuals aged 50 or older who met criteria for any subtype of mild cognitive impairment (MCI) using Jak/Bondi criteria, which require performance of one standard deviation below normative

expectations on two different assessments within a single cognitive domain (i.e., memory, attention, language, executive functioning), or cognitively normal (NC) control participants. Exclusion criteria included: (1) presence or history of medical or neurological disorders that may affect brain function (e.g., stroke, epilepsy, Parkinson's disease), (2) presence of dementia, (3) history of unconsciousness for a period greater than 15 min, (4) significant impairment of vision (e.g., blindness, glaucoma, vision uncorrectable to 20/40, color blindness) or hearing (e.g., hearing loss) that would interfere with their ability to complete the study protocol, (5) presence of intellectual disability (defined as IQ < 70), (6) current diagnosis of substance use disorder, (7) or presence or history of a psychotic disorder or bipolar disorder.

Data were collected across three sites between December 2020 and December 2021: The University of Texas at Dallas (UTD), University of California San Diego (UCSD), and University of Miami Miller School of Medicine (UM), resulting in a total of 94 participants (48 MCI, 46 NC). UTD participants were recruited from community advertisements and previous participation in aging-related research studies at the Center for Vital Longevity at UTD. UCSD participants were recruited from word of mouth and posting in the Stein Institute for Successful Aging monthly newsletter. UM participants were recruited from the clinical programs at the Miller School of Medicine Memory Disorders Center, the Florida ADRC, and through advertisements and previous study participants.

2.2 Procedures

The study was approved by each University's respective Institutional Review Board, and all participants provided written informed consent. After a brief phone screen, participants completed a baseline visit either remotely *via* Microsoft Teams or Zoom or in-person. During the baseline visit, participants completed a neuropsychological battery. Research staff held a bachelor's degree or higher, and were trained over the course of several weeks, within and across sites, to administer and score the neuropsychological tests accurately. Jak/Bondi diagnostic criteria for MCI were applied to the neuropsychological test data to determine MCI status. The Jak/Bondi diagnostic criteria show a good balance of sensitivity, specificity, and reliability compared to other conventional MCI criteria (18). Study eligibility, all neuropsychological test scores, and diagnoses were reviewed by the first author (RCM). Once eligibility and group status were confirmed, staff contacted participants to set up their smartphones for the EMCT period. Participants could either complete the EMCTs using their personal smartphone or, if they requested or did not own a smartphone, they were provided with a study-owned Android smartphone. Those using study-provided smartphones were trained to operate the device and given a user manual to reduce technological issues. Participants were trained on the EMCT protocol and completed a mock EMA survey and mobile cognitive testing session to allow for technical questions and troubleshooting.

For the following 30 days, participants completed the EMCT protocol using the NeuroUX platform (19). Participants were sent text message notifications to take the EMA surveys three times per day. Every other day, participants were asked to complete the three different mobile cognitive tests (i.e., Variable Difficulty List Memory Test, Memory Matrix, Color Trick Task) of varied difficulty along with each of their EMA surveys. The mobile cognitive tests were counterbalanced throughout the EMA period by test type and difficulty level, resulting in a total of 5 easy, 5 medium, and 5 hard conditions of each of the three mobile cognitive tests (see **Figure 1**). To encourage EMA adherence and help troubleshoot any difficulties, researchers contacted participants if they missed more than three surveys in a row. Participants were compensated up to $190 total for completing the baseline visit ($50) and EMCT sessions (EMA questions only – $0.88; EMA + mobile cognitive tests – $2.25).

2.2.1 Remote visit task modifications

Due to evolving restrictions on in-person data collection during the height of the COVID-19 pandemic, some individuals participated in-person ($n = 28$) whereas others participated *via* remote visits ($n = 66$). For remote appointments, all tasks were completed *via* video conferencing using Microsoft Teams or Zoom meetings and required minimal modification. Participants were asked to complete the visit in a quiet environment away

Mobile Cognitive Test	Study Day (and administration order)															
	1	3	5	7	9	11	13	15	17	19	21	23	25	27	29	
VLMT	1	2	3	2	1	3	3	1	2	2	3	1	1	2	3	
Color Trick	2	1	2	3	3	1	2	2	1	3	1	3	2	3	1	
Memory Matrix	3	3	1	1	2	2	1	3	3	1	2	2	3	1	2	

FIGURE 1
Protocol of mobile cognitive testing administration. *Note.* Difficulty levels are depicted as green (easy), yellow (medium), and red (hard).

from distractions (e.g., away from other individuals, powering off/silencing unrelated devices) and a screening measure was completed to ensure participants could hear the researcher well and see the PowerPoint materials on their desktop, laptop, or iPad. Researchers also asked participants to refrain from utilizing any performance aids, such as writing down stimulus items, searching for answers on the internet, or seeking help from other individuals.

Tasks that were typically administered orally (Hopkins Verbal Learning Test – Revised (HVLT-R), Number Span Test: Forward) were implemented as is. Tasks that required visual presentations (Wide Range Achievement Test-4 (WRAT-4), Delis-Kaplan Executive Function System Color-Word Interference Test (D-KEFS), Brief Visuospatial Memory Test - Revised (BVMT-R)) were administered *via* video call using a PowerPoint screenshare function. Prior to the baseline visit, research staff instructed participants to prepare four blank pieces of printer paper for the BVMT-R task. Additionally, during the BVMT-R task, after the participant completed each trial drawing, the researcher asked the participant to hold the paper in front of the camera so that a photo could be taken, then instructed them to flip the paper over and place it out-of-sight before beginning the next trial.

2.3 Measures

2.3.1 Traditional Neuropsychological measures (lab or remote administered at baseline)

To determine premorbid IQ, the Wide Range of Achievement Test 4 (WRAT-4; 20) word reading subtest was used. The Montreal Cognitive Assessment-BLIND version 7.1 (MoCA-BLIND; 21) was administered to screen for the presence of dementia using established cutoff scores. This version of the MoCA was used for participants who completed virtual visits as well as participants who completed in-person visits. To determine MCI eligibility, the following tests were administered: Hopkins Verbal Learning Test – Revised (HVLT-R; 22), Brief Visuospatial Memory Test – Revised (BVMT-R; 23), Oral Trail Making Test- A and B (24), Digit Span Forward (25), Verbal Fluency – Letter and Animals (25), Multilingual Naming Test (MINT; 26), Number Span Test: Forward (25), and the D-KEFS-Color Word Interference Test (27).

For validity analyses in the current study, we used non-demographically adjusted scores from the HVLT-R (verbal memory), BVMT-R (visual memory), Letter-Number Span (attention/working memory), and D-KEFS Color-Word Interference Test (executive function).

2.3.2 EMA surveys

Each EMA survey asks participants questions about their daily functioning, including where they are (dichotomized as "at home" versus "away") and who they are with (dichotomized as "alone" versus "with others"). The EMA surveys also generally queried participants' mood, cognitive concerns, substance use, pain, and sleep as additional questions but data are not reported here.

2.3.3 Mobile cognitive tests

See **Table 1** for a list of the mobile cognitive tests, the cognitive domains assessed, completion times, and screenshots.

2.3.3.1 Mobile variable difficulty list memory test (VLMT)

The VLMT has been described and validated by Parrish et al. (2020). For this task, participants are presented with a list of words (list length varies between 6, 12, or 18) on 3 separate trials for 30 s each. Immediately following each trial, participants are shown target and distractor words one-by-one and asked to identify whether the word appeared on the list (matched number of target and distractor words presented). Each trial is scored by number of words correctly recalled or based on a percentage of correct target items (range 0%–100%).

2.3.3.2 Memory matrix

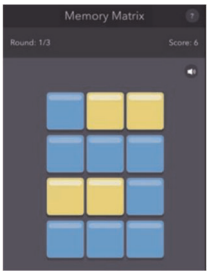
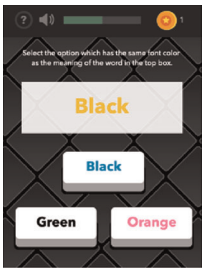
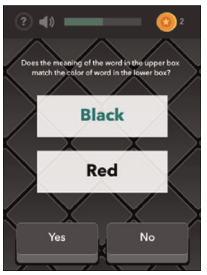
During the Memory Matrix task, participants are presented with a matrix of blue tiles. A pattern of yellow tiles is then displayed, and the participant is asked to memorize the location of the yellow tiles. After 1.5 s, the yellow tiles are then switched back to blue, and the participant is asked to tap the tiles that were previously yellow. Matrix sizes are varied across administration days so that participants complete 5 days of 6-tile matrices, 5 days of 12-tile matrices, and 5 days of 18-tile matrices. Each administration also includes three trials of 9 patterns each. Participants earn 1 point for each pattern correctly recreated for a score range of 0–9 per trail and 0–27 per administration.

2.3.3.3 Color trick

The Color Trick task was modelled after the Stroop-type paradigm (Stroop, 1935). Participants completed three different conditions of this task (Meaning-to-Meaning, Meaning-to-Color, Yes-No Mechanic) divided across the 15 days of EMCTs such that each condition was administered 5 times. Each condition includes three trials of 9 items/questions for a total of 27 items per administration. Each item in each condition shows a word in an upper box of the smartphone screen and between 1 and 3 words on the lower half of the screen. The font colors and actual meanings of the upper and lower words are either the same or different colors. The first condition type is *Meaning-to-Meaning*, in which participants are presented with one word in an upper box on their screen and 2–3 word choices on the lower half of their screen and asked to select the word choice that has the same meaning as the word in the top box (e.g., matching top word "pink" with bottom word "pink"). The second condition type

10.3389/fdgth.2022.946685

TABLE 1 Mobile cognitive tests.

Mobile Cognitive Test	Cognitive Domain Assessed	Time to Complete	Screenshot of Task
Variable Difficulty List Memory Test (VLMT)	Recognition Memory	30 s for list presentation	
Memory Matrix	Visual Working Memory	Variable; 3 trials; approximately 1–2 min (Mean completion time: 1.5 min	
Color-Trick: Meaning-to-Meaning	Executive Function	Variable; 3 trials; approximately 1.5–3 min (Mean completion time: 2.25 min)	
Color-Trick: Meaning-to-Color	Executive Function	Variable; 3 trials; approximately 2–3.5 min (Mean completion time: 2.75 min)	
Color-Trick: Yes-No Mechanic	Executive Function	Variable; 3 trials; approximately 2.5–3.5 min (Mean completion time: 3 min)	

is *Meaning-to-Color*, in which participants are presented with one word in an upper box on their screen and 2–3 word choices on the lower box of their screen and asked to select the word choice that has the same font color as the meaning of the word in the top box (e.g., matching top word "pink" with bottom word printed in pink font). The third condition

type is *Yes-No Mechanic*, in which participants are presented with one word in an upper box on their screen and one word in a lower box on their screen, and asked, "Does the meaning of the word in the upper box match the color of the word in the lower box?" and the participant can choose either "yes" or "no." Each trial is scored based on the number of items correct (range 0–9) and average response time for correct items.

2.4 Statistical analyses

Demographic differences between groups (MCI+ vs. NCs) and administration formats (in-person vs. remote) were assessed using independent samples *t*-tests or Chi-Square tests (x^2) as appropriate. Adherence was calculated as the percentage of EMA surveys completed by the total number possible (90), as well as the percentage of each of the three mobile cognitive tests completed by the total number possible (15 each). Adherence differences between groups and administration formats were assessed using independent samples *t*-tests. In addition, Pearson's *r* correlations were used to estimate relationships between adherence and demographic differences.

To further assess whether adherence changed over time, we computed missing data variables for the EMCTs that denoted whether participants skipped a test that they were scheduled to take (0 = completed test, 1 = missed test). We then estimated fatigue effects for each of the EMCTs (i.e., whether participants' odds of missing a test was greater on later versus earlier study days) using growth-curve models specified with multilevel logistic regression model in Mplus v. 8.4 (28). Using maximum likelihood estimation, each model regressed participants' log odds of missing a test on time (scaled such that 0 is the midpoint of the EMA period and a one-unit change corresponds to the total change in the log odds of missing a test across the EMA period), MCI status (effect coded such that −1 = NC and 1 = MCI), and the interaction of time with MCI status. Each model also included an unstructured variance-covariance matrix for the random intercepts and slopes. These specifications enabled us to estimate the average probability of missing a test across the EMA period (*via* the threshold value[1]), the average fatigue effect in the sample (*via* the first-order effect of time[2]), whether the average log odds of missing a test across the EMA period differs between NC and MCI (*via* the first-order effect of MCI status), and whether fatigue effects differ between NC and MCI (*via* the interaction of time and MCI status).

We next investigated participants' average performance on the EMCTs across the EMA period. To evaluate group differences (i.e., NC vs. MCI) on EMCT performance across trials, we conducted independent samples t-tests.

The final sets of analyses provided additional psychometric evidence for each EMCT – namely, reliability, convergent validity, ceiling effects, and practice effects. We first calculated Intraclass Correlation Coefficients (ICCs) for each EMCT to quantify the proportion of variance in the tests attributed to trait vs. state components across the EMA period. We then used generalizability theory (see Ref. 29) to estimate the reliability of stable between-person individual differences (R_{KF}) as well as the reliability of within-person change (R_C) in the EMCT measures that contained multiple trials (i.e., list-learning and matrix memory). These analyses used the Minimum Norm Quadratic Unbiased Estimate (MINQ) method within SPSS v. 26 to estimate the variance components linked to the factorial combination of participant, day, and item (where only participant was treated as a random factor).

We then evaluated the convergent validity evidence for each EMCT by estimating correlations between participants' average performance on a given EMCT and their parallel performance on a similar lab-based measure. Ceiling effects for each EMCT were subsequently evaluated by counting the number of participants who earned the maximum score consistently across the EMA period. Practice effects for each of the EMCTs (i.e., whether participants' performance on the measures systematically changed across the course of the EMA period) were then assessed *via* growth-curve models specified with linear multilevel regression in Mplus v. 8.4 (28). Using maximum likelihood estimation with robust standard errors, each model regressed participants' test scores on time, MCI status, and the interaction of time with MCI status (we used the same scaling for time and MCI status as our analyses investigating fatigue effects). When sufficient variability was present, we specified an unstructured variance-covariance matrix for the random intercepts and slopes. These specifications enabled us to estimate participants' average performance on the EMCT (*via* the intercept), the average practice effect in the sample (*via* the first-order effect of time), whether average levels of performance for an EMCT differs between NC and MCI (*via* the first-order effect of MCI

[1]Participants' average probability of missing an EMCT item was computed as $1/(1 + \exp(\tau))$, where τ = threshold.

[2]In cases where there was evidence for a fatigue effect, we used the following formula to determine participants' average probability of missing an EMCT item at the beginning (time = −0.50) and end (time = 0.50) of the study:

$$\pi = \frac{\exp[-(\tau) + \beta_1 X_i]}{1 + \exp[-(\tau) + \beta_1 X_i]}$$

where τ = threshold, β1 = Slope reflecting fatigue effect, and Xi = the specific value of time.

status), and whether practice effects differ between NC and MCI (*via* the interaction of time and MCI status).

3 Results

3.1 Sample characteristics

Demographic and clinical characteristics by MCI status are displayed in **Table 2**. Groups were comparable on demographics and did not significantly differ on age, sex, race, ethnicity, or years of education. Groups were also comparable on type of phone used, with 55% of MCI participants and 62% of NCs using iPhones, while the other participants used Android devices (Chi-Square = 11.3, $p = 0.334$; **Supplementary Table S1**).

Sixty-six participants completed the lab-based neuropsychological visit remotely *via* telehealth, while 28 completed this visit in-person. There were no demographic differences for participants who completed this visit remotely versus in-person except for fewer Hispanic individuals in the in-person group ($\chi^2 = 6.4$, $p = 0.01$). Additionally, there were no significant differences in MCI status ($\chi^2 = 0.59$, $p = 0.44$) or performance on any of the neuropsychological tests based on remote vs. in-person participation (all $ps > 0.09$).

3.2 Adherence

For the whole sample, adherence to EMA surveys was 86% (SD = 15.8%; range = 24%–100%). In regard to the mobile cognitive tests, adherence to the VLMT was 84% (SD = 19.3%; range = 7%–100%), adherence to Memory Matrix was 85% (SD = 18%; range = 20%–100%), and adherence to Color Trick was 85% (SD = 17%; range = 13%–100%). Adherence to EMA surveys did not differ by diagnostic status, $t = 1.21$, $p = 0.23$, and neither did completion rates of the mobile cognitive tests (VLMT: $t = 0.83$, $p = 0.41$; Memory Matrix: $t = 1.56$, $p = 0.12$; Color Trick: $t = 0.97$, $p = 0.33$). Further, there was no difference in EMA adherence or mobile cognitive test completion rates for participants who completed the lab-visit remotely or in-person (all $ps > 0.19$). Age, education, and estimated IQ (measured by the WRAT-4) did not correlate with adherence to EMCTs nor with percentage of surveys completed at home or alone, except for a small negative correlation between years of education and completion of the Memory Matrix test. Higher adherence was positively correlated with answering more surveys when home and when alone (see **Table 3**).

3.3 Fatigue effects

Because we used varying list lengths for the VLMT, we included list length (*via* two effect-codes that treated the 18-word list length as the reference group) and its interaction with time as covariates in the VLMT fatigue effect analyses. On average, participants' probability of missing (i.e., failing to complete) a list-learning item was 0.08 for Trial 1 (threshold = 2.40, SE = 0.23, $p < 0.001$), 0.08 for Trial 2 (threshold = 2.38, SE = 0.22, $p < 0.001$), and 0.09 for Trial 3 (threshold = 2.34, SE = 0.22, $p < 0.001$), where trials refer to trials within the same test (e.g., for the VLMT, there were three trials administered at each session). We found no evidence of a fatigue effect for Trial 1 (logit = 0.46, SE = 0.53, $p = 0.39$; OR = 1.58), Trial 2 (logit = 0.56, SE = 0.53, $p = 0.29$, OR = 1.74), or Trial 3 (logit = 0.52, SE = 0.52, $p = 0.32$, OR = 1.68). Moreover, MCI participants did not significantly differ from controls on their log odds of missing a list-learning item vs. not missing the item for Trials 1, 2, or 3 (all $p's > 0.12$) or their fatigue effects for Trials 1, 2, or 3 (all $p's > 0.57$).

Similar to the VLMT, participants' average probability of missing a Memory Matrix item across the EMA period was 0.08 for Trial 1 (threshold = 2.45, SE = 0.21, $p < 0.001$), 0.08 for Trial 2 (threshold = 2.43, SE = 0.21, $p < 0.001$), and 0.08 for Trial 3 (threshold = 2.42, SE = 0.21, $p < 0.001$). Unlike the VLMT, however, we found evidence of fatigue effects for the Memory Matrix items across the three trials. In particular, participants' odds of missing a Memory Matrix item vs. not missing a Memory Matrix item from the beginning to the end of the EMA period increased approximately 3.23-fold for Trial 1 (logit = 1.174, SE = 0.52, $p = 0.023$), approximately 3.47-fold for Trial 2 (logit = 1.244, SE = 0.51, $p = 0.014$), and approximately 3.42-fold for Trial 3 (logit = 1.231, SE = 0.50, $p = 0.014$). That is, whereas participants' probability of missing a Memory Matrix item was 0.05 at the beginning of the EMA period for Trials 1, 2, and 3, their probability of missing a Memory Matrix item at the end of the EMA period was 0.13 for Trials 1 and 2 and 0.14 for Trial 3. Nonetheless, MCI participants did not significantly differ from controls on their log odds of missing a Memory Matrix item vs. not missing the item for Trials 1, 2, or 3 (all $p's > 0.06$) or on their fatigue effects for Trials 1, 2, or 3 (all $p's > 0.59$).

Participants' average probability of missing a Color Trick item across the EMA period was 0.09 for Trial 1 (threshold = 2.285, SE = 0.19, $p < 0.001$), 0.09 for Trial 2 (threshold = 2.269, SE = 0.19, $p < 0.001$), and 0.09 for Trial 3 (threshold = 2.256, SE = 0.19, $p < 0.001$). We found no evidence of a fatigue effect for Trial 1 (logit = 0.299, SE = 0.46, $p = 0.514$, OR = 1.35), Trial 2 (logit = 0.242, SE = 0.46, $p = 0.598$, OR = 1.27), or Trial 3 (logit = 0.269, SE = 0.45, $p = 0.55$, OR = 1.31). MCI participants also did not significantly differ from controls on their log odds of missing a Color Trick item vs. not missing the item for Trials 1 to 3 (all $p's > 0.07$) or on their fatigue effects for Trials 1 to 3 (all $p's > 0.20$).

10.3389/fdgth.2022.946685

TABLE 2 Demographics and clinical characteristics by mild cognitive impairment (MCI) status.

	MCI ($n = 48$)	Cognitively Normal (CN) ($n = 46$)	Test-statistic[a]	p-value
Demographics				
Age in years, *M* (SD); range	72 (7.7); 54–85	70 (6.6); 60–87	0.96	0.34
Sex (% F)	27 (56%)	34 (73%)	3.22	0.07
Race (%)				
White	45 (94%)	41 (89%)	4.81	0.09
Black/African American	1 (2%)	5 (11%)		
More than one race	2 (4%)	0 (0%)		
Ethnicity (% Hispanic/Latino)	8 (17%)	5 (11%)	0.66	0.42
Education (years), *M* (SD)	16.1 (2.5)	16.2 (2.1)	0.26	0.80
Premorbid IQ (WRAT-4 SS), *M* (SD)	110.2 (15.1)	109.9 (12.0)	0.11	0.91
Employment status				
Retired	26 (54%)	32 (70%)	2.64	0.45
Umemployed	2 (4%)	1 (2%)		
Part-time employment or volunteer	14 (29%)	8 (17%)		
Full-time employment or volunteer	6 (13%)	5 (11%)		
Residential Status				
Independent/Financially Responsible	48 (100%)	44 (96%)	2.13	0.14
Independent/Not Financially Responsible	0 (0%)	2 (4%)		
Smartphone used for study				
Personal iPhone	27 (56%)	31 (67%)	4.36	0.11
Personal Android	17 (36%)	15 (33%)		
Study Loaned Android	4 (8%)	0 (0%)		
Remote Participation	32 (67%)	34 (74%)	0.59	0.44
Lab-Based Neuropsychological Scores[b]				
Hopkins Verbal Learning Test (HVLT) – Immediate Recall	40.7 (9.9)	51.4 (10.0)	5.24	<0.001
Brief Visuospatial Memory Test-R (BVMT-R) – Immediate Recall	50.8 (9.7)			
Letter Number Span	45.1 (8.9)	49.6 (9.3)	2.4	0.02
D-KEFS Interference	54.7 (11.4)	56.4 (10.0)	0.73	0.47
Mobile Cognitive Tests – Mean aggregated scores[c]				
VLMT 6 words (% Correct)	94.5 (5.7)	95.6 (5.0)	1.04	0.30
VLMT 12 words (% Correct)	85.0 (8.5)	87.3 (6.1)	1.50	0.14
VLMT 18 words (% Correct)	76.6 (9.3)	80.8 (6.9)	2.41	0.02
Memory Matrix (Total Score)	7.3 (0.93)	7.4 (0.83)	0.97	0.33
Color Trick: Meaning-to-Meaning (Total Score)	8.2 (0.51)	8.5 (0.46)	2.13	0.04
Color Trick: Meaning-to-Color (Total Score)	8.6 (0.41)	8.7 (0.42)	1.50	0.07
Color Trick: Yes-No Mechanic (Total Score)	8.6 (0.41)	8.7 (0.28)	1.19	0.24

Note. Values are presented as mean (SD) or *n* (%).
[a]*T*-tests for continuous variables; Chi square for dichotomous variables.
[b]Demographically-adjusted *T*-Scores from lab-based neuropsychological scores are reported.
[c]Raw scores are reported.

3.4 EMCT performance and group differences

Table 2 presents average mobile cognitive test performance for the MCI and NC groups across the EMA period. As expected, participants generally committed more errors on the VLMTs when the list length was greater. Participants' performance on the Memory Matrix and Color Trick tests was also quite high. While participants with MCI scored lower on all EMCTs, they only performed significantly worse than the NC participants on the 18-word VLMT and the Color Trick: Meaning-to-Meaning task.

TABLE 3 Correlations between adherence and demographic characteristics in the whole sample ($N = 94$).

	Age	Education	Estimated IQ	% surveys completed at home	% surveys completed alone
EMA Adherence	−0.122	−0.167	−0.129	0.582**	0.286**
VLMT Adherence	−0.029	−0.023	−0.075	0.536**	0.249*
Memory Matrix Adherence	−0.158	−0.205*	−0.129	0.511**	0.274**
Color Trick Adherence	−0.117	−0.114	−0.132	0.363**	0.381**

Note. *$p < 0.05$; **$p < 0.01$.

We also examined performance differences by phone type. In the overall sample, there were no significant performance differences based on phone type (**Supplementary Table S2**). When examining the effects of both phone type and group (and their interaction) on mobile cognitive test performance, no main effects were found for the VLMT 6- or 12-word list, Memory Matrix, Color Trick Meaning-to-Color, or Color Trick Yes-No Mechanic (all p's > 0.05). Further, there were no significant interactions between phone type and group on any of the mobile cognitive tests (all p's > 0.05). For the VLMT 18-word list, a main effect for group was observed, such that NC participants performed better than participants with MCI ($F = 6.53$, $p = 0.01$); there was no main effect for phone type ($F = 0.53$, $p = 0.47$). Lastly, there was a main effect for group on Color Trick Meaning-to-Meaning, such that MCI participants performed worse than NC participants ($F = 5.23$; $p = 0.03$), but there was no main effect for phone type ($F = 0.11$, $p = 0.74$).

3.5 EMCT psychometrics: Reliability, convergent validity, ceiling effects, and practice effects

3.5.1 Psychometric evidence for VLMT

Aggregated across trials, the Intraclass Correlation Coefficients (ICCs) for each trial length of the VLMT were 0.22, 95% CI [0.11, 0.32] for the 6-word list, 0.33, 95% CI [0.22, 0.44] for the 12-word list, and 0.32, 95% CI [0.20, 0.42] for the 18-word list. Thus, most of the variance on VLMT can be attributed to within-person differences in performance across trials. Using generalizability theory, we further found that the reliability of stable between-person individual differences in VLMT scores across list lengths and trials was quite high ($R_{KF} = 0.94$). In contrast, the reliability of within-person change across list lengths and trials was somewhat low ($R_C = 0.57$).

To examine convergent validity, we examined relationships between the VLMT with immediate recall scores from the HVLT and BVMT (see **Table 4**). We examined the VLMT data in two ways: percentage correct by trial length and overall correct across all trial lengths. In the overall sample, percent of items correct on the 18-item VLMT list was positively correlated with the HVLT ($r = 0.33$, $p < 0.001$). The relationships between the 6- and 12-item percent correct VLMT lists were not significantly related to HVLT performance. When looking at the overall correct data across all three list lengths, the VLMT was positively associated with HVLT ($r = 0.26$, $p = 0.012$). When comparing the VLMT to the BVMT, percent of items correct on the 6-item VLMT list was positively correlated with the BVMT ($r = 0.27$, $p = 0.01$); 12- and 18-item VLMT lists were unrelated to the BVMT. The VLMT overall correct scores (across all three list lengths) was positively correlated with BVMT performance ($r = 0.27$, $p = 0.01$).

We next examined whether there were ceiling effects at any of the VLMT list lengths. At length 6, there was some evidence for ceiling effects such that on Trial 1, 13 (28%) NC and 15 (31%) MCI participants consistently scored 100%; on Trial 2, 23 (50%) NC and 26 (54%) MCI consistently scored 100%; and on Trial 3, 29 (63%) NC and 27 (56%) MCI participants consistently scored 100%. No ceiling effects were observed for list length 12 or 18.

Practice effects were subsequently investigated with linear mixed effect models to determine whether participants' performance on the VLMT systematically changed across the EMA period.[3] Note that all effects were adjusted for list length. On average, participants recognized 10.06 out of an average of 12 words (i.e., average of 6, 12, and 18) correctly (SE = 0.08), averaging across the list lengths. Moreover, participants showed a systematic decline in the number of words they got correct for the list-learning task across the EMA period (on average, participants' total change = −0.84, SE = 0.14, $p < 0.001$). Although MCI participants ($M = 9.87$) significantly differed from controls ($M = 10.25$) on their average number of words correct across the trials ($b = -0.19$, SE = 0.08, $p = 0.015$), participants' systematic change in words correct across the EMA period was not significantly related to MCI status ($b = -0.10$, SE = 0.14, $p = 0.471$).

[3]Practice effects were treated as fixed effects as opposed to random given limited variability in the data set.

TABLE 4 Correlations between mobile cognitive tests and in-lab neuropsychological performance in whole sample ($N = 94$).

Mobile Cognitive Tests (Raw Scores)	Demographic Characteristics				Lab Administered Neuropsychological Tests				
	Age	Sex	Race	Education	WRAT-4	HVLT-Immediate Recall	BVMT-Immediate Recall	Letter Number Span	D-KEFS Color-Word Interference Test (time)
VLMT 6 words (% Correct)	−0.27*	0.25*	0.11	0.04	0.07	0.12	0.27**	0.23*	−0.29*
VLMT 12 words (% Correct)	−0.17	0.09	0.11	0.04	−0.03	0.13	0.09	0.07	−0.17
VLMT 18 words (% Correct)	−0.12	0.24*	0.02	0.04	−0.04	0.33**	0.17	0.03	−0.020
VLMT Overall Mean (all trials)	−0.01	0.37**	0.01	0.08	0.17	0.26**	0.27**	0.10	−0.29*
Memory Matrix (Total Score)	−0.43**	0.09	0.11	0.21*	0.04	0.20	0.17	0.38**	−0.26*
Color Trick: Meaning-to-Meaning (Total Score)	−0.12	0.24*	0.13	0.28**	0.30**	0.28**	0.32**	0.24*	−0.33**
Color Trick: Meaning-to-Color (Total Score)	−0.05	0.18	0.03	−0.25*	0.22*	0.21*	0.29**	0.18	−0.19
Color Trick: Yes-No Mechanic (Total Score)	−0.04	0.23*	0.07	0.33**	0.28**	0.21*	0.19	0.20	−0.18

Note. *$p < 0.05$; **$p < 0.01$.

3.5.2 Psychometric evidence for the Memory Matrix task

The ICC for the average Memory Matrix score across trials was 0.07, 95% CI [0.03, 0.11], indicating that the majority of the variance on this measure can be attributed to within-person differences in performance across trials. Generalizability theory analyses further showed that the reliability of stable between-person individual differences was 0.97. The reliability of within-person change was also satisfactory, with a value of 0.72.

To assess convergent validity, we looked at associations between the Letter-Number Span and performance on Memory Matrix. Memory Matrix scores were positively and significantly correlated with Letter-Number Span ($r = 0.38$, $p < 0.001$). Relationships with demographics and the other lab-administered tests are presented in Table 3.

Although we did not find any evidence of a ceiling effect for Memory Matrix, we nonetheless decided to modify our analyses for the practice effects to account for the possibility of right-hand censoring in the data. Because participants' average scores on these EMCTs tended to be close to the maximum number correct, we wanted to ensure that the growth-curve analyses could accurately capture systematic changes in performance across the EMA period in spite of any measurement limitations. As such, these analyses use Mplus v. 8.4 to estimate what the scores would be if there was not an upper limit (e.g., scores can be greater than 9).

Averaging across trials, participants were estimated to get 8.43 items correct on average out of 10 (SE = 0.12, $p < 0.001$). Moreover, participants showed systematic change in the number of Memory Matrix items they got correct across the EMA period (on average, participants' total change = 1.75, SE = 0.20, $p < 0.001$). However, MCI participants did not significantly differ from NCs on either the intercepts ($b = -0.16$, SE = 0.13, $p = 0.22$) or the slopes ($b = -0.11$, SE = 0.20, $p = 0.577$). In addition, although the data suggest evidence of a practice effect, closer inspection of participants' trajectories *via* spaghetti plots suggests that participants' performance on the Memory Matrix ebbs and flows throughout the EMA period. Specifically, there appears to be a slight decrease in performance from days 1 to 13, then a marked improvement in performance from days 13 to 21, and then a slight decrease in performance from days 21 to 30.

3.5.3 Psychometric evidence for the Color Trick task

We computed ICCs for participants' accuracy on each version of the Color Trick task: Meaning-to-Meaning, ICC = 0.13, 95% CI [0.07, 0.18]); Meaning-to-Color, ICC = 0.17, 95% CI [0.10, 0.23]; and Yes-No Mechanic, ICC = 0.23, 95% CI [0.15, 0.30]), indicating that the majority of the variance on these measures can be attributed to within-person differences in performance across trials. Table 3 presents associations between the Color Trick tasks with demographics and lab-

based assessments. As can be seen, the D-KEFS Interference Trial showed a moderate negative correlation with the Meaning-to-Meaning Color Trick task, such that faster performance on the D-KEFS was related to better performance on Meaning-to-Meaning.

We next examined whether there were ceiling effects for participants' accuracy on any of the Color Trick tasks. There was some evidence for ceiling effects, such that 5 (11%) NC and 1 (2%) MCI participants consistently scored 100% for the Meaning-to-Meaning task; 8 (17%) NC and 4 (8%) MCI participants consistently scored 100% for the Meaning-to-Color task; and 5 (11%) NC and 5 (10%) MCI participants consistently scored 100% for the Yes-No Mechanic task. To account for the possibility of right-hand censoring in the data, we adapted our practice effect analyses for the color trick tasks to be consistent with the modifications we made for the memory matrix task analyses.

For Meaning-to-Meaning trials, participants were estimated to get 9.86 items correct on average (SE = 0.16, $p < 0.001$). Moreover, participants showed systematic change in the number of items they got correct across the EMA period (on average, participants' total change = 2.19, SE = 0.32, $p < 0.001$). Although MCI participants ($M = 9.51$) significantly differed from NCs ($M = 10.21$) on their average number of items correct across the EMA period ($b = -0.35$, SE = 0.14, $p = 0.011$), participants' systematic change in the number of items that they got correct across the EMA period was not significantly related to MCI status ($b = -0.12$, SE = 0.30, $p = 0.677$).

For Meaning-to-Color trials, participants were estimated to get 11.01 items correct on average (SE = 0.23, $p < 0.001$). Moreover, participants showed systematic change in the number of items they got correct across the EMA period (on average, participants' total change = 1.75, SE = 0.42, $p < 0.001$). Similar to performance on Meaning-to-Meaning trials, MCI participants ($M = 10.62$) significantly differed from NCs ($M = 11.40$) on their average number of items correct across the EMA period ($b = -0.39$, SE = 0.16, $p = 0.015$). In addition, participants' systematic change in the number of items correct across the EMA period was not significantly related to MCI status ($b = 0.53$, SE = 0.37, $p = 0.154$).

Lastly, for Yes-No Mechanic trials, participants were estimated to get 11.05 items correct on average (SE = 0.21, $p < 0.001$). Participants also showed systematic change in the number of items they got correct across the EMA period (on average, participants' total change = 0.91, SE = 0.43, $p = 0.035$). MCI participants did not significantly differ from NC on either the intercepts ($b = -0.29$, SE = 0.15, $p = 0.06$) or the slopes ($b = -0.04$, SE = 0.36, $p = 0.902$).

4 Discussion

This study evaluated the feasibility and validity of three mobile cognitive tests among persons with and without MCI. Adherence to this 30-day, fully remote, ecological momentary cognitive testing protocol was very good, with 86% of assigned EMA sessions completed and 84–85% of mobile cognitive testing sessions completed. In this sample of cognitively normal and cognitively impaired older adults, adherence did not differ by MCI status. Further, these findings indicate adherence does not differ by demographic characteristics. Participants who had higher adherence answered more surveys when home and alone compared to people with lower adherence.

We found mixed findings of a fatigue effect at the level of the individual tests, such that there was no evidence of a fatigue effect for the VLMT or Color Trick tests, but participants were more likely to miss Memory Matrix tests over the course of the 30-day protocol (with no difference by NC vs MCI). In another study using the VLMT and Memory Matrix test (14-day protocol in participants with bipolar disorder and control participants) we found an overall fatigue effect for the EMCT protocol, such that participants were more likely to miss a test as study day increased (no differences by diagnostic status), but we did not examine fatigue effects at the level of the individual test (30). Of note, the prior study had a more intensive protocol than the current study, with participants pinged to complete 2–3 mobile cognitive tests three times daily for 14-days. When designing EMCT protocols there is always a frequency and duration trade-off when considering participant burden and capturing outcomes of interest. Our prior work has shown that a 14-day period is sufficient to capture cognition and mood data across various contexts (e.g., 31–35), and other groups have demonstrated strong feasibility and psychometric properties for measuring cognition in as few as 7–8 days (e.g., 14, 16). In general, the 30-day EMCT protocol in this study was largely well tolerated and provides further support for the feasibility of remote, smartphone-based cognitive testing among older adults. Participants had higher rates of adherence than has been reported with other digital health apps (36), which is likely due to a combination of factors including incentives for completing each testing session, brief, gamified tests that varied in difficulty, establishment of good rapport with the study team, and a time-limited engagement with the app.

The psychometric properties of the tasks in this sample were generally good. The reliability of stable between-person individual differences for the VLMT and Memory Matrix were very high, indicating that participants' averaged scores on each mobile cognitive test across the EMA period can reliably assess differences between participants' average levels of the variables. In addition, although the reliability of within-person change (i.e., the consistency in the degree of systematic within-person change across multiple items over time) for Memory Matrix was adequate, the corresponding reliability estimate for the VLMT was not.

Of note, the reliability of within-person change would likely increase if there were more trials, but this would also increase participant burden. As hypothesized, the VLMT overall percentage correct score had an overall moderate positive correlation with the HVLT and BVMT, demonstrating convergent validity. Further, MCI participants recognized significantly fewer words on this task than CN participants. The trajectories of word recognition did not differ by group status across the 30-day study period, but rather, on average, the participants with MCI remembered fewer words overall. In the whole sample, females performed significantly better than males on both the VLMT and HVLT, which is consistent with the female verbal memory advantage highlighted in the Alzheimer's disease literature (e.g., 37), and further supports utility of the VLMT in people with MCI.

Also consistent with our hypotheses, Memory Matrix had a moderately positive correlation with Letter Number Span. Group differences in Memory Matrix performance were not found, although the data did demonstrate variability in performance on this task over the 30-day study period, and future work is needed to examine whether context (e.g., home vs. away from home; alone vs. with others; time of day effects) affected performance on this task. Lastly, data from the Meaning-to-Meaning condition of the Color Trick task was related to faster performance on the D-KEFS Interference Trial. The other two Color Trick conditions were not significantly related to D-KEFS performance. For the Meaning-to-Meaning and Meaning-to-Color trials, MCI participants performed significantly worse than NCs. There was some evidence for ceiling effects, especially among the NC participants, for all versions for Color Trick, and future development of this task, such as increasing the number of trials at each administration or increasing difficulty of the task, may be beneficial if this task is to be adopted in a cognitively normal sample. It is worth noting that traditional neuropsychological tests, albeit used as the "gold standard" comparison for mobile cognitive tests in this study, are limited in that they only provide a snapshot of cognitive abilities at one time point. We would not expect a high correlation between once-administered tests and averaged mobile cognitive testing performance. Additional research is needed to examine whether one testing method is superior to the other when examining clinical outcomes such as disease progression, medication effects, reversion rates, and associations with pathology.

This study is not without limitations. Our sample was largely White and highly educated, which may limit generalizability. There were significantly more women in the cognitively normal group compared to the MCI group, which could have an effect on our findings, especially given the female advantage to verbal memory. Future work is needed with larger and more representative samples to determine whether these tests would be appropriate to detect differences based on cognitive status in randomized controlled trials. Additionally, data were collected during the COVID-19 pandemic, and we did not measure how pandemic-related factors may have influenced performance on these tasks. Another limitation that applies to all ambulatory mobile cognitive testing is that it is difficult to identify suspected cheating, such as whether the participant or someone else took the tests. Relatedly, it is difficult to assess effort on mobile cognitive tests. However, aggregating mobile cognitive test scores can reduce error associated with instances of low effort, as evidenced by the construct validity findings of our mobile cognitive tests with lab-based tests. We did observe evidence of ceiling effects on the VLMT 6-item list and the Color-Trick task in the whole sample, and these trials could possibly be adapted to be made more difficult or used as performance-validity tests in future EMCT protocols. A final limitation is that while we were able to examine differences by smartphone make (iOS vs. Android), we did not have a sufficient sample size to examine differences by smartphone model or OS version, service providers, connectivity, and screen size, all of which may impact response times. Touch sensitivity and latency can differ by up to 100 ms between difference devices, especially between newer and older devices (38, 39). In this study none of the mobile cognitive test outcomes were based on speed. In future work examining timing of responses, these smartphone differences should be examined.

In conclusion, our data add to the extant literature on self-administered mobile cognitive testing in older adults, and is one of the first studies examining an EMCT protocol in people with MCI. The tests are automatically scored, integrated with EMA surveys, and available on iOS and Android operating systems for ease of use by other investigators. Adherence to the EMCTs was high, and the psychometric data are promising. Thus, the three mobile cognitive tests in this study, and particularly the VLMT, may serve as useful tools in future clinical trials with cognition as an endpoint, especially in persons with increased risk for Alzheimer's disease such as those with MCI.

Data availability statement

The original contributions presented in the study are included in the article/**Supplementary Files**, further inquiries can be directed to the corresponding author/s.

Ethics statement

The studies involving human participants were reviewed and approved by UCSD IRB, UTD IRB, UM IRB. The

patients/participants provided their written informed consent to participate in this study.

Author contributions

AEP, RCM, RAA, CAD, and PDH contributed to the conception, design, and obtaining funding for this study. RAA and AEP performed the statistical analysis. RCM wrote the manuscript. RAA, MTR, and LMC wrote sections of the manuscript. AEP, CAD, and PDH provided critical edits. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by grant R01 MH112620-03S1 from the National Institute of Mental Health. LMC was supported by grant F31 AG067869 from the National Institute on Aging.

Conflict of interest

R.C.M. is a co-founder of KeyWise AI, Inc. and a consultant for NeuroUX. P.D.H. has received consulting fees or travel reimbursements from Alkermes, Bio Excel, Boehringer Ingelheim, Karuna Pharma, Merck Pharma, Minerva Pharma, and Sunovion (DSP) Pharma in the past year. He receives royalties from the Brief Assessment of Cognition in Schizophrenia (Owned by WCG Verasci, Inc. and contained in the MCCB). He is Chief Scientific Officer of i-Function, Inc. and Scientific Consultant to EMA Wellness, Inc. No other authors report conflicts.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2022.946685/full#supplementary-material.

References

1. Lancaster C, Koychev I, Blane J, Chinner A, Wolters L, Hinds C. Evaluating the feasibility of frequent cognitive assessment using the mezurio smartphone app: observational and interview study in adults with elevated dementia risk. *JMIR Mhealth Uhealth*. (2020) 8:e16142. doi: 10.2196/16142

2. Moore RC, Swendsen J, Depp CA. Applications for self-administered mobile cognitive assessments in clinical research: a systematic review. *Int J Methods Psychiatr Res*. (2017) 26:e1562. doi: 10.1002/mpr.1562

3. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and validity of ambulatory cognitive assessments. *Assessment*. (2016) 25(1):14–30. doi: 10.1177/1073191116643164.

4. Weizenbaum E, Torous J, Fulford D. Cognition in context: understanding the everyday predictors of cognitive performance in a new era of measurement. *JMIR Mhealth Uhealth*. (2020) 8:e14328. doi: 10.2196/14328

5. Alzheimer's Association. 2017 Alzheimer's disease facts and figures. *Alzheimer's Dementia*. (2017) 13:325–73. doi: 10.1016/j.jalz.2017.02.001

6. Aging NIO. NIA-Funded Active Alzheimer's and Related Dementias Clinical Trials and Studies. U.S. Department of Health & Human Services (2022). Available at: https://www.nia.nih.gov/research/ongoing-AD-trials#:~:text='The%20National%20Institute%20on%20Aging,dementias%20(AD%2FADRD) (Accessed May 4, 2022).

7. Harvey PD, Cosentino S, Curiel R, Goldberg TE, Kaye J, Loewenstein D, et al. Performance-based and observational assessments in clinical trials across the Alzheimer's disease Spectrum. *Innov Clin Neurosci*. (2017) 14:30–9.

8. Kueper JK, Speechley M, Montero-Odasso M. The Alzheimer's disease assessment scale–cognitive subscale (ADAS-cog): modifications and responsiveness in pre-dementia populations. a narrative review. *J Alzheimer's Dis*. (2018) 63:423–44. doi: 10.3233/JAD-170991

9. Posner H, Curiel R, Edgar C, Hendrix S, Liu E, Loewenstein DA, et al. Outcomes assessment in clinical trials of Alzheimer's disease and its precursors: readying for short-term and long-term clinical trial needs. *Innov Clin Neurosci*. (2017) 14:22–9.

10. Weizenbaum EL, Fulford D, Torous J, Pinsky E, Kolachalama VB, Cronin-Golomb A. Smartphone-based neuropsychological assessment in Parkinson's disease: feasibility, validity, and contextually driven variability in cognition. *J Int Neuropsychol Soc*. (2021) 28(4):401–13. doi: 10.1017/S1355617721000503

11. Zlatar ZZ, Campbell LM, Tang B, Gabin S, Heaton A, Higgins M, et al. Daily level association of physical activity and performance on ecological momentary cognitive tests in free-living environments: a Mobile health observational study. *JMIR Mhealth Uhealth*. (2022) 10:e33747. doi: 10.2196/33747

12. Center PR. Mobile Fact Sheet: Mobile phone ownership over time (2021). Available at: https://www.pewresearch.org/internet/fact-sheet/mobile/ (Accessed May 4. 2022).

13. Schweitzer P, Husky M, Allard M, Amieva H, Peres K, Foubert-Samier A, et al. Feasibility and validity of mobile cognitive testing in the investigation of age-related cognitive decline. *Int J Methods Psychiatr Res*. (2016) 26(3):e1521. doi: 10.1002/mpr.1521

14. Thompson LI, Harrington KD, Roque N, Strenger J, Correia S, Jones RN, et al. A highly feasible, reliable, and fully remote protocol for mobile app-based cognitive assessment in cognitively healthy older adults. *Alzheimers Dement*. (2022) 14:e12283. doi: 10.1002/dad2.12283

15. Cerino ES, Katz MJ, Wang C, Qin J, Gao Q, Hyun J, et al. Variability in cognitive performance on Mobile devices is sensitive to mild cognitive impairment: results from the einstein aging study. *Front Digit Health*. (2021) 3:758031. doi: 10.3389/fdgth.2021.758031

16. Wilks H, Aschenbrenner AJ, Gordon BA, Balota DA, Fagan AM, Musiek E, et al. Sharper in the morning: cognitive time of day effects revealed with high-frequency smartphone testing. *J Clin Exp Neuropsychol*. (2021) 43:825–37. doi: 10.1080/13803395.2021.2009447

17. Charalambous AP, Pye A, Yeung WK, Leroi I, Neil M, Thodi C, et al. Tools for app- and web-based self-testing of cognitive impairment: systematic search and evaluation. *J Med Internet Res*. (2020) 22:e14551. doi: 10.2196/14551

18. Jak A, Bondi M, Delano-Wood L, Wierenga C, Corey-Bloom J, Salmon D, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. *Am J Geriatr Psychiatry*. (2009) 17:368–75. doi: 10.1097/JGP.0b013e31819431d5

19. Neuroux (2022). Available at: https://www.getneuroux.com/ (Accessed May 4, 2022).

20. Wilkinson GS, Robertson GJ. *Wide Range Achievement Test Fourth Edition (WRAT-4) Professional Manual* Lutz, FL: Psychological Assessment Resources (2004).

21. Nasreddine Z, Phillips N, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal cognitive assessment. MoCA: A brief screening tool for mild cognitive impairment. *Am J Geriatr Psychiatry*. (2005) 53:695–9. doi: 10.1111/j.1532-5415.2005.53221.x

22. Benedict RH, Schretlen D, Groninger L, Brandt J. Hopkins verbal learning test–revised: normative data and analysis of inter-form and test-retest reliability. *Clin Neuropsychol*. (1998) 12:43–55. doi: 10.1076/clin.12.1.43.1726

23. Benedict RHB, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the brief visuospatial memory test: studies of normal performance, reliability, and validity. *Psychol Assess*. (1996) 8:145–53. doi: 10.1037/1040-3590.8.2.145

24. Ricker JH, Axelrod BN. Analysis of an oral paradigm for the trail making test. *Assessment*. (1994) 1:47–51. doi: 10.1177/1073191194001001007

25. Besser L, Kukull W, Knopman DS, Chui H, Galasko D, Weintraub S, et al. Version 3 of the national Alzheimer's coordinating center's uniform data set. *Alzheimer Dis Assoc Disord*. (2018) 32:351–8. doi: 10.1097/WAD.0000000000000279

26. Gollan T, Weissberger G, Runnqvist E, Montoya R, Cera C. Self-ratings of spoken language dominance: a multilingual naming test (MINT) and preliminary norms for young and aging spanish–English bilinguals. *Biling: Lang Cogn*. (2012) 15:594–615. doi: 10.1017/S1366728911000332

27. Delis DC, Kaplan E, Kramer JH. *Delis-Kaplan Executive Function System (D–KEFS)*. San Antonio: APA PsycTests (2001).

28. Muthén LK, Muthén BO. *Mplus User's Guide: eighth Edition*. Los Angeles, CA: Muthén & Muthén (1998–2017).

29. Cranford JA, Shrout PE, Iida M, Rafaeli E, Yip T, Bolger N. A procedure for evaluating sensitivity to within-person change: can mood measures in diary studies detect change reliably? *Pers Soc Psychol Bull*. (2006) 32:917–29. doi: 10.1177/0146167206287721

30. Moore RC, Parrish EM, Van Patten R, Paolillo EW, Filip T, Bomyea J, et al. *Initial Psychometric Properties of 7 NeuroUX Remote Ecological Momentary Cognitive Tests Among People with Bipolar Disorder*. JMIR mHealth Uhealth (2022). In Press.

31. Bomyea JA, Parrish EM, Paolillo EW, Filip TF, Eyler LT, Depp CA, et al. Relationships between daily mood states and real-time cognitive performance in individuals with bipolar disorder and healthy comparators: a remote ambulatory assessment study. *J Clin Exp Neuropsychol*. (2021) 43:813–24. doi: 10.1080/13803395.2021.1975656

32. Campbell LM, Paolillo EW, Heaton A, Tang B, Depp CA, Granholm E, et al. Daily activities related to Mobile cognitive performance in middle-aged and older adults: an ecological momentary cognitive assessment study. *JMIR Mhealth Uhealth*. (2020) 8:e19579. doi: 10.2196/19579

33. Jones SE, Moore RC, Pinkham AE, Depp CA, Granholm E, Harvey PD. A cross-diagnostic study of adherence to ecological momentary assessment: comparisons across study length and daily survey frequency find that early adherence is a potent predictor of study-long adherence. *Pers Med Psychiatry*. (2021) 29–30:100085. doi: 10.1016/j.pmip.2021.100085

34. Moore RC, Campbell LM, Delgadillo JD, Paolillo EW, Sundermann EE, Holden J, et al. Smartphone-Based measurement of executive function in older adults with and without HIV. *Arch Clin Neuropsychol*. (2020) 35:347–57. doi: 10.1093/arclin/acz084

35. Moore RC, Paolillo EW, Sundermann EE, Campbell LM, Delgadillo J, Heaton A, et al. Validation of the mobile verbal learning test: illustration of its use for age and disease-related cognitive deficits. *Int J Methods Psychiatr Res*. (2021) 30:e1859. doi: 10.1002/mpr.1859

36. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. *NPJ Digit Med*. (2020) 3:21. doi: 10.1038/s41746-020-0224-8

37. Caldwell JZK, Berg JL, Cummings JL, Banks SJ. Moderating effects of sex on the impact of diagnosis and amyloid positivity on verbal memory and hippocampal volume. *Alzheimers Res Ther*. (2017) 9:72. doi: 10.1186/s13195-017-0300-8

38. Siegal J. Here's why typing on Android phones is harder than typing on an iPhone. BGR (2013a). Available at: http://bgr.com/2013/09/20/iphone-android-touchscreen-responsiveness/ (Accessed).

39. Siegal J. Study: iPads are the most responsive tablets in the world (2013b). Available at: http://bgr.com/2013/10/09/tablet-touch-screen-responsiveness/ (Accessed).

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 23 August 2022
DOI 10.3389/fdgth.2022.920706

OPEN ACCESS

EDITED BY
Timo Jämsä,
University of Oulu, Finland

REVIEWED BY
Jack Holman,
University of Nottingham, United Kingdom
Alejandro Domínguez-Rodríguez,
Valencian International University, Spain

*CORRESPONDENCE
Federica Klaus
fklaus@health.ucsd.edu

SPECIALTY SECTION
This article was submitted to Human Factors and Digital Health, a section of the journal Frontiers in Digital Health

RECEIVED 14 April 2022
ACCEPTED 08 August 2022
PUBLISHED 23 August 2022

CITATION
Klaus F, Peek E, Quynh A, Sutherland AN, Selvam D, Moore RC, Depp CA and Eyler LT (2022) Mobile survey engagement by older adults is high during multiple phases of the COVID-19 pandemic and is predicted by baseline and structural factors.
Front. Digit. Health 4:920706.
doi: 10.3389/fdgth.2022.920706

Mobile survey engagement by older adults is high during multiple phases of the COVID-19 pandemic and is predicted by baseline and structural factors

Federica Klaus[1]*, Elizabeth Peek[1], Avery Quynh[1], Ashley N. Sutherland[1,2], Divya Selvam[1], Raeanne C. Moore[1], Colin A. Depp[1,2] and Lisa T. Eyler[1,2]

[1]Department of Psychiatry, UC San Diego, La Jolla, CA, United States, [2]VA San Diego Healthcare System, Mental Illness Research, Education, and Clinical Center (MIRECC), La Jolla, CA, United States

Digital surveys, such as mobile phone ecological momentary assessment (EMA), bear the potential to assess and target individual wellbeing in a personalized, real-time approach and allow for interaction in situations when in-person contact is not possible, such as during the coronavirus pandemic. While the use of digital technology might especially benefit research in older adults who find themselves in circumstances of reduced mobility, little is known about their barriers to adherence. We investigated baseline and structural factors that predict study withdrawal and adherence from daily smartphone EMA self-report surveys in the StayWELL Study. The StayWELL study is a longitudinal, observational study on the relationship between social restrictions during the coronavirus pandemic and mental well-being in 95 community-dwelling older aged adults (67–87 years) who were participants in a randomized clinical trial using EMA. Withdrawal was associated with less research staff changes and less likely in participants that reached the study mid-point. No baseline characteristics predicted withdrawal. Main reasons for withdrawal were communication issues, i.e. staff not being able to contact participants. We found an adherence rate of 82% and no fatigue effects. Adherence was predicted by education status, study participation duration, reaching the study midpoint and time between study start and enrollment. COVID infections or supporting people in the household was not related to adherence. To conclude, it is feasible to conduct an EMA study in older people without impacting engagement during a pandemic. Furthermore, personal characteristics and smartphone operating system (Android vs. iOS) used did not relate to engagement, allowing for a broad distribution of digital health technologies. Our study adds information on single predictive variables relevant for adherence and withdrawal from EMA smartphone surveys in older people that can inform the design of future digital EMA research to maximize engagement and reliability of study results.

KEYWORDS

ecological momentary assessment, survey, adherence, withdrawal, stay-at-home, mobile phone, older adults, pandemic

Introduction

Ecological Momentary Assessment (EMA) repeatedly samples participants' current mood, behaviors, and sense of wellbeing in real time in their natural environments (1, 2). Surveys delivered multiple times daily *via* mobile phone apps with text messaging reminders can capture self-reports in real-time and minimize recall bias while preserving ecological validity (1, 2). However, to best capture momentary feelings and behaviors, it is necessary for participants to answer as many surveys as possible and to stay in the study throughout its duration. Missing data due to low EMA adherence rate and, as an extreme form, study withdrawal, are important since they lead to low statistical power, response bias, and increased cost to the researchers (3). It is important to understand typical rates of withdrawal and adherence in order to power future studies to discover factors that predict adherence and withdrawal in order to devise protocols that maximize participation.

In the general population, one predictor of high overall adherence rates is high adherence in the early phase of the protocol (3). Contextual factors, such as time of day, may play a role in missing survey reminders whereas increased training on how to use EMA is correlated with higher adherence levels (4). Some research suggests that older age, healthy mood and affect were correlated with higher levels of adherence (4–9).

Some early research has been done using EMA with older adults; but the scope of the literature is narrow. This can partially be attributed to the fact that in 2015, only 27% of American older adults owned a smartphone; however smartphone ownership among older adults is growing rapidly with 83% of those 50–64 years old and 61% of those 65 years and older owning a smartphone in the US in 2021 (10–13). While older adults are more reluctant to use new technology, they are more likely to utilize new technologies when they understand the benefits (14, 15). A review on EMA in aging research reported a general adherence rate of over 80% in most studies assessed (16), which is higher compared to younger adults, where a survey adherence rate of 75% with a withdrawal rate of 15% was reported (5, 17). One study used EMA over a 14-day burst to look at daily activities and neurocognitive health in 103 older American adults and observed an adherence rate of 91% (18). Another study looked at EMA engagement for African American older adults, a group that was expected to be more wary of EMA and see it as surveillance; however they also had high adherence rates of on average 92%–98% and a withdrawal rate of 9% over the whole study (19). Graham et al. examined American users of a digital healthcare platform with multiple interfaces, and found that on an aggregate level, older users (65 years and older) utilized the app more than younger users (35–64 years) (20). In a sample of older participants (aged 50–70 years), 95 older adults were sent six surveys per day for a week and had a 91.5% response rate (21). In a study measuring older adults' (60–98 years) physical activity using EMA, an adherence of 92% and withdrawal rate of 2% was observed (8).

To understand the reasons for withdrawal from an EMA study and what influences withdrawal from and adherence to EMA surveys, more detailed information on personal baseline factors and study structural factors are needed. In the previously cited studies, systematic information on reasons for study withdrawal in older adults are rarely reported, with one study mentioning withdrawal due to medical emergency or participant burden in older adults (8) and another describing as a primary withdrawal reason that participants did not fully understand what they were supposed to do (22). Moderators of adherence are reported for some, but not all studies and focus mainly on personal factors, such as age, sex, relationship status, residence, number of people in residence, education, employment status, income, and location. Most studies do not find an association between personal baseline characteristics and study adherence (8, 19). One study in older adults (50–70 years) observed higher adherence in female participants aged 50–59 years (93.3%) vs. male participants of the same age (84.5%), but overall adherence levels were above the recommended 80% (21). Another study in older people reported no relation of age to completion or response rate, but found that older participants were more likely to report not being alerted to surveys and that issues with survey alerts were independently related to Android operating system (23). The same study also observed higher response rates among iOS vs. Android users (23). However, information on other structural factors specific to older adults' adherence to EMA studies is sparse. One study reported a higher likelihood of missing an EMA survey in the afternoon compared to the morning in healthy older adults, providing information on time-varying factors (8).

The above-described studies provide evidence that withdrawal from EMA studies among older adults is generally low and EMA adherence is generally high in typical settings. However, it is unclear whether the same would be seen during a global pandemic such as that caused by COVID-19. Due to social distancing measures during the pandemic, studies suggest that older adults increased their technology consumption. One study showed that older adults used technology to connect socially and two-thirds of participants learned a new communication technology (24). Another study showed that 73% of German nursing homes self-reported increased opportunities for residents to connect virtually (25). Another study demonstrated that there was an increase in older adults who ordered groceries using mobile delivery (26). This shows that there was increased use of digital technology during the pandemic among older people. Additionally, there was an increase in using EMA methods during the pandemic with researchers being unable to conduct in-person visits due

to social distancing; however, there is limited literature about adherence levels for older adults in this context. Most recently, a study in 47 older adults (age 45–78 years) during the pandemic sent six surveys per day for a week and had a 84% vs. 54% completion and a 64% vs. 54% response rate among experienced vs. inexperienced EMA users (23).

We created the Stay-at-home Wellness EMA in Late Life (StayWELL) Study to understand the wellbeing of older adults during the pandemic. StayWELL enrolled a well-characterized sample of older adults (>65 years) who had previously completed a randomized controlled trial which included EMA sampling and collection of detailed demographic and psychological assessments. As a completely virtual, longitudinal study, StayWELL collected self-report data on mental wellbeing and daily activities, using online questionnaires and EMA *via* mobile surveys, throughout the pandemic. EMA data was collected during two 2-week bursts of assessments in Summer/Autumn 2020 and in Summer of 2021. The data from this study therefore provides a unique opportunity to assess factors that predict withdrawal and adherence to EMA in older adults over the course of a seventeen-month period, beginning during the sudden and long-term shutdown of normal behavior and routines and a forced shift to digital technology. Given the lack of longitudinal studies on older adults during the coronavirus pandemic, our study provides the opportunity to capture a wide breadth and depth of data on potential predictors of EMA study adherence and study withdrawal factors of older adults in situations of reduced mobility due to pandemic-related social restrictions.

We investigated study withdrawal and adherence to daily EMA surveys in the StayWELL study. The first aim was to explore reasons for withdrawal and if any personal baseline factors predicted withdrawal in older adults, including age, gender, race/ethnicity, employment, residence, number of people in residence, education, location, generations in the household, how often the house was left pre-pandemic and previous participation in an active treatment arm of the prior study randomized controlled trial (vs. a control condition). Further, we examined if study structural factors, including study research staff turnover, smartphone operating system, time elapsed between study start and enrollment and reaching the study mid-point predicted withdrawal.

The second aim was to examine if any personal baseline or study structural factors predicted adherence—per burst—in older adults, based on the same personal and structural factors explored in relation to withdrawal. We hypothesized that adherence will be predicted by age and gender with highest adherence in younger women based on specific findings on older adults (21).

The final aim was to understand how the rate of completed EMA surveys from the first two bursts (burst 1 and 2; administered before significant lifting of restrictions, i.e. stay-at-home and social distancing order, in Summer/Fall 2020) compared to the last two bursts (burst 3 and 4; after lifting of restrictions in Summer of 2021). We hypothesized that adherence would be significantly less during burst 3 and 4 due to participant's reengaging in activities outside the home.

Materials and methods

Participants and recruitment

Ninety-five community-dwelling older aged adults (67–87 years) were included in the StayWELL study. All had previously participated in the Mindfulness, Education, and Exercise (MEDEX) study (27) and were concurrently enrolled in an extension trial. The MEDEX study was a now-concluded 18-month randomized controlled trial to assess the effects of in-person interventions, which included three active treatments, i.e. mindfulness-based meditation, exercise and their combination, and one active control group, i.e. health education, on cognition in older adults. MEDEX also included daily EMA assessments on study-provided tablets of self-reports (such as positive and negative affect) during four 10-day periods for all groups. Most participants who completed the in-person study then continued to take part in the extension of the randomized trial, where participants received once per month a virtual booster session of the same intervention that they had received in the previous trial. The extension of MEDEX started in October 2018 and is ongoing planned until September 2023. Participants were recruited by MEDEX staff during the monthly virtual booster sessions and when they received MEDEX study related information *via* mail on a rolling basis. After receiving advertisement about the StayWELL Study, 124 MEDEX study participants contacted the StayWELL study team and 95 of these participants decided to participate after hearing what study participation entailed. Couples were allowed to participate in the study and enrolled occasionally (approximately 2 couples).

Inclusion criteria were previous participation in the MEDEX study and current enrollment in the MEDEX extension study (St. Louis or San Diego), and the possession of a mobile device (Android or iOS operating system) with touch screen and internet access. All procedures were approved by UCSD's Institutional Review Board before protocol implementation, and all participants provided oral informed consent.

Measures and procedures

Study design

The fully virtual StayWELL study began in June 2020 and was completed in October 2021.

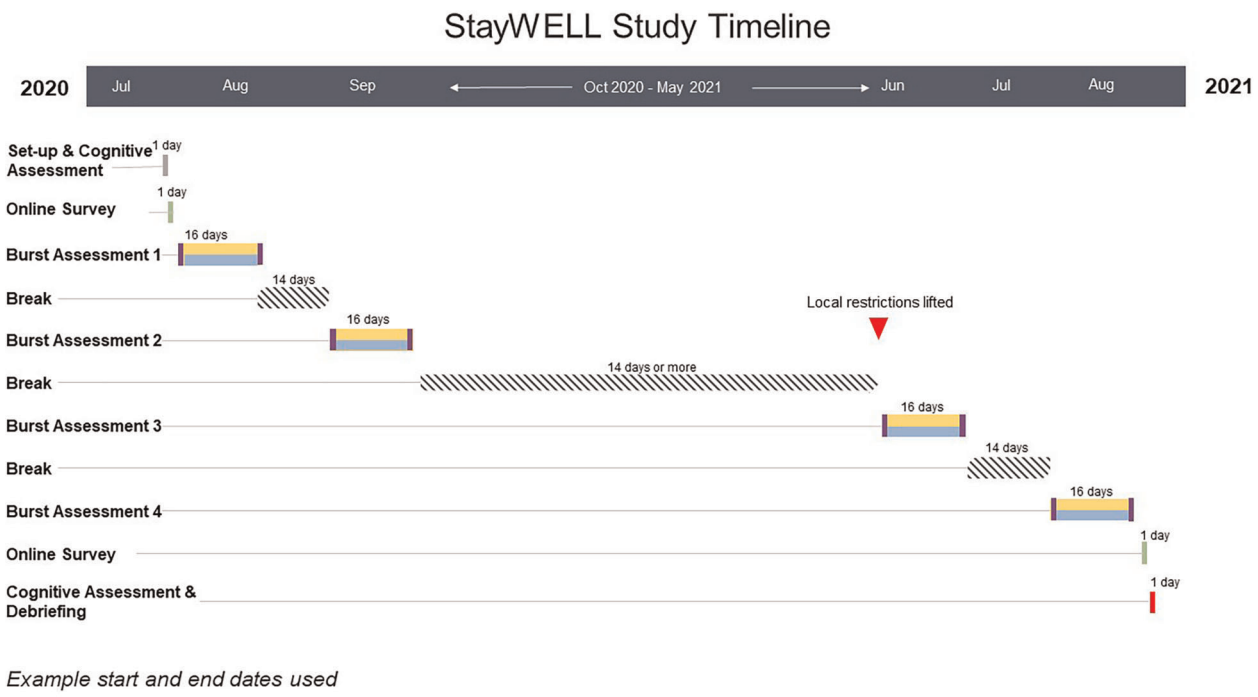


FIGURE 1
Exemplary timeline of StayWELL study design.

One study visit at the beginning and end of the study included a Set-up/Final Call and cognitive assessment, of which the data is not analyzed at this point, and direction to an online questionnaire. In-between, four EMA bursts, which lasted each 16 (14 + 2)-days, with a long break in-between burst 2 and 3 (mid-study) took place (See **Figure 1**). Due to the rolling advertisement with participants contacting the study staff at their will, time between study start and actual study enrollment varied between participants.

Self-report data, including mental wellbeing and daily activities, was collected using online questionnaires and EMA *via* mobile surveys throughout the pandemic. All study visits were conducted *via* videoconferencing using the software Zoom or *via* phone at the beginning and end of the study and baseline information (demographics etc.) was assessed using the online questionnaire at the beginning and end of the study. The number of available staff or volunteer research associates (research assistants (RA's)) did oscillate due to the unexpected long duration of the study with the restrictions during the pandemic remaining in place longer than anticipated. When this longitudinal study was conceived, a duration of the pandemic of a few weeks and therefore a study duration of maximum half a year was generally expected, which was proven wrong by the prolonged nature of the pandemic. These events led to unexpected personal circumstances that were reflected in the availability of RA's. At the beginning of the study, each participant was assigned to one of the available ten RAs. During the study, some participants were re-assigned to other RA's due to the reduced availability or change of position of some (due to personal and educational duties). Short communication by RA's (e.g. for confirmation of appointments or check-in visits) was mostly conducted using Skype phone calling, Google voice calling or Google voice texting or email. Longer visits/trouble-shooting visits were conducted by sending a Zoom link to the participants *via* email and then conducting a video call using Zoom. Participants received a $30 Amazon gift card at the conclusion of the study. If they complete greater than 85% of the daily EMA surveys, they received a bonus of $20 in the form of an additional gift card.

Smartphone set-up visit

After giving initial information to the potential study participants and obtaining consent *via* a phone call, an RA was assigned to each participant. The RA contacted the participant *via* Zoom (video call) and conducted a 1-hour long set-up visit, during which the participants were guided to install all necessary applications on their own mobile devices and participants were provided an individualized live tutorial delivered by an RA that also used recorded video sequences of how to install the apps on their mobile phone (iOS or Android specific) and how to complete EMA surveys. Further, they were given, *via* email, a written and illustrated document explaining how to use the apps on their smartphone. At the end of the setup call, the participants practiced an example EMA survey that came through to their mobile device. Participants were provided contact information in the event

they experienced technological difficulties and were contacted if a drop in completed surveys suggested difficulties. The platform and application mEMA (Illumivu) was used for the smartphone-based ecological momentary assessment and could be downloaded from the respective AppStores by the participants on their mobile devices.

EMA bursts

Four 16-day total mobile burst assessments took place (bursts 1–4). The first two bursts started after enrollment with an approximate gap of 2 weeks in-between (depending on participant's availability) and were completed in Summer/Fall 2020. The third burst was planned to occur upon significant lifting or pandemic-related social restrictions and started in Summer/Fall 2021 upon lifting of restrictions and the fourth burst followed after an approximate 2-week gap.

Each burst started and ended with a longer survey (20 min, scheduled to arrive in the morning) asking about thoughts, behaviors and events related to well-being and the coronavirus pandemic over the past 2 weeks on the first and last day. Each burst then included 14 days of twice-daily brief momentary surveys (5–10 min long, scheduled in morning and evening randomly within a 1-hour time window according to participant's preference). The brief daily surveys asked about wellbeing, mood, compassion, empathy, social isolation, mindfulness, resilience, and loneliness and behavior in the moment. A pop-up notification and a sound alert (if not set to mute) was sent that reminded the participant to take the surveys. The surveys started at random times (15 min, 30 min, 45 min or 60 min apart from the last survey) within the specified 1-hour window and were available to answer until 1 h after the notification first appeared (see **Figure 1**).

All participants were contacted mid-study to enhance retention and explain the protocol for the remaining bursts. The final timing of assessments differed from what was anticipated at study start due to the protracted course of the pandemic. Thus, although four bursts had been planned all along, the total time in the study was much longer than expected by participants when they enrolled. In addition to beginning and end-of study contacts, RA's contacted participants a week in advance before each burst started, at the start date of each burst, during the burst if needed, at the end of each burst and during the longer mid-study break to remind participants that the study was still ongoing.

Data analysis

Time frame and variables used

Number of surveys completed out of the 4×14-days short daily surveys in the morning and evening were analyzed in this study and morning and evening surveys were collapsed into a daily average.

Adherence was calculated as percentage of surveys validly answered per burst and per burst category (burst 1/2 and burst 3/4 were collapsed into burst categories before and after lifting of stay-at-home restrictions).

Analysis of adherence was conducted in all participants that contributed EMA data. Analysis of withdrawal included all participants that initially consented to the study. Only data from participants who completed burst 1 and/or 2 and burst 3 and/or 4 were used for the analysis of adherence differences between burst.

Demographic variables were collected at baseline. The original choices for people in the same household (number of people) were collapsed into categories: (I live alone, 1 person, 2 or more people). Information about COVID-19 diagnosis and support given to family members was derived from the longer survey at the end of each burst covering the past 2 weeks.

Quantitative assessment of withdrawal reasons was performed by classifying the reasons given for withdrawal by the participants into the following categories: participant was hard to reach/communication issues, unknown/reason not given, no longer interested/too busy, technical difficulties, personal/unspecified and health reasons. Qualitative assessment of feedback given by participants is based on the notes taken of the conversations by the RAs who received the note of withdrawal by the participant.

Statistical analysis

Statistical analyses were computed with SPSS version 25 (IBM Corp., SPSS Inc., Chicago IL, USA). Demographic variables were compared using chi-square, Mann-Whitney U or two-tailed t-tests, as appropriate. Generalized linear mixed model analyses were conducted with subject ID as random effects in all analyses. Level of significance was set at $p < 0.05$. Multiple testing was accounted for using the false discovery rate (FDR) (28). Data are presented as mean and standard deviation (SD) if not noted otherwise. Data of all participants who consented to the study were analyzed for withdrawal analyses. Only data from participants where EMA survey were set up were analyzed in analyses investigating adherence.

Specifically, the following methods were applied for each study aim:

1) To assess predictors of and reasons for withdrawal, data of all participants that consented to study participation were entered in the analysis. Separate models with baseline predictors as fixed effects were ran with age, gender, race/ethnicity, employment status, type of residence, education, location (San Diego vs. St. Louis), number of people living in the same residence and assignment to the MEDEX treatment group as well as generations in the household and how often the house was left (before the pandemic).
   To assess the structural variables, number of changes of RA, operating system, time elapsed between study start to

enrollment and reaching burst 3 (i.e. the second half of the study/study mid-point) were used as fixed effects. All significant predictors were then combined into one model to investigate potential dependencies. Spearman correlations due to non-normal distribution of date variables were used to assess relationships of number of RA changes with study participation duration, since participants that were longer in the study might have experienced more RA changes. A multivariate model with significant predictors and study participation duration was used to account for dependency of RA change on study participation duration length.

2) To assess predictors of adherence, data of all participants that had EMA surveys set up were entered in the analysis. Separate models with burst and each baseline predictor as fixed effects were used with age, gender, race/ethnicity, employment status, type of residence, education, location (San Diego vs. St. Louis), number of people living in the same residence and assignment to the MEDEX treatment group as well as generations in the household and how often the house was left (before the pandemic). To assess the structural variables, number of changes of RA, operating system, time elapsed between study start to enrollment and reaching burst 3 (i.e. the second half of the study) were used in separate models as fixed effects in addition to burst. Correction for multiple testing using FDR was applied (denoted as adjusted p (adj.p)). Finally, all individual variables that remained significant after FDR correction were entered into one model to investigate potential dependencies.

   To follow up on a hypothetical moderation of the relationship between people living in residence and adherence by COVID-19 infections in the family and giving support to family members, two separate models were calculated. Fixed effects were people in the residence and support given to family members, or people in the residence and COVID-19 infections of family members and their respective interaction terms.

   To predict adherence per burst based on structural factors, separate mixed model analyses were used with burst, number of changes of RA, operating system, time elapsed between study start to date of individual study enrollment, and study participation duration as fixed effects. Spearman correlations due to non-normal distribution of date variables were used to assess relationships of number of RA changes with time elapsed between study start and enrollment and with study participation duration, since participants that were longer in the study might have experienced more RA changes.

3) To compare the rate of completed EMA surveys from the first two to the last two bursts within those participants that contributed data to burst 1 and/or 2 and to burst 3 and/or 4, mixed model analyses were used with burst and burst category as fixed effects respectively.

Results

Sample characteristics

Ninety-five participants enrolled in the study and 47 participants completed study procedures ((49.5%, 67–87 years) whereas the other 48 participants (50.5%, 67–83 years) withdrew.

The mean (SD) age of all enrolled participants at baseline was 74 (4.3) years, 78% were women, and the mean (SD) education was 16.6 (2.0) years. The racial distribution was 84% White, 8% Black/African American, 4% Asian, and 3% More than One Race; Hispanic participants were 5% of the sample (See **Table 1** for details.)

Among the participants that completed the study, 42 participants contributed data to burst 1 and/or 2 and to burst 3 and/or burst 4 (See **Supplementary Table S1** for details.).

Participants who reached the study mid-point were less likely to withdraw after that time-point

Analysis of withdrawal were conducted in the whole dataset containing 95 participants. Statistical details on estimates for each level of predictor can be found in the **Supplementary information**.

Of all participants that withdrew, 46 participants withdrew before burst 3 and 2 participants withdrew after burst 3. See **Table 2** for exact timing of withdrawal. Main reasons for withdrawal were staff- participant communication issues (31%), not enough time for the study (21%) and no reasons reported (21%). Technical difficulties accounted for 15% of withdrawals, partially contrary to our initial hypothesis (see **Table 3**).

Baseline structural characteristics did not predict withdrawal, specifically withdrawal was not predicted by gender ($F(1,93) = 1.6$, $p = 0.21$), age ($F(1,93) = 0.022$, $p = 0.88$), race ($F(3,91) = 0.99$, $p = 0.40$), ethnicity ($F(1,93) = 1.56$, $p = 0.21$), employment status ($F(3,75) = 0.01$, $p = 0.99$), residence type ($F(2,76) = 0.23$, $p = 0.79$), education status ($F(1,91) = 0.06$, $p = 0.81$), location ($F(1,93) = 1.25$, $p = 0.27$), generations in household ($F(2,76) = 0.05$, $p = 0.95$), how often the house was left before the pandemic ($F(3,75) = 1.18$, $p = 0.32$), number of people living in residence ($F(2,76) = 0.18$, $p = 0.84$) or MEDEX intervention group ($F(3,91) = 1.05$, $p = 0.37$).

Structural factors significantly related to withdrawal were number of changes of RA's ($F(2,92) = 6.9$, $p = 0.002$) and making it beyond the mid-study point, i.e. reaching burst 3 ($F(1,93) = 34.5$, $p < 0.001$), which remained significant after

TABLE 1 Demographics of study participants who withdrew and who completed the study.

	Participants who withdrew ($n = 48$)		Participants who completed the study ($n = 47$)		Test statistics	
	Mean (range) or n (%)	SD	Mean (range) or n (%)	SD	$t/\chi^2/U$	p
Age (years)	73.8 (67-83)	3.9	73.7 (67-87)	72.3	$t = -1.5$	0.9
Gender (F (%))	40 (83%)		34 (72%)		$\chi^2 = 1.7$	0.2
Education (years)	16.6	2.4	16.7	1.6	$U = 1075$	0.9
Race (n (%))					$\chi^2 = 3.4$	0.4
White	39 (81%)		41 (87%)			
Black/African American	6 (13%)		2 (4%)			
Asian	1 (2%)		3 (6%)			
More than one Race	2 (4%)		1 (2%)			
Ethnicity (Non-Latino-Hispanic (%))	44 (91%)		46 (97%)		$\chi^2 = 1.8$	0.18
Employment status (Retired (%))	27 (56%)		42 (89%)		$\chi^2 = 3.1$	0.4
Number of people living in same household (0/1/2 or more)	21%/40%/6%		26%/66%/8%		$\chi^2 = 2.6$	0.6
Number of generations living in same household	0.2	0.4	0.4	0.0	$U = 1210$	0.4
How often house was left per day before pandemic (days per week)	4.3	0.9	4.0	0.9	$U = 1733$	0.1
COVID19 Diagnosis of family member (burst 1/2/3/4)	31%/9%/0%/0%		9%/20%/16%/18%			
Support to family members given (burst 1/2/3/4)	23%/27%/0%/0%		30%/29%/35%/23%			
Location (San Diego (%))	22 (45%)		27 (57%)		$\chi^2 = 1.3$	0.3
Medex intervention group					$\chi^2 = 3.4$	0.3
MBSR (mindfulness-based meditation)	11 (23%)		16 (34%)			
Exercise	12 (25%)		14 (30%)			
MBSR + Exercise	18 (38%)		10 (21%)			
Health Education (comparison group)	7 (15%)		7 (15%)			
Number of research assistant (RA) changes	0.3	0.6	0.7	0.7	$U = 1885$	**$p < 0.001$**
Operating System (Android/iOS (%))	18 (38%)/30 (62%)		16 (34%)/31 (66%)		$\chi^2 = 1.2$	0.7
Bursts completed (0/1/2/3/4)	11/12/16/0/0		1/2/3/10/31			
Last completed burst (0/1/2/3/4)	20/12/15/1/0		1/2/3/3/38			
Withdrawal timing (until during burst 2/between burst 2 and 3/during or after burst 3)	27/19/2		–		–	–
Study participation duration (years)	0.5	0.4	1.1	0.2	$U = 1336$	**$p < 0.001$**
Time elapsed between study start and date of enrollment (months)	1.5	0.9	1.5	0.8	$U = 2226$	0.8

Significant results with p-values < 0.05 are bolded.
χ^2, Chi-square; U, Mann-Whitney U; t, two-tailed t-test.

adjusting for multiple testing ($adj.p = 0.003$ and $adj.p < 0.001$, respectively). Completing the study was associated with having 1 RA change vs. 0 ($p < 0.001$, odds ratio = 0.14) and with reaching burst 3 ($p < 0.001$, odds ratio = 0.006). There was no relation of withdrawal with operating system ($F(1,93) = 0.12$, $p = 0.73$) and time elapsed between study start and enrollment ($F(1,93) = 0.07$, $p = 0.79$). When combining the two significant predictors into one model, the relationship of both RA changes ($F(2,91) = 3.23$, $p = 0.04$) and reaching burst 3 ($F(1,91) = 31.1$, $p < 0.001$) to withdrawal remained significant, with having no vs. 1 or 2 RA changes predicting withdrawal ($p = 0.04$, odds ratio = 0.15 and $p = 0.03$, odds ratio = 0.09 respectively). Because there were significant positive correlations of number of RA changes with study participation duration ($r_s(95) = 0.31$, $p = 0.003$), we also examined a model with additionally study participation duration included. In that model, number of RA changes was no longer significant ($F(2,90) = 1.53$, $p = 0.22$), while reaching burst 3 remained a significant predictor of withdrawal ($F(1,90) = 22.7$, $p = 0.04$), indicating that RA changes are dependent on the other two predictors.

Adherence to EMA surveys was not significantly associated with study withdrawal ($F(1,72) = 3.59$, $p = 0.06$), the direction of the trend-level relationship was such that poor adherence was associated with greater likelihood of withdrawal (odds ratio = 0.98, 95% confidence interval [0.96, 1.00]).

TABLE 2 Timing and reason for withdrawal.

Timing of withdrawal	Reason for withdrawal						
	no longer interested/too busy	technical difficulties	personal/ unspecified	health reasons	hard to reach/ communication issues	unknown/reason not given	Total
before burst 1	0	2	1	0	5	4	**12**
during burst 1	1	1	0	0	1	0	**3**
after burst 1 before burst 2	2	2	2	0	3	1	**10**
during burst 2	1	1	0	0	0	0	**2**
after burst 2 before burst 3	5	1	2	1	5	5	**19**
during burst 3	0	0	0	0	0	0	**0**
after burst 3 before burst 4	1	0	0	0	1	0	**2**
during burst 4	0	0	0	0	0	0	**0**
Total	**10**	**7**	**5**	**1**	**15**	**10**	

Significant results with p-values < 0.05 are bolded.

TABLE 3 Reasons for study withdrawal.

Reason for withdrawal	Count	Percent per total withdrawals/ total participants
hard to reach/communication issues (participant could not be contacted)	15	31% / 15.8%
unknown/reason not given	10	21% / 10.5%
no longer interested/too busy	10	21% / 10.5%
technical difficulties	7	15% / 7.4%
personal/unspecified	5	10% / 5.3%
health reasons	1	2% / 1.1%

Adherence was predicted by education status, study participation duration, time between study start and enrollment and reaching the second half of the study

Analysis of adherence was conducted in all participants that contributed EMA data ($n = 74$). Statistical details on estimates for each level of predictor can be found in the **Supplementary information**.

Among *baseline characteristics*, education status significantly predicted adherence after correction for multiple testing ($F(1,206) = 6.27$, $p = 0.013$, adj.$p = 0.04$)), with more years of education being associated with higher adherence. The significant relation of adherence to number of people living in residence ($F(2,192) = 4.48$, $p = 0.01$ (adj.$p = 0.05$), living with 2 people or more vs. living alone was associated with poorer adherence), participation in the MEDEX active treatment groups ($F(3, 210) = 3.96$, $p = 0.009$ (adj.$p = 0.11$), being in an active control group vs. being in a MBSR or MBSR plus exercise, but not exercise alone group, was associated with poorer adherence) and race ($F(3,210) = 3.86$, $p = 0.01$ (adj.$p = 0.06$)), being White vs. Black or African American, but not Asian or more than one race, was associated with higher adherence) did not remain significant after correction for multiple testing. Adherence was not predicted by gender ($F(1, 212) = 3.02$, $p = 0.08$), age ($F(1, 212) = 0.06$, $p = 0.80$), ethnicity ($F(1,212) = 0.36$, $p = 0.55$), employment status ($F(3,205) = 0.08$, $p = 0.97$), residence type ($F(2,206) = 1.1$, $p = 0.34$), location ($F(1,212) = 0.58$, $p = 0.45$), generations in household ($F(2,206) = 0.03$, $p = 0.97$) or how often the house was left before the pandemic ($F(3, 205) = 1.07$, $p = 0.36$).

To follow up whether the trend-level association of lower adherence with people living with 2 or more people vs. living alone might be moderated by pandemic-related events, we examined whether increased care duties for family members or the presence of COVID-19 infections in the household might moderate this relationship.

However, in multivariate models, we found no relationship of adherence with a COVID-19 diagnosis of a family member ($F(1,170) = 0.02$, $p = 0.89$) and no interaction of the number of people in residence with COVID 19 infections of family members ($F(2, 170) = 0.056$, $p = 0.95$). Further, adherence was not related to whether the participant was giving support to a family member ($F(1,160) = 0.42$, $p = 0.52$) and there was no interaction effect of number of people in residence with giving support to a family member ($F(2, 160) = 0.56$, $p = 0.57$), indication no moderating effects of variables related to COVID19.

Structural factors related to adherence were duration of study participation ($F(1, 212) = 11.7$, $p < 0.001$, adj.$p = 0.004$) with a longer study participation associated with higher adherence, reaching burst 3 (($F(1, 212) = 9.84$, $p = 0.002$ (adj.$p = 0.054$), with higher adherence in participants that made it to burst 3), and time elapsed between study start and enrollment ($F(1,212) = 5.18$, $p = 0.02$ (adj.$p = 0.004$), with a faster enrollment associated with a higher adherence). Structural factors not related to adherence were number of changes of RA's ($F(1, 211) = 2.36$, $p = 0.09$) and

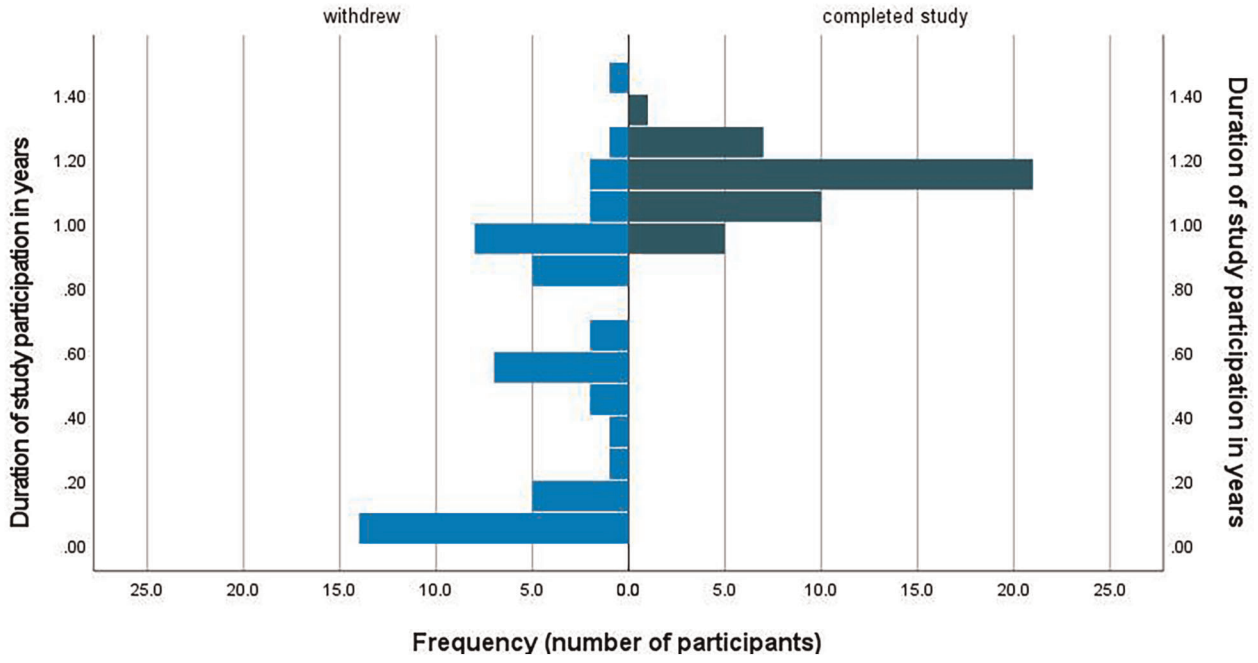


FIGURE 2
Distribution of study participation duration among participants who withdrew and who did not withdraw.

TABLE 4 Adherence rates across bursts in subsample of participants that contributed data to burst 1 and/or 2 and burst 3 and/or 4.

Completed EMA surveys	n (participants)	Adherence (mean %)	Standard deviation
burst1	42	82.6	22.7
burst2	40	80.8	23.8
burst3	38	82.8	18.6
burst4	38	80.1	22.0

operating system ($F(1,212) = 0.125$, $p = 0.26$). When combining all significant predictors into one model, study participation duration and time between study start and enrollment remain significant predictors of adherence, while reaching burst 3 was not significant anymore ($F(1, 210) = 2.97$, $p = 0.08$), indicating its dependency on time elapsing (see **Figure. 2** for a histogram of study participation duration in participants that withdrew and that did not withdraw).

Adherence to EMA surveys does not differ between bursts

Within those participants ($n = 42$) that did not withdraw and that did contribute data to burst 1 and/or 2 and to burst 3 and/or 4 we found, contrary to our hypothesis, that adherence did not significantly differ between bursts 1/2 (81.7% ± 23.1) and bursts 3/4 (81.4% ± 20.3; $F(1, 116) = 0.002$, $p = 0.965$). When analyzing the four bursts separately, adherence during all four bursts was above 80% (see **Table 4**).

Conclusions and discussion

This longitudinal, observational study was the first of its kind, using daily smartphone EMA self-report surveys to investigate baseline and structural factors that predict study withdrawal and adherence in older adults across an extended time period covering multiple phases of the COVID 19 pandemic. The objective of this study was to assess factors that predict withdrawal and adherence to EMA in older adults over the course of a seventeen-month period, beginning during the sudden and long-term shutdown of normal behavior and routines and a forced shift to digital technology during the pandemic.

Our *main results* are that withdrawal was associated with less research staff changes and was less likely among participants who reached the study mid-point. No baseline characteristics predicted withdrawal. Main reasons for withdrawal were communication issues, i.e. staff not being able to contact participants.

We found an adherence rate of 82% and no fatigue effects. Adherence was predicted by education status, study participation duration, reaching the study midpoint and time

between study start and enrollment. COVID infections or supporting people in the household was not related to adherence.

Our *first aim* was to explore reasons for *withdrawal* and if any personal baseline factors predicted withdrawal in older adults, including age, gender, race/ethnicity, employment, residence, number of people in residence, education, location, generations in the household, how often the house was left pre-pandemic and previous participation in an active treatment arm of the prior study randomized controlled trial (vs. a control condition). Further, we examined if study structural factors, including study research staff turnover, smartphone operating system, time elapsed between study start and enrollment and reaching the study mid-point predicted withdrawal. Based on the available literature in older adults, we hypothesized to find withdrawal due to participant burden (8) or communication issues (22).

Withdrawal was not predicted by baseline factors, but by number of changes of RA's and making it beyond the mid-study point. While it seems intuitive that people who complete the study reached the study mid-point, it is interesting that completing the study was associated with having 1 RA change,—and when combining significant structural factors into one model- with 2 RA changes compared to experiencing no change in RA. One potential explanation for this finding could be that participants who withdrew before/during the first burst did not experience any change in RA's, however when excluding participants who withdrew before or during the first burst, the results remain unchanged, indicating that this finding is not solely driven by having less opportunity to experience an RA change in the case of an earlier withdrawal. Further explanations could be that our findings suggest that once participants take part in the study for a certain amount of time (i.e. reached the study mid-point), they are likely to complete the study, potentially due to a sense of duty to finish what they started, even in the face of changes in personnel which was experienced more often the longer one was in the study. Furthermore, while changes in RA's depend on study participation duration, experiencing RA changes does not necessarily predict withdrawal. A possible interpretation could be that being assigned to a new RA led to more contact with study personnel and to different ways of explaining instructions on how to handle the study app, which might have been pleasing and helpful for participants, motivating them to stay in the study. Another explanation could be that participants did not want to withdraw upon a new contact in order to avoid making the newly assigned RA feel responsible for their withdrawal. While our study experienced a complex flow of personnel during the pandemic with many changes in the longer break in the middle of the study, our findings demonstrate that personnel change in itself is not predictive of withdrawal of participants and may actually help keep people in the study. We further demonstrate that the main reasons for withdrawal were communication issues, underlining the importance of forming reliable relationships with participants, that might even overcome technical issues, which we found to be only the fourth most frequent reason for withdrawal. Anecdotally, participants were very tolerant of technical issues if the situation was communicated transparently. This is in line with the literature emphasizing the importance of trust for the adaptation of digital technologies by older people (29) and with a study describing as a primary withdrawal reason that participants did not fully understand what they were supposed to do (22). Another withdrawal reason that was mentioned in one study was medical emergency (8), which we do not find in our study as a main withdrawal reason, possibly due to the very good initial health of our participants or a potential reduction of non-urgent medical treatments conducted during the pandemic.

These findings on withdrawal can be seen parallel to a previous finding on adherence, with early adherence predicting study-long adherence in adults (3). We further observed that adherence did not predict withdrawal, indicating that different factors might be at play in the two situations.

The *second aim* was to examine if any personal baseline or study structural factors predicted *adherence*—per burst—in older adults, based on the same personal and structural factors explored in relation to withdrawal. We hypothesized that adherence will be predicted by age and gender with highest adherence in younger women based on the literature on specific findings on older adults (21), but not by other personal factors (8, 19). Based on the literature, we further hypothesized that an Android operating system would be related to lower adherence (23).

Consistent with the EMA literature in older adults (8, 16, 19, 21), we found an adherence rate of 82% and no fatigue effects. Inconsistent with our hypothesis, we found no relationship between gender and age with adherence, which is line with a recent study in older people that reported no relation of age to completion or response rate, but found that older participants were more likely to report not being alerted to surveys (23): inconsistencies in findings might be due because adherence in older adults might be malleable by prompt support if participants contact research staff if they notice that their alerts do not come through as planned. Further, we found that more years of education were associated with higher adherence, potentially due to an increased exposure of our participants to research along their educational path and therefore potentially a sense of duty to contribute to research studies. There was a trend towards lower adherence among those living with two or more people compared to living alone. As a follow-up analysis, we hypothesized that people whose family member had a COVID-19 diagnosis or who were giving support to family members might be more likely to show such a negative

relationship between the number of people living in the household and adherence rate. We observed no moderating effect of these two variables, and they also did not predict adherence themselves, adding novel evidence to the literature that EMA surveys can be a reliable tool during a pandemic in older adults and further that the adherence of older adults to EMA surveys was little impacted by the pandemic.

Regarding structural factors, the association of higher adherence with longer study participation and reaching the study mid-point is in line with the literature (3). Further we observed higher engagement in participants that enrolled faster in the study, potentially because participants that were highly motivated contacted the study staff faster, which is also in line with the literature (23). Additionally, we found no relationship of smartphone operating system with adherence, which is opposed to a previous finding of higher EMA response rate in iOS vs. Android users in a week-long study in older adults (23). This difference might be due to the difference in the duration of our study which lasted over several months vs. a 1-week long study, in such a way that a longer study duration allowed for more occasions to practice and get familiar with potential pitfalls of a specific operating system.

The *final aim* was to understand how the rate of completed EMA surveys from the first two bursts compared to the last two bursts. We hypothesized that adherence would be significantly less during burst 3 and 4 due to participant's reengaging in activities outside the home after spending more time at home during the pandemic (30).

Inconsistent with our hypothesis, interestingly, we did not find a difference in adherence between different time-points in the pandemic among those participants that did continue their study participation beyond the study mid-point– those who remained in the study were as adherent after lifting of restrictions as the larger group had been during early months of the pandemic, a finding that is to some extent in line with a previous finding that early adherence predicted study-long adherence in adults (3) and contrary to our initial hypothesis that more activities, as the stay-at-home restrictions lift, might lead to less adherence. This also adds weight to the interpretation that participants develop a loyalty towards the study, as their participation continues, potentially due to increasing trust and relationships with study personnel.

Strengths of this study include the extended time period covered during multiple phases of the pandemic and the availability of detailed pre-pandemic and during-pandemic baseline and structural variables.

Limitations of our study include the small sample size and bias of our sample: our sample has a high percentage of female participants and is not very diverse in terms of race, ethnicity and socioeconomic background. This is relevant especially because people from different backgrounds might have been affected differently by the pandemic. Further, all our participants owned their own smartphone, which might not be the case for socioeconomic environments different from the US, with rates of smartphone ownership among older adults below 35% in emerging economies (31), which further limits the generalizability of our results. Another factor is our study design that included a bonus payment of $20 for the completion of more than 85% of surveys, which was however not very high given the length of the study. Additionally, it is important to interpret the findings of our study in the light of relatively few surveys per day (2) and the possibility for scheduling a time window of 1 h for the random timing of the surveys, supporting a higher adherence. Further, it is a highly specific subgroup of participants who were concurrently enrolled in the MEDEX study that included some EMA components. However, the MEDEX EMA surveys were completed on study-provided tablets, while the StayWELL study used the participant's smartphone, making the present EMA experience different to some extent in regard to the technical aspects. Therefore, participants in our sample were familiar with completing EMA surveys and had previously demonstrated good adherence, since all the recruited participants could be considered previous study completers and currently active study participants of the MEDEX study. While we did not assess this question, it was previously found that the majority of experienced EMA participants preferred to use their own smartphone vs. study-provided devices (23), which might add, on top of the training effect of previous EMA study participation, a familiarity with handling of the participant's own phone. Whether these benefits outweigh the benefits of study-provided phones in preventing technical difficulties should be carefully considered in future studies, taking into consideration also costs of study-phones and the downside of having to manage two phones at once for the participants. In addition to familiarity with EMA procedures, the participants seemed to have found their previous study participation to be beneficial,, which seemed to motivate them to take part in the StayWELL study, which is in line with a previous finding, that previous study participation has led to higher enrollment and completion of an EMA study compared to inexperienced participants (23), potentially due to more contact/stronger relationships with research personnel and positive experiences. Furthermore, we did not follow up in detail about withdrawal reasons once withdrawal was communicated. Anecdotally, some of the detailed reasons related to communication issues were that some participants were difficult to reach for study staff *via* phone, due to e.g. not picking the phone up for non-identified caller IDs or due to being out/traveling. Reasons related to technical difficulties included running out of mobile data on the phone, too high battery drain or that the surveys did not come through to the phone. Based on individual, not systematically collected, more detailed feedback, it also became evident that for some participants it was not clear

that their surveys would arrive at random times within the specific 1-hour time interval, and they were irritated because the surveys did not arrive on time with regular schedules. Further, the longer break between the second and third burst led some participants to believe that the study had ended without timely communication with them.

Future methodological questions arising are how to address missingness of data, whether to focus on subject withdrawal or to report adherence in the form of missing data per participant, since few outliers can have a large effect on the overall adherence rates (3, 16). Strategies for averaging data over the day or week have to be considered carefully, so information is not lost that might be valuable in predicting changes in adherence (2). Further, additional factors influencing adherence need to be considered: In EMA studies in healthy adults (18–65 years), higher anxiety and depression variables were correlated with lower adherence (3). A future outlook for our study is therefore to investigate the relationship of mood states with EMA adherence in older adults in order to contribute to recommendations how to best design EMA surveys that assess wellbeing in older adults (2). Additionally, an important next step will be to investigate whether these findings also apply to a diverse group of participants.

To *conclude*, EMA can potentially reach research participants unable to attend in-person study visits (e.g. due to restricted mobility) and allow for immediate collection with less bias. We hereby demonstrate that EMA surveys in older adults during unusual circumstances, such as a pandemic, to assess daily experiences are feasible and that engagement of older adults with EMA was high and little impacted by the phase of the pandemic. By investigating the factors underlying engagement and adherence to assess mental wellbeing using EMA in older adults in a situation of reduction of mobility, we identify barriers to consider when designing digital health technologies for research and clinical use in older people. We found that withdrawal and adherence are robust to study alterations and not associated with staff changes when accounting for study duration. This should be kept in mind when recruiting participants on the one hand, but also when building a cohort of participants and patients that might be willing to participate in several or longitudinal digital health studies or treatments on the other hand. We further demonstrate that communication issues were the largest contributing factor to withdrawal, suggesting that future digital health interventions should invest in an easy-to-use communication strategy for participants or patients that would like to contact their care provider or study personnel, a recommendation that also applies to study personnel, for which ease of use of digital platforms will also be a relevant factor in the success of a study. A further important finding is that the mobile software used by participants did not affect adherence or withdrawal, allowing for a broad distribution of digital health technologies. Our study adds information on single predictive variables that affect compliance to and withdrawal from EMA smartphone surveys in older people that can inform the design of future digital survey studies to maximize engagement and reliability of studies using EMA.

Contribution to the field statement

Digital surveys can potentially reach people unable to attend in-person visits and allow for immediate collection with less bias. By investigating the factors that predict withdrawal from and adherence to a study on mental wellbeing using smartphone-based, momentary surveys in older adults during the COVID-19 pandemic, our study adds significant detail to the limited literature on factors that potentially affect engagement in this specific population. We demonstrate novel findings that digital surveys can be a reliable tool during a pandemic in older adults and that the adherence of older adults was little impacted by the pandemic. We contribute to the identification of barriers to the participation in digital technologies for older people and recommend that future interventions should invest in an easy-to-use communication strategy for participants or patients that would like to contact their care provider or study personnel. Our study adds information on single predictive variables that affect compliance to smartphone surveys in older people to inform the design of future digital health technologies and maximize engagement and reliability of results of studies using digital momentary surveys.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author (FK) upon reasonable request and after filing an institutional data sharing agreement. The data are not publicly available due to the privacy of research participants. Requests to access the datasets should be directed to Federica Klaus, fklaus@health.ucsd.edu.

Ethics statement

The studies involving human participants were reviewed and approved by UCSD's Institutional Review Board. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

Conception and design of study: FK, LTE, RCM. CAD. Acquisition of data: FK, EP, AQ, AS, DS. Analysis of data: FK, LTE. Drafting the manuscript: FK, EP, AQ, LTE. Revising the manuscript: FK, EP, AQ, AS, DS, RCM, CAD, LTE. All authors contributed to the article and approved the submitted version.

Funding

Early. Postdoc Mobility Fellowship of the Swiss National Science Foundation (SNSF) (grant number P2ZHP3_181506) (FK) and Novartis Foundation for Medical-Biological Research Fellowship (FK). T. Denny Sanford Institute for Empathy and Compassion, Center for Research on Empathy and Compassion at University of California San Diego Seed Grant Program 2021. Salary support for RCM was provided, in part, by the NIA (R01AG062387; R01AG070956).

Acknowledgments

We thank Alan De La Cruz, Alyssa Arnold, Eleonora Vaccariello, Emma Parrish, Isabella Davis, Jeremy Delgadillo, Jessica Carrasco, Kyle Mitchell, Marlen García, Mary K. Duff, Nancy Mendoza, Rosa Montoya, Yanlin Liu for their contribution to data collection and study management. We thank the MEDEX team, especially Michelle Voegtle and Andrew Scott for support in recruitment and the participants for their time and engagement.

Conflict of interest

RM is a co-founder of KeyWise, Inc. and a consultant for NeuroUX. This study received funding from the Novartis Foundation for Medical-Biological Research. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication. All authors declare no other competing interests.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fdgth.2022.920706/full#supplementary-material.

References

1. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. *Annu Rev Clin Psychol.* (2008) 4:1–32. doi: 10.1146/annurev.clinpsy.3.022806.091415

2. de Vries LP, Baselmans BML, Bartels M. Smartphone-based ecological momentary assessment of well-being: a systematic review and recommendations for future studies. *J Happiness Stud.* (2021) 22(5):2361–408. doi: 10.1007/s10902-020-00324-7

3. Jones SE, Moore RC, Pinkham AE, Depp CA, Granholm E, Harvey PD. A cross-diagnostic study of adherence to ecological momentary assessment: comparisons across study length and daily survey frequency find that early adherence is a potent predictor of study-long adherence. *Pers Med Psychiatry [Internet].* (2021) [cited 2021 Nov 18] 29. p. 100085. doi: 10.1016/j.pmip.2021.100085. Available from: https://www.personalizedmedpsych.com/article/S2468-1717(21)00017-X/fulltext

4. Williams-Kerver GA, Schaefer LM, Hazzard VM, Cao L, Engel SG, Peterson CB, et al. Baseline and momentary predictors of ecological momentary assessment adherence in a sample of adults with binge-eating disorder. *Eat Behav.* (2021) 41:101509. doi: 10.1016/j.eatbeh.2021.101509

5. Jones A, Remmerswaal D, Verveer I, Robinson E, Franken IHA, Wen CKF, et al. Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. *Addict Abingdon Engl.* (2019) 114(4):609–19. doi: 10.1111/add.14503

6. Wen CKF, Junghaenel DU, Newman DB, Schneider S, Mendez M, Goldstein SE, et al. The effect of training on participant adherence with a reporting time frame for momentary subjective experiences in ecological momentary assessment: cognitive interview study. *JMIR Form Res.* (2021) 5(5):e28007. doi: 10.2196/28007

7. Gershon A, Kaufmann CN, Torous J, Depp C, Ketter TA. Electronic Ecological Momentary Assessment (EMA) in youth with bipolar disorder: demographic and clinical predictors of electronic EMA adherence. *J Psychiatr Res.* (2019) 116:14–8. doi: 10.1016/j.jpsychires.2019.05.026

8. Maher JP, Rebar AL, Dunton GF. Ecological momentary assessment is a feasible and valid methodological tool to measure older Adults' physical activity and sedentary behavior. *Front Psychol.* (2018) 9:1485. doi: 10.3389/fpsyg.2018.01485

9. Jones SL, Hue W, Kelly RM, Barnett R, Henderson V, Sengupta R. Determinants of longitudinal adherence in smartphone-based self-tracking for chronic health conditions: evidence from axial spondyloarthritis. *Proc ACM Interact Mob Wearable Ubiquitous Technol.* (2021) 5(1):16:1–24. doi: 10.1145/3448093

10. Fact Sheet Pew Research Center. Demographics of Mobile Device Ownership and Adoption in the United States [Internet]. Mobile Fact Sheet Pew Research Center. [cited 2022 Feb 24]. Available from: https://www.pewresearch.org/internet/fact-sheet/mobile/

11. Smith A, Pew Research Center. U.S. Smartphone Use in 2015 [Internet]. Pew Research Center: Internet, Science & Tech. (2015) [cited 2022 Feb 24]. Available from: https://www.pewresearch.org/internet/2015/04/01/us-smartphone-use-in-2015/

12. Monica Anderson, Perrin Andrew, Pew Research Center. Technology use among seniors [Internet]. Pew Research Center: Tech adoption climbs among older adults. (2017) [cited 2022 Feb 24]. Available from: https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/

13. Pew Research Center. *Share of those 65 and older who are tech users has grown in the past decade.* Pew Research Center. (2021) [cited 2022 Jun 29].

Available from: https://www.pewresearch.org/fact-tank/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/

14. Vaportzis E, Giatsi Clausen M, Gow AJ. Older adults perceptions of technology and barriers to interacting with tablet computers: a focus group study. *Front Psychol.* (2017) 8:1687. doi: 10.3389/fpsyg.2017.01687

15. Berkowsky RW, Czaja SJ. Challenges associated with online health information seeking among older adults. In: Pak R, McLaughlin AC, editors. *Aging, technology and health*. San Diego, CA, United States: Elsevier (2018) [cited 2022 Feb 24]. p. 31–48. Available from: https://linkinghub.elsevier.com/retrieve/pii/B9780128112724000026

16. Cain AE, Deep CA, Jeste DV. Ecological momentary assessment in aging research: a critical review. *J Psychiatr Res.* (2009) 43(11):987–96. doi: 10.1016/j.jpsychires.2009.01.014

17. Bell IH, Lim MH, Rossell SL, Thomas N. Ecological momentary assessment and intervention in the treatment of psychotic disorders: a systematic review. *Psychiatr Serv.* (2017) 68(11):1172–81. doi: 10.1176/appi.ps.201600523

18. Campbell LM, Paolillo EW, Heaton A, Tang B, Depp CA, Granholm E, et al. Daily activities related to Mobile cognitive performance in middle-aged and older adults: an ecological momentary cognitive assessment study. *JMIR MHealth UHealth.* (2020) 8(9):e19579. doi: 10.2196/19579

19. Fritz H, Tarraf W, Saleh DJ, Cutchin MP. Using a smartphone-based ecological momentary assessment protocol with community dwelling older African Americans. *J Gerontol Ser B.* (2017) 72(5):876–87. doi: 10.1093/geronb/gbw166

20. Graham SA, Stein N, Shemaj F, Branch OH, Paruthi J, Kanick SC. Older adults engage with personalized digital coaching programs at rates that exceed those of younger adults. *Front Digit Health.* (2021) 3:93. doi: 10.3389/fdgth.2021.642818

21. Liu H, Lou VWQ. Developing a smartphone-based ecological momentary assessment protocol to collect biopsychosocial data with community-dwelling late-middle-aged and older adults. *Transl Behav Med.* (2019) 9(4):711–9. doi: 10.1093/tbm/iby096

22. Hnatiuk SH. Experience sampling with elderly persons: an exploration of the method. *Int J Aging Hum Dev.* (1991) 33(1):45–64. doi: 10.2190/541V-1036-8VFK-J0MU

23. Burke L, Naylor G. Smartphone app-based noncontact ecological momentary assessment with experienced and naïve older participants: feasibility study. *JMIR Form Res.* (2022) 6(3):e27677. doi: 10.2196/27677

24. Sin F, Berger S, Kim IJ, Yoon D. Digital social interaction in older adults during the COVID-19 pandemic. *Proc ACM Hum-Comput Interact.* (2021) 5 (CSCW2):380:1-380:20. doi: 10.1145/3479524

25. Hoel V, Seibert K, Domhoff D, Preuß B, Heinze F, Rothgang H, et al. Social health among German nursing home residents with dementia during the COVID-19 pandemic, and the role of technology to promote social participation. *Int J Environ Res Public Health.* (2022) 19(4):1956. doi: 10.3390/ijerph19041956

26. Ellison-Barnes A, Moran A, Linton S, Chaubal M, Missler M, Pollack CE. Limited technology access among residents of affordable senior housing during the COVID-19 pandemic. *J Appl Gerontol.* (2021) 40(9):958–62. doi: 10.1177/07334648211013634

27. Wetherell JL, Ripperger HS, Voegtle M, Ances BM, Balota D, Bower ES, et al. Mindfulness, Education, and Exercise for age-related cognitive decline: study protocol, pilot study results, and description of the baseline sample. *Clin Trials.* (2020) 17(5):581–94. doi: 10.1177/1740774520931864

28. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol.* (1995) 57(1):289–300. ISSN: 0035-9246. Available from https://www.jstor.org/stable/2346101

29. Fox G, Connolly R. Mobile health technology adoption across generations: narrowing the digital divide. *Inf Syst J.* (2018) 28(6):995–1019. doi: 10.1111/isj.12179

30. Portegijs E, Keskinen KE, Tuomola EM, Hinrichs T, Saajanaho M, Rantanen T. Older adults' activity destinations before and during COVID-19 restrictions: from a variety of activities to mostly physical exercise close to home. *Health Place.* (2021) 68:102533. doi: 10.1016/j.healthplace.2021.102533

31. Silver L. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally. Pew Research Center's Global Attitudes Project. (2019) [cited 2022 Aug 1]. Available from: https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 31 August 2022
DOI 10.3389/fdgth.2022.869790

OPEN ACCESS

EDITED BY
Raeanne Moore,
University of California San Diego, United States

REVIEWED BY
Sarah Anne Graham,
Lark Health, United States
Matthew O'Connell,
King's College London, United Kingdom

*CORRESPONDENCE
Kaitlin B. Casaletto
kaitlin.casaletto@ucsf.edu

SPECIALTY SECTION
This article was submitted to Human Factors and Digital Health, a section of the journal Frontiers in Digital Health

RECEIVED 05 February 2022
ACCEPTED 08 August 2022
PUBLISHED 31 August 2022

CITATION
VandeBunte A, Gontrum E, Goldberger L, Fonseca C, Djukic N, You M, Kramer JH and Casaletto KB (2022) Physical activity measurement in older adults: Wearables versus self-report.
Front. Digit. Health 4:869790.
doi: 10.3389/fdgth.2022.869790

Physical activity measurement in older adults: Wearables versus self-report

Anna VandeBunte[1,2], Eva Gontrum[1], Lauren Goldberger[1], Corrina Fonseca[1], Nina Djukic[1], Michelle You[1], Joel H. Kramer[1] and Kaitlin B. Casaletto[1]*

[1]Department of Neurology, Memory and Aging Center, University of California, San Francisco, San Francisco, CA, United States, [2]Department of Psychology, Palo Alto University, Palo Alto, CA, United States

Physical activity (PA) is associated with preserved age-related body and brain health. However, PA quantification can vary. Commercial-grade wearable monitors are objective, low burden tools to capture PA but are less well validated in older adults. Self-report PA questionnaires are widely accepted and more frequently used but carry inherent limitations. We aimed to compare these commonly used PA measures against one another and examine their convergent validity with a host of relevant outcomes. We also examined the factors that drive differences in PA self-reporting styles in older adults. 179 older adults completed 30-day Fitbit Flex2™ monitoring and reported PA levels *via* two widely used PA questionnaires: PASE and CHAMPS-METs (metabolic expenditure calories burned). Participants also completed measures of cardiometabolic (hypertension diagnosis, resting heart rate, A1C levels), cognitive (memory, processing speed, executive functioning), and brain MRI (medial temporal lobe volume) outcomes. The discrepancy between objective Fitbit monitoring and self-reported PA was evaluated using a sample-based z difference score. There were only modest relationships across all PA metrics. Fitbit step count demonstrated a stronger association with the PASE, whereas Fitbit calories burned was more strongly associated with CHAMPS-MET. Fitbit outcomes had more consistent convergence with relevant outcomes of interest (e.g., cardiometabolic and brain health indices) when compared to subjective measures; however, considerable heterogeneity within these associations was observed. A higher degree of overreporting was associated with worse memory and executive performances, as well as hypertension diagnoses. We build on prior findings that wearable, digital health indicators of PA demonstrate greater construct validity than self-report in older adults. We further show important clinical features (e.g., poorer cognitive status) of older adults that could contribute to a higher level of overreporting on self-report measures. Characterization of what PA measures truly operationalize will help elucidate relationships between most relevant facets of PA and outcomes of interest.

KEYWORDS

healthy aging, fitbit, actigraphy, CHAMPS, PASE

1. Introduction

Physical activity (PA) is associated with preserved vascular health, brain structure, and cognition with age (1–3). In the context of neurodegenerative diseases, active lifestyles are also linked with less functional decline and reduced risk of dementia (2, 4, 5). However, physical activity is a broad construct, and its operationalization can vary widely depending on the measures employed, possibly leading to imprecision and/or inconsistencies in the literature linking physical activity to brain health.

The most commonly utilized measures of PA may capture several aspects of activity. Actigraphy monitors are useful for objective assessment of free-living movement levels, ranging from everyday chores to structured exercise. Given their relative low cost and user friendly interfaces, commercially available wearables (like Fitbit actigraphy monitors) have garnered increasing attention and validation of metrics for reliable measurement of PA (6, 7). For instance, Fitbit monitors have demonstrated inter-device reliability with other actigraphy monitors (i.e., Actigraph GT3X+ accelerometer) and positive correlations with observed step count and gait speed on a treadmill (8–10). Within the Fitbit suite of outcomes, average steps has demonstrated the most robust validity though there is less validity for energy expenditure as a measurement of PA (11). However, few studies have been conducted in older adults who may have varying experience using wearable devices. As digital health tools expand, understanding their utility in populations that may be most vulnerable (e.g., older adults) is needed. Moreover, considering the importance of engaging in PA for healthy aging, it important to understand the convergence across measures, how measures of PA may differ, and what are predictors of discrepancies between measures.

In the absence of actigraphy metrics, standard self-report measures of PA are used in older adults to capture more structured activities and exercise routines (e.g., duration, type of exercise). The Community Healthy Activities Model Program for Seniors (CHAMPS) and Physical Activity Scale for the Elderly (PASE) are two such widely used self-report questionnaires of PA in older adults. However, these subjective measures may carry inherent limitations and capture only segments of PA. For instance, the PASE evaluates the level of activity (e.g., frequency and duration), whereas the primary outcome for CHAMPS is metabolic expenditure. When considering utility, each is relatively quick, well-validated, and inexpensive to use. PASE has previously been associated with portable accelerometer readings, walking steps, and energy expenditure (12, 13). Higher PASE scores (indicating a greater degree of PA) have also been correlated with reduced likelihood of cardiometabolic, neurological, and psychological health conditions (14, 15). PASE has also demonstrated meaningful associations with health status and physiologic measures, such as heart rate and static balance (13, 16). Similarly, the CHAMPS has been shown to relate to a host of relevant outcomes, such as physical functioning and psychological health (17). Similar to the PASE, the CHAMPS has demonstrated positive associations with total minutes of movement measured by an accelerometer and corresponding intensity, as well as measures of fitness capacity (i.e., the 6-min walk test) and lower body physical functioning (i.e., Short Physical Performance Battery) (17–19). However, such self-report tools may be limited by scope and subjectivity. For instance, self-report measures typically underestimate sedentary time compared to real-time digital health measures, such as an accelerometer or inclinometer (20). This is particularly important when using subjective measurements to assess older adults with cognitive difficulties, as there is greater risk of recall bias (21). Furthermore, individuals who have been encouraged to engage in exercise (e.g., many older adults by their physicians) have demonstrated the tendency to engage in more overreporting, perhaps related to well-known effects of social desirability bias (22).

To date, there is a gap in the literature directly comparing and evaluating PA as assessed across multiple standardly employed measures. Studies have not pragmatically demonstrated how metrics within the Fitbit suite compare against widely used self-report measures of PA in older adults. Similarly, research has yet to compare commonly used self-report measures (i.e., CHAMPS and PASE) alongside a comprehensive panel of relevant neurologically relevant aging outcomes (e.g., cognition, MRI outcomes) to characterize their convergent validity for use in brain aging studies. Furthermore, it is uncertain whether there are discrepancies in reporting styles across self-report measures, and if particular participant characteristics systematically predict older adults who over or under report. The current study will begin to elucidate some of these relationships to contribute to our understanding PA measurement tools in older adults.

In the current study, we aimed to (1) determine the comparability across commonly used self-reported measures of physical activity (PA) and Fitbit-based actigraphy metrics, (2) examine the convergent validity of Fitbit, PASE, and CHAMPS using a comprehensive panel of demographic, cardiometabolic, cognitive, and brain structural outcomes, and (3) examine the person-specific factors that characterize overreporting on self-report measures We hypothesized that objective measures of PA *via* Fitbit would demonstrate the best construct validity and that the degree of overreporting would relate to poorer neurobehavioral status.

2. Materials and methods

One hundred seventy-nine older adults enrolled in the UCSF Memory and Aging Center's Longitudinal Aging Study

TABLE 1 Descriptive statistics.

	n	% or *M* (*SD*)
Sex, % female	105	58.66%
Race		
White	153	85.47%
Black	2	1.12%
Asian	19	10.61%
Other	5	2.79%
Age (years)	179	73.50 (8.23)
Education (years)	179	17.57 (1.85)
Fitbit steps (daily average)	179	7840.77 (3365.11)
Fitbit calories (daily average)	179	1862.27 (426.51)
PASE (possible range 0 to >500)	105	126.10 (60.66)
CHAMPS-MET (max calories burned in a week)	85	4062.76 (2275.75)
Hypertension, % yes	116	37.93%
Resting heart rate (bpm)	165	66.62 (9.51)
Hemoglobin A1C (%, normal range 4.3–5.6)	97	5.47 (0.33)
Technology Familiarity Questionnaire (Q8)[d]	128	4.50 (0.68)
Technology Familiarity Questionnaire (Q9)[e]	128	4.86 (0.41)
Memory (z-score)[a]	124	−0.07 (0.87)
Executive functioning (z-score)[c]	132	0.77 (0.58)
Processing speed (z-score)[b]	126	−2.60 (1.58)
Medial temporal lobe volume (voxels, 1 cm^3)	72	9.80 (1.05)

Note. $N = 175$.
[a]z-scores on these tests represent performances compared to the larger Hillblom Aging cohort of older adults.
[b]z-score represents performance compared to young adults (20–30 years old).
[c]z-score derived from EXAMINER normative study group (adults aged 18–80+).
[d]Question 8: "How much difficulty do you have using computers?" (Range 1–5, 1 = extreme difficulty, 5 = no difficulty).
[e]Question 9: "How anxious (or nervous) do you typically feel when using a computer, tablet, or smartphone?" (Range 1–5, 1 = extremely anxious, 5 = not anxious).

who choose to participate in 30-day Fitbit monitoring (average daily steps and calories burned), and who completed at least one measure of self-reported physical activity levels (PASE, $n = 105$; CHAMPS-MET, $n = 85$) were included in the study (see Table 1). Participants completed comprehensive neurological and neuropsychological evaluations, as well as structural neuroimaging, cardiometabolic measures, and a study partner interview. Following evaluations, participants were reviewed at a case conference with board certified neurologists and neuropsychologists. Inclusion criteria for enrollment consisted of: (1) no current evidence of a memory or neurological condition (e.g., stroke, epilepsy), (2) no functional decline as operationalized as a Clinical Dementia Rating (CDR) scale of 0–0.5 *via* study partner interviews, (3) no history or evidence of DSM-5 major psychiatric disorders, active substance abuse, hepatitis C, blindness, deafness, HIV, and syphilis. Participants had very minimal cardiovascular medical histories (Myocardial Infarction, $n = 4$, Cerebrovascular Accident, $n = 2$, Transient Ischemic Attack, $n = 2$; note, all cardiovascular events occurred >5 years prior to study participation). The study was approved by the institutional review board of the University of California, San Francisco and is conducted in accordance with the latest Declaration of Helsinki, including written informed consent from all participants.

2.1. Procedure

At the in-person visit, participants neuropsychological evaluations, brain MRI, self-reported measured, and Fitbit set-up. Participants then completed 30 days of subsequent Fitbit monitoring before mailing their device back to study personnel. All data from Fitbit devices were synced to the Fitabase platform and data were downloaded for quality control, cleaning, and analysis.

2.2. Actigraphy monitoring

The FitBit Flex2™ (Fitbit Inc., San Francisco, CA, USA; https://www.fitbit.com) recorded average daily steps and calories burned. The FitBit Flex2™ is a thin, flexible, Bluetooth fitness tracker with no visible record of physical activity measurements, with a three-axis acceleration sensor, and with the capability to store 7 days of detailed motion data. Participants were blinded to all notifications and indication of the duration of exercise for the 30-day time period. They were instructed to wear the Fitbit during all waking hours, and to synchronize nightly with their smartphone *via* Bluetooth 4.0 before charging at night. In cases where the participant did not have a Fitbit-compatible smartphone, the Fitbit was synchronized to an iPad and aggregate daily physical activity data was collected at the completion of the 30-day period. Fitbit accounts for each participant were connected to Fitabase, a platform specifically tailored for wearable research data management.

Average daily steps and average daily calories burned were selected as the primary objective outcomes of interest. Fitbit averages were calculated by taking the average daily steps and calories burned for the first 20 days of available monitoring data. Individual days with fewer than 100 steps were removed from the analyses to control for nonadherence. Participants were only included if they had at least 14 days of available monitoring data, and the first 20 days of the 30-day monitoring data were used in the analyses.

2.3. Physical activity questionnaires

2.3.1. Physical activity scale for the elderly

Self-reported physical activity was measured using the Physical Activity Scale for the Elderly (PASE), a widely

validated measure of self-reported activity levels for older adults. Participants were asked to rate the frequency, duration, and intensity of activity in three domains (leisure, household, and work-related activity) over the past seven days with the 11-item questionnaire. Utilizing the scoring manual, activity scores were computed by multiplying activity frequencies by the task-specific weights (16). Activity scores were then summed to obtain a total score representing overall physical activity level, with higher values indicating greater activity.

2.3.2. CHAMPS-MET

The Community Healthy Activities Model Program for Seniors physical activity questionnaire (CHAMPS) was administered to assess the variety of physical activities that older adult participants may engage in, from less intensive forms such as walking or stretching to more vigorous exercise routines (17). The questionnaire includes 41 items to evaluate the frequency and duration of light, moderate, and vigorous activities that were performed weekly over the last four weeks. Participants reported whether they participated in an activity during the four-week period and then selected the hours per week spent participating in the activity, rating the duration on a six-point scale from less than 1 to 9 or more hours. Each activity corresponds to a metabolic weight or MET value. Estimated caloric expenditure was calculated by multiplying the estimated duration of each activity by the corresponding MET value, in alignment with published guidelines (17).

2.4. Cognitive outcomes

Participants completed a neuropsychological battery assessing cognitive outcomes hypothesized to be associated with physical activity (23, 24). This brief standardized battery has been previously described and validated to be neuroanatomically sensitive to age-related neurodegeneration (25, 26).

2.4.1. Episodic memory

Verbal episodic memory was measured by the California Verbal Learning Test (CVLT-II) and a modified version of the Benson Figure Memory test. The CVLT-II includes a 16-item list presented over five learning trials, followed by free and cued recall of the list after an interference trial, and then again after a 20-min delay. Following the long delay, participants were given a list of 44 words and asked to discriminate between the target word and a distractor item (recognition trial). Outcome metrics included words correctly recalled after delays and recognition discrimination performance.

To assess visual memory, participants were asked to draw the modified Benson figure from memory after a 10-min delay. Recall of the figure was scored on a 17-point scale (25).

Sample-based z-scores were created for outcomes on both measures and averaged together to create an episodic memory composite.

2.4.2. Processing speed

Processing speed was assessed through five computerized tests of reaction time to different visual stimuli (dots, lines, search, shapes, abstract matching 1, abstract matching 2) (27). All tasks included a practice trial period where the participant had to perform at greater than 70% accuracy in order to continue to the test trials. Sample-based z-scores were created for each of the five tasks to calculate a processing speed composite score.

2.4.3. Executive functioning

Executive functions were measured by the NIH EXAMINER (28). NIH EXAMINER includes a composite score of five computer-based tests of working memory (dot counting, 1-back, 2-back), response inhibition (enclosed flanker), and set shifting (set shifting), and two verbally mediated tests of generativity (D-word and animal fluency). All computerized tasks included at least three practice trials.

2.5. Cardiometabolic outcomes

2.5.1. Hemoglobin A1C

Whole blood and serum samples were collected and stored in 0.5 ml aliquots at −80 °C following baseline 12 h fasting blood draws, until used for biochemical processing. All laboratory analyses were performed by UCSF Clinical Laboratories, a CLIA-certified, CAP-accredited laboratory at UCSF Mission Bay Hospital. Hemoglobin A1C (HbA1C) levels were determined from whole blood by means of an Abbott Architect c8000 enzymatic immunoassay.

2.5.2. Resting heart rate

Participant resting heart rates were measured by a clinician or study staff. A normal resting heart rate for adults ranges from 60 to 100 beats per minute (29). Elevated heart rate is a risk factor for cardiovascular morbidity and mortality (30).

2.5.3. Hypertension

Personal medical history, including presence of hypertension, was collected *via* self-report during the clinical history gathering with a neurologist. If the participant did not report a history of hypertension, but indicated taking an antihypertensive drug, she/he was asked by the clinician or study staff if they were prescribed the medication for their blood pressure. If answered yes, the participant was marked as having a history of hypertension.

2.6. Brain MRI

2.6.1. Structural neuroimaging

Magnetic Resonance Imaging (MRI) scans were performed at the UCSF Neuroscience Imaging Center using a Siemens Prisma fit 3 T scanner. Magnetization prepared rapid gradient-echo (MPRAGE) sequences were used to obtain whole brain T1-weighted images (TR/TE/TI = 2300/2.9/900 ms, $\alpha = 9°$). The field of view was 240 mm × 256 mm, with 1 mm × 1 mm in-plane resolution and 1 mm slice thickness with a sagittal orientation.

Before processing, all T1-weighted images were visually inspected for quality control and those with excessive motion or image artifact were excluded. Magnetic field bias was corrected using the N3 algorithm (31). Tissue segmentation was performed using unified segmentation in SPM12 (32). Each participant's gray matter segmentation was warped to create a study-specific template using Diffeomorphic Anatomical Registration using Exponentiated Lie algebra (DARTEL) (33). Participants' native space gray and white matter segmentations were then normalized and modulated to study-specific template space using nonlinear and rigid-body transformations. Images were smoothed using a Gaussian kernel of 4-mm full width half maximum. Each participant's segmentation was carefully inspected to ensure the robustness of the process.

For statistical purposes, linear and nonlinear transformations between DARTEL's space and ICBM space were applied (34). Quantification of volumes in specific brain regions was accomplished by transforming a standard parcellation atlas into International Consortium for Brain Mapping (ICBM) space and summing all modulated gray matter within each parcellated region of interest (ROI) (35). Total intracranial volume was calculated for each participant as the sum of the gray matter, white matter, and cerebrospinal fluid segmentations. Medial temporal lobe volume was selected as our brain MRI outcome, as exercise engagement in older adults has previously been associated with greater volume in this particular region (36). For the purpose of this study, medial temporal lobe volumes included the following bilateral regions: hippocampus, entorhinal cortex, and the parahippocampal gyrus.

2.7. Statistical analyses

First, we examined associations among the four PA measures (Fitbit steps, Fitbit calories burned, PASE, CHAMPS-MET) with Spearman's rank correlations to evaluate comparability. Next, we evaluated relationships between PA measures with demographic variables of interest *via* independent samples *t*-tests, Spearman's rank correlations, and ANOVA. We tested construct validity by evaluating the relationship between each PA measure and the cardiometabolic and cognitive outcomes of interest *via* Spearman's rank correlations or independent samples *t*-tests (e.g., HTN diagnosis), as appropriate. Lastly, we examined relationships between each PA measure and medial temporal volume outcomes *via* linear regression modeling adjusting for total intracranial volume. We reported effect sizes as Spearman's correlations, Cohen's *d*, or standardized betas, as necessary.

In order to identify the characteristics of older adult participants who were over or under reporting, we performed a discrepancy score analysis. Given our data indicated that Fitbit total steps demonstrated the best construct validity out of all PA measures examined, we utilized Fitbit total steps as our "gold standard" metric in these analyses. First, within participants that completed all three PA metrics ($n = 75$), we computed sample-based z-scores for each PA measure (Fitbit steps, PASE, and CHAMPS-MET). We computed individual discrepancy scores separately by subtracting Fitbit total steps z-scores from each self-report PA measure z-score (e.g., PASE or CHAMPS-MET). Distribution of discrepancy scores approximated normality (**Figure 1**). In this manner, higher discrepancy scores indicated greater overreporting compared to Fitbit. Next, we evaluated relationships between each discrepancy score with demographic (i.e., gender, age, education), cardiometabolic, cognitive, and brain volume outcomes. We tested relationships and group differences *via* independent samples *t*-tests, linear regression, and Spearman's rank correlations. Interpretation of effect sizes was in alignment with Cohen's *Statistical Power Analysis for the Behavioral Sciences*: coefficients of 0.10 "small,".30 "medium," and those of 0.50 "large" in terms of the magnitude (37).

3. Results

On average, participants were 74 years old, 58% female, and took 7,841 average daily steps during the monitoring period. Older age was associated with less physical activity across metrics, but only reached significance for Fitbit outcomes (**Table 2**). Females had lower levels of quantified physical activity for Fitbit daily steps [$t(177) = 2.43$, $p = 0.02$], Fitbit daily calories [$t(177) = 15.77$, $p < 0.001$], and lower reported physical activity for CHAMPS-MET [$t(83) = 2.98$, $p < 0.001$], but not PASE [$t(103) = 0.72$, $p = 0.47$]. Education did not meaningfully associate with the physical activity metrics (**Table 2**).

3.1. Associations among physical activity measures

All of the PA metrics were positively correlated, though demonstrated only small-to-medium effect sizes (**Table 2**). Greater Fitbit step count was more strongly associated with PASE compared to CHAMPS-MET, whereas Fitbit calories burned was more strongly associated with CHAMPS-MET

compared to PASE scores. Self-reported CHAMPS-MET and PASE scores demonstrated a medium effect size correlation with one another.

3.2. Cognitive outcomes

Higher Fitbit step count [$\rho(132) = 0.28$, $p < 0.001$] and calories burned [$\rho(132) = 0.23$, $p = 0.01$], but not PASE [$\rho(78) = -0.10$, $p = 0.39$] were associated with better performances on measures of executive functioning. Unexpectedly, CHAMPS-MET demonstrated an inverse relationship with executive functioning [$\rho(63) = -0.41$, $p < 0.01$]. Only Fitbit measures demonstrated expected positive associations with processing speed, though effect sizes were small (ρ range = 0.05–0.07, ps < 0.43). Fitbit step count [$\rho(124) = 0.20$, $p = 0.03$], but not Fitbit calories burned [$\rho(124) = -0.08$, $p = 0.38$], PASE [$\rho(76) = -0.26$, $p = 0.02$], or CHAMPS-MET [$\rho(62) = -0.31$, $p = 0.02$] was associated with better scores on tests assessing memory.

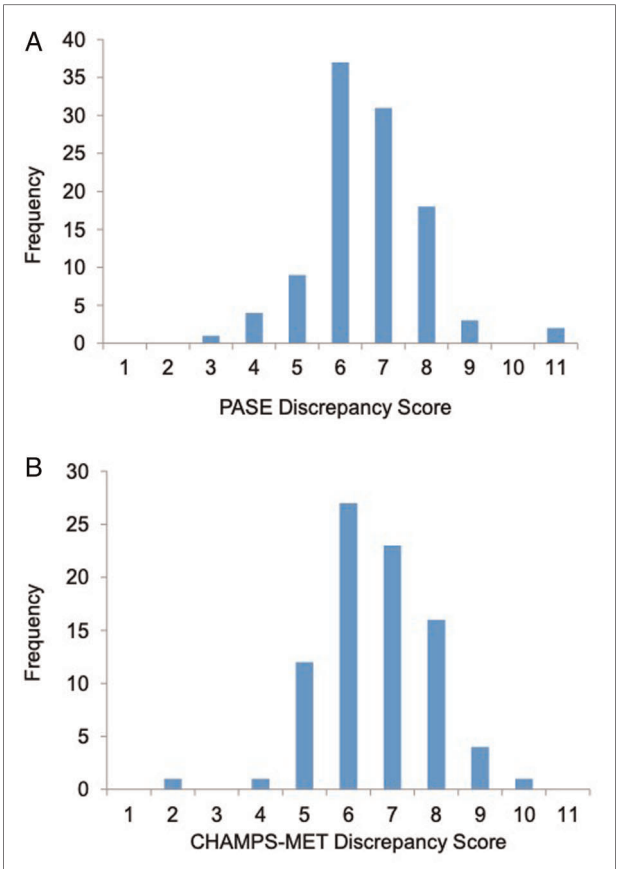


FIGURE 1
Distribution of PASE and CHAMPS discrepancy scores.

3.3. Cardiometabolic outcomes

Lower resting heart rate was significantly associated with greater daily calories burned ($\rho = -0.29$, $p < 0.001$), but less strongly with daily steps ($\rho = -0.14$, $p = 0.08$) and showed minimal associations with reported PA as measured by the PASE ($\rho = 0.02$, $p = 0.85$) or CHAMPS-MET ($\rho = -0.03$, $p = 0.80$) (Figure 2). Greater Fitbit step count ($t = 3.058$, $p < 0.001$), but not Fitbit calories burned ($t = -0.37$, $p = 0.71$), PASE ($t = 1.43$, $p = 0.16$) or CHAMPS-MET ($t = -1.75$, $p = 0.09$) was associated with a lower likelihood of hypertension. Each PA measure demonstrated expected negative associations with hemoglobin A1C that did not reach statistical significance (ρ range = −0.01 to −0.17, ps > 0.05).

3.4. Brain MRI outcomes

Lastly, greater Fitbit step count ($\beta = 0.35$, $p < 0.001$) and calories burned ($\beta = 0.43$, $p < 0.001$), but not PASE ($\beta = 0.15$, $p = 0.17$) or CHAMPS-MET ($\beta = 0.07$, $p = 0.56$), were associated with larger medial temporal lobe volumes (see Figure 3).

Given there were sex and age-related differences, we ran models adjusting for these factors and the patterns remained the same, with the exception of the relationship between Fitbit calories burned and memory, which showed a positive association ($\beta = 0.31$, $p = 0.07$)

TABLE 2 Correlations between physical activity measures, age, and education.

	1. Age	2. Education	3. Fitbit steps	4. Fitbit calories	5. PASE
1. Age					
2. Education					
3. Fitbit steps	−0.36*	0.02			
4. Fitbit calories	−0.38*	0.13	0.50*		
5. PASE	−0.13	−0.11	0.35*	0.20*	
6. CHAMPS-MET	−0.03	0.04	0.20	0.31*	0.44*

Note. *Statistically significant at $p < 0.05$.

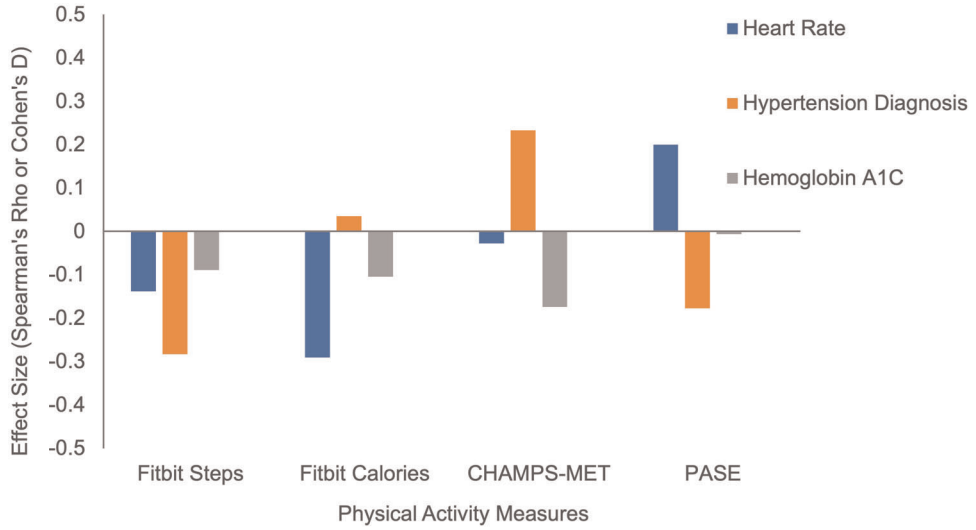


FIGURE 2
Physical activity measures with cardiometabolic outcomes. Note. *Statistically significant at $p < 0.05$.

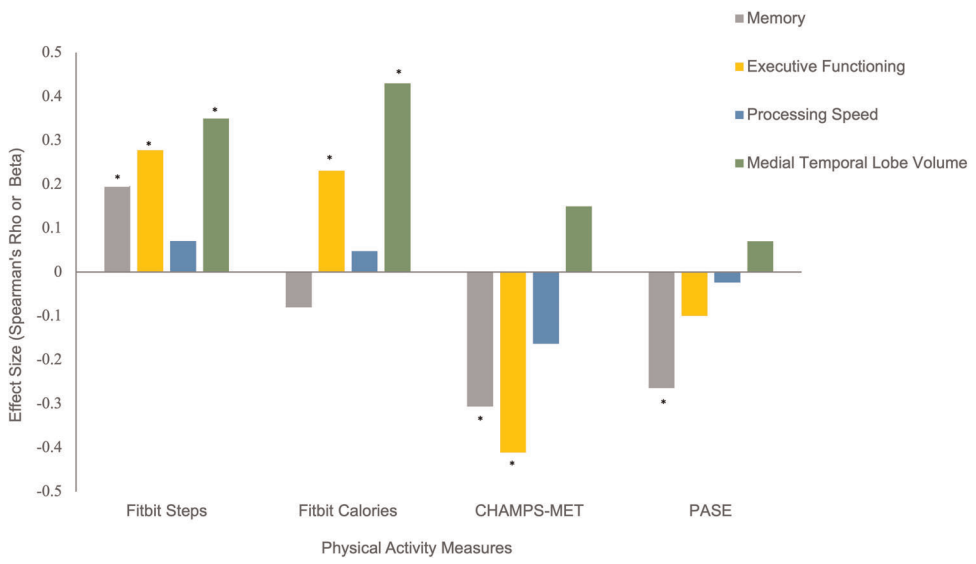


FIGURE 3
Physical activity measures with cognitive and brain MRI outcomes. Note. *Statistically significant at $p < 0.05$.

3.5. Discrepancy scores

Using Fitbit step count as the standard, discrepancy score analysis showed that females tended to overreport to a greater degree on both self-report PA scales, though this effect reached significance only for PASE (PASE Cohen's $d = 0.61$, $p = 0.02$; CHAMPS Cohen's $d = 0.22$, $p = 0.45$), see **Figure 4**. Older age and overreporting showed small associations that did not reach significance (ρ range = 0.09–0.17, $ps > 0.14$). Additionally, those with a diagnosis of hypertension overreported on CHAMPS-MET (PASE Cohen's $d = 0.20$, $p = 0.45$; CHAMPS Cohen's $d = 1.02$, $p < 0.01$), see **Figure 4**. Generally, worse cognition was associated with greater degree of overreporting particularly for measures of memory (ρ range = −0.23 to −0.26, $ps = 0.05$–0.09) and executive functioning (PASE $\rho = -0.15$, $p = 0.28$; CHAMPS $\rho = -0.32$, $p < 0.01$), see **Figure 5**.

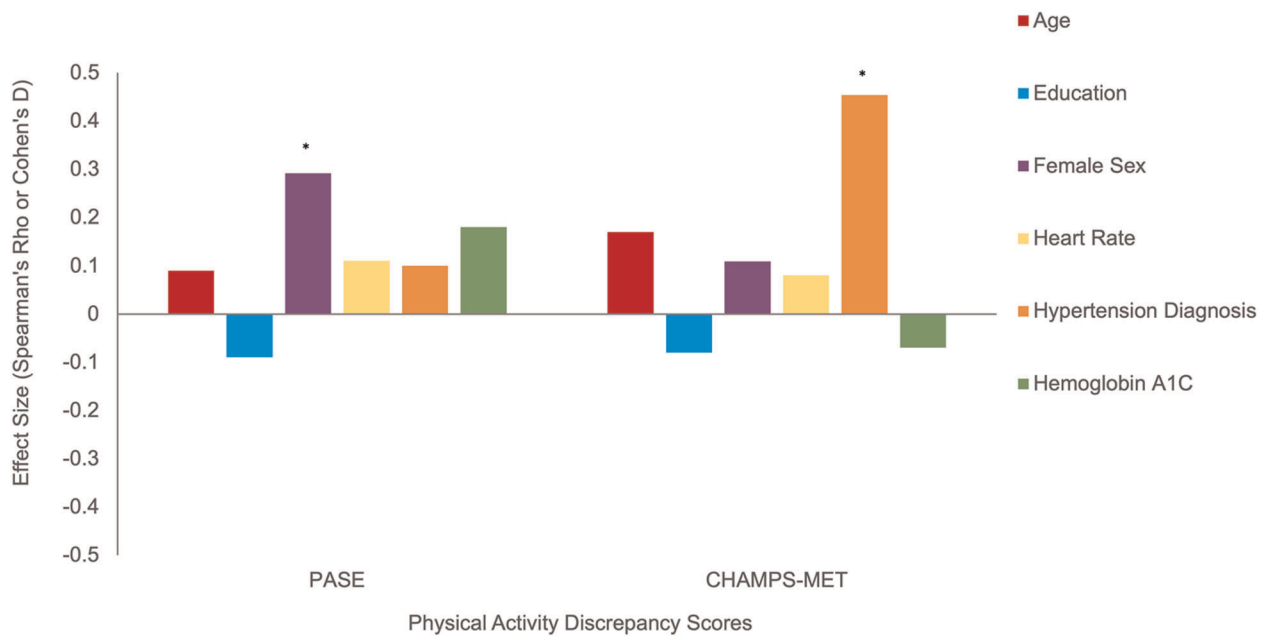


FIGURE 4
Demographic and cardiometabolic correlates of physical activity overreporting positive discrepancy score indicates greater overreporting). Note. *Statistically significant at $p < 0.05$.

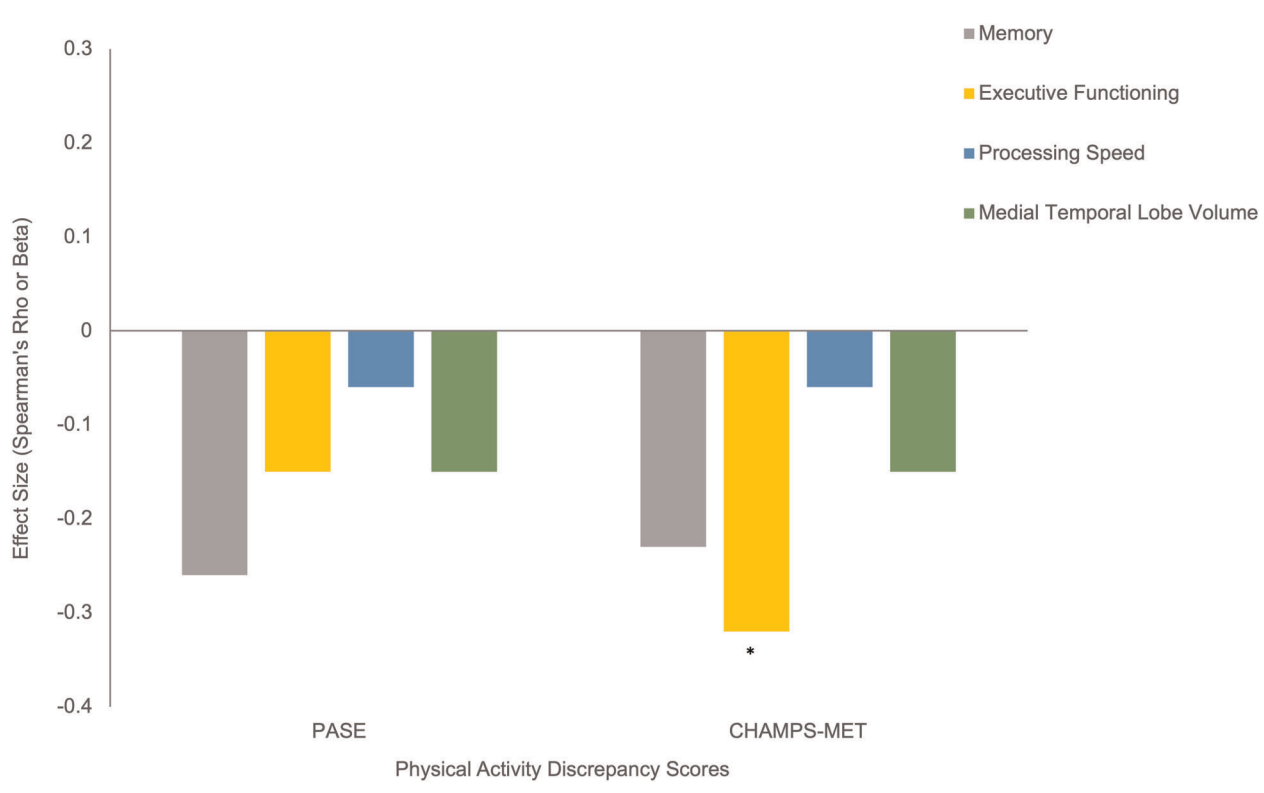


FIGURE 5
Cognitive correlates of physical activity overreporting (positive discrepancy score indicates greater overreporting). Note. *Statistically significant at $p < 0.05$.

Decreased medial temporal lobe volume showed small associations with increased overreporting, though they did not reach statistical significance (β range = −0.14 to −0.15 $ps > 0.19$), see **Figure 5**.

4. Discussion

We build on previous findings that wearable, objective indicators of PA demonstrate greater overall construct validity compared to self-report measures now extending to older adults, and further detail the characteristics that may impact reporting styles on PA measures. We also highlight several novel findings on the limited interchangeability of PA metrics. Both Fitbit steps and calories burned demonstrated more consistent convergence with relevant outcomes of interest when compared to subjective measures. These results are consistent with prior studies in younger adults showing closer approximation of measuring physical activity with objective monitors in comparison to self-report (38, 39). Within Fitbit metrics, step count particularly demonstrated stronger and more consistent expected associations with each cardiometabolic, cognitive, and brain volume outcome of interest compared to calories burned, consistent with previous studies (11). These findings converge with previous literature that has demonstrated distinct associations between physical activity and executive functioning, memory, and MTL volume (36, 40, 41). While we again identify that objective digital health measures show preferable construct validity over subjective measures of physical activity, objective measures still demonstrated substantial variability with expected outcomes of interest (e.g., hypertension diagnosis, memory) and more work is warranted to elucidate factors that contribute to variability in physical activity metrics for older adults.

Notably, the associations found among the four physical activity metrics used in this study suggest a lack of cohesion within the field's current standard measurement approaches. For instance, our results demonstrated only modest relationships across the four physical activity measures (i.e., $r = 0.2–0.30$) indicating small variance explained between each measure (i.e., $R^2 = 4\%–9\%$ variance explained). These data suggest only minimal overlap among metrics that were created to quantify the same construct. There was also variation between Fitbit metrics and their associations with the self-report measures. Fitbit step count demonstrated a greater association with subjective reports of physical activity measured by the PASE, whereas Fitbit calories was more strongly associated with physical activity measured by CHAMPS-MET. These findings illustrate an important implication. While these subjective and objective measures were intended to capture the same construct, there is clearly variability in how each measure assesses physical activity. However, it is worth noting that CHAMPS-MET appears to be more closely related to the construct of calories burned, as expected. It is important to continue to clarify what each measure is assessing (with regard to physical activity) in order to pinpoint what factors of physical activity are beneficial in the context of brain and age-related health.

In addition to variation across the physical activity measures, our results also indicated considerable heterogeneity with associations between each activity measure and outcomes of interest. For example, Fitbit step count showed a stronger relationship to decreased likelihood of hypertension, while Fitbit calories demonstrated a stronger association with lower resting heart rate. Medial temporal lobe volume, and hemoglobin A1C demonstrated consistent, expected associations with each measure of PA; however, effects were small and often did not reach statistical significance (with the exception of Fitbit metrics and brain volume). Overall, there was notable variability in the strength and direction of examined relationships, particularly between self-reported physical activity measures and each outcome. This again suggests that physical activity may be comprised of several constructs that are differentially tapped into by self-report questionnaires and wearable devices, and/or there is imprecision when capturing physical activity across measurement tools. These findings are particularly relevant in scientific research when comparing across studies, as these objective and subjective measures do not appear to be equivalent to one another. Our results revealed that metrics capturing physical activity intensity *via* calories burned (e.g., Fitbit calories, CHAMPS-MET) may not be interchangeable with metrics of overall movement (i.e., Fitbit steps, PASE). Our data highlight how challenging measurement of physical activity can be, and that there is still room for improvement, even within "gold-standard" objective measures.

Fitbit steps demonstrated the greatest construct validity. However, there are still mixed findings in current literature with regard to its validity and reliability (42). For example, one study found that Fitbit total steps underestimated activity in healthy adults walking at faster treadmill speeds, but overestimated total steps at slower speeds (43). Similarly, another study found that Fitbit underestimated caloric expenditure in comparison to CHAMPS (39). The results of the current study, in conjunction with the variability in results from previous studies examining physical activity measures, highlight the importance of improving the current standard measurements of physical activity. In addition, these inconsistencies point to a need for refinement of the operational definitions associated with each of these physical activity measures to best understand what they are evaluating.

Notably, more precise measurement and specification of the broad range of physical activity constructs currently

being utilized would allow for greater understanding of the specific movement patterns that are most critical for brain health. To date, there is not strong evidence for a particular movement (e.g., walking) or intensity (calories burned, heart rate during exercise) that is most impactful for brain health trajectories. For example, in a meta-analysis of exercise randomized control trials (RCTs) in older adults, type of exercise was not a significant predictor of cognitive benefit (44). Indeed, activities ranging from tai chi to jogging demonstrated comparable benefit. Some of the earliest evidence linking physical activity with cognitive aging demonstrated beneficial effects even with low impact activities, such as walking (45). However, several epidemiologic and RCTs indicate that cardiorespiratory activities aimed at increasing VO_2 max may be particularly beneficial for cognitive outcomes and future dementia risk (2, 46, 47). Nonetheless, more work needs to be done in this area to understand what particular aspects of physical activity are most important for brain health.

Because of the degree of variability found within associations between measures of physical activity and important outcomes, we elected to more closely examine which factors could be driving these relationships. More specifically, we leveraged a discrepancy score analysis to identify factors that characterize participants who may overreport on physical activity questionnaires. The results demonstrated a greater degree of overreporting in females, particularly for the PASE. This finding may impact the ability of studies to determine sex-related differences in physical activity for brain and age-related health. In addition, participants who were most discrepant generally had higher levels of physical activity across all measures, which may suggest that self-report measures are less accurate in detecting physical activity levels for those who are very active.

We also found that individuals who engaged in the greatest degree of overreporting were older adults who performed worse on cognitive assessments, suggesting subjective measures may be systematically confounded by cognitive ability when assessing activity level in the aging population. This finding converges with prior studies demonstrating that utilization of self-report measures in older adults for measuring physical activity is less accurate (48). In addition, cognitive functioning declines with age as a group, which in turn may increase risk of inaccurate responses on self-report measures (49). In addition to decline in cognition, a greater degree of overreporting was also associated with smaller medial temporal lobe volume. Generally, our findings demonstrated that individuals with poorer vascular and cognitive health tended to overreport to a greater degree. These novel findings increase our understanding of possible factors that could be contributing to reporting bias in older adults, namely their vascular and cognitive health. Future studies should examine whether these discrepancy scores could be used to predict individuals at risk for adverse brain aging.

Our study is not without limitations. Simply wearing a Fitbit may have increased participants' motivation to move, which could lead to possible a possible confound in our data. However, we noted that all feedback from the Fitbit was removed from the device so participants were otherwise blinded to real time activity levels. Our sample was not representative of the general population, with 85 percent of participants identifying as White, limiting generalizability. Furthermore, our sample was relatively small and had some variability across measures, which may have biased the outcomes. There are also inherent biases that can occur when utilizing digital technology with older adults; notably, data collection is contingent on successful device use, which may be impacted by technology familiarity. Therefore, study findings have limited generalizability to older adults with low levels of technology literacy. In addition, the high educational attainment of this cohort also impacts our generalizability and is particularly important considering higher education has been associated with lower risk of developing neurodegenerative disease. It is possible that our results are limited by a slight gap in time between obtaining self-report data and gathering metrics from Fitbit monitors. At the visit, participants reported physical activity from the last seven days *via* PASE, and from the prior 30 days *via* CHAMPS. Fitbit monitoring took place in the following 30 days post visit. However, physical activity levels are a generally stable trait (50), so it is unlikely this time discrepancy significantly affected our findings.

Our study is also limited by the lack of other "gold-standard" objective measures of physical activity and fitness, such as the 6-min walk or a research-grade accelerometer (e.g., Actigraph GT3X+). With these other objective measures, we could have more comprehensively examined Fitbit as an actigraphy metric. In addition, a longitudinal study design would have provided the opportunity to understand the reliability of subjective and objective PA measures over time and meaningfully track changes in cognitive and vascular health related to physical activity.

Without precision and specificity, it is difficult to pinpoint which aspects of physical activity contribute to brain health. Our results suggest that objective quantification of physical activity demonstrates the best validity and high clinical relevance. Moreover, issues regarding reporting bias may be especially important in older adults with lower vascular, cognitive and brain structural statuses. These findings also begin to broaden our understanding of what physical activity metrics represent to facilitate better-informed recommendations for healthy aging.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by UCSF Human Research Protection Program (HRPP). The patients/participants provided their written informed consent to participate in this study.

Author contributions

AV and KC contributed to the conception and design of the study and performed the statistical analysis. ND and MY organized the database. AV wrote the manuscript. LG, EG, and CF wrote sections of the manuscript. JK provided critical edits and study funding. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the NIH-NIA (K23 AG058752, R01AG032289, and R01AG048234), and the UCSF ADRC P30AG062422. Our work was also supported by the Alzheimer's Association (AARG-20-683875, PI: KC) and the Larry L. Hillblom Foundation Network (2014-A-004-NET).

Acknowledgments

The authors would like to thank the National Institute on Aging, the Alzheimer's Association, and the Larry L. Hillblom Foundation Network for funding this study. We would also like to thank the participants and family members involved.

Conflict of interest

Joel H. Kramer receives royalties from Pearson's Inc. The authors report no other conflict of interest.

Publisher's note

References

1. Chen F-T, Hopman RJ, Huang C-J, Chu C-H, Hillman CH, Hung T-M, et al. The effect of exercise training on brain structure and function in older adults: a systematic review based on evidence from randomized control trials. *J Clin Med*. (2020) 9(4):914. doi: 10.3390/JCM9040914

2. Hörder H, Johansson L, Guo X, Grimby G, Kern S, Östling S, et al. Midlife cardiovascular fitness and dementia: a 44-year longitudinal population study in women. *Neurology*. (2018) 90(15):e1298–305. doi: 10.1212/WNL.0000000000005290

3. Nystoriak M, Bhatnagar A. Cardiovascular effects and benefits of exercise. *Front Cardiovasc Med*. (2018) 5:135. doi: 10.3389/FCVM.2018.00135

4. Buchman AS, Boyle PA, Yu L, Shah RC, Wilson RS, Bennett DA. Total daily physical activity and the risk of AD and cognitive decline in older adults. *Neurology*. (2012) 78(17):1323–9. doi: 10.1212/WNL.0B013E3182535D35

5. Casaletto KB, Staffaroni AM, Wolf A, Appleby B, Brushaber D, Coppola G, et al. Active lifestyles moderate clinical outcomes in autosomal dominant frontotemporal degeneration. *Alzheimer's Dement*. (2020) 16(1):91–105. doi: 10.1002/alz.12001

6. Noah JA, Spierer DK, Gu J, Bronner S. Comparison of steps and energy expenditure assessment in adults of fitbit tracker and ultra to the actical and indirect calorimetry. *J Med Eng Technol*. (2013) 37(7):456–62. doi: 10.3109/03091902.2013.831135

7. Dontje ML, de Groot M, Lengton RR, van der Schans CP, Krijnen WP. Measuring steps with the fitbit activity tracker: an inter-device reliability study. *J Med Eng Technol*. (2015) 39(5):286–90. doi: 10.3109/03091902.2015.1050125

8. Diaz KM, Krupka DJ, Chang MJ, Peacock J, Ma Y, Goldsmith J, et al. Fitbit®: an accurate and reliable device for wireless physical activity tracking. *Int J Cardiol*. (2015) 185:138–40. doi: 10.1016/J.IJCARD.2015.03.038

9. Jones D, Crossley K, Dascombe B, Hart HF, Kemp J. Validity and reliability of the fitbit FLEX™ and actigraph gt3x+ at jogging and running speeds. *Int J Sports Phys Ther*. (2018) 13(5):860. doi: 10.26603/ijspt20180860

10. Reid RER, Insogna JA, Carver TE, Comptour AM, Bewski NA, Sciortino C, et al. Validity and reliability of Fitbit activity monitors compared to ActiGraph GT3X+ with female adults in a free-living environment. *J Sci Med Sport*. (2017) 20(6):578–82. doi: 10.1016/J.JSAMS.2016.10.015

11. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int J Behav Nutr Phys Act*. (2015) 12(1):1–22. doi: 10.1186/S12966-015-0314-1

12. Hagiwara A, Ito N, Sawai K, Kazuma K. Validity and reliability of the physical activity scale for the elderly (PASE) in Japanese elderly people. *Geriatr Gerontol Int*. (2008) 8(3):143–51. doi: 10.1111/J.1447-0594.2008.00463.X

13. Washburn RA, Ficker JL. Physical activity scale for the elderly (PASE): the relationship with activity measured by a portable accelerometer. *J Sports Med Phys Fitness*. (1999) 39(4):336–40. PMID: 10726435

14. Chad KE, Reeder BA, Harrison EL, Ashworth NL, Sheppard SM, Schultz SL, et al. Profile of physical activity levels in community-dwelling older adults. *Med Sci Sports Exerc*. (2005) 37(10):1774–84. doi: 10.1249/01.MSS.0000181303.51937.9C

15. Logan SL, Gottlieb BH, Maitland SB, Meegan D, Spriet LL. The physical activity scale for the elderly (PASE) questionnaire; does it predict physical

health? *Int J Environ Res Public Heal*. (2013) 10(9):3967–86. doi: 10.3390/IJERPH10093967

16. Washburn RA, Smith KW, Jette AM, Janney CA. The physical activity scale for the elderly (PASE): development and evaluation. *J Clin Epidemiol*. (1993) 46 (2):153–62. doi: 10.1016/0895-4356(93)90053-4

17. Stewart L, Mills KM, King AC, Haskell WL, Gillis D, et al. CHAMPS physical activity questionnaire for older adults: outcomes for interventions. *Med Sci Sports Exerc*. (2001) 33(7):1126–41. doi: 10.1097/00005768-200107000-00010

18. Hekler EB, Buman MP, Haskell WL, Conway TL, Cain KL, Sallis JF, et al. Reliability and validity of CHAMPS self-reported sedentary-to-vigorous intensity physical activity in older adults. *J Phys Act Heal*. (2012) 9(2):225–36. doi: 10.1123/JPAH.9.2.225

19. Feldman LS, Kaneva P, Demyttenaere S, Carli F, Fried GM, Mayo NE. Validation of a physical activity questionnaire (CHAMPS) as an indicator of postoperative recovery after laparoscopic cholecystectomy. *Surgery*. (2009) 146 (1):31–9. doi: 10.1016/J.SURG.2009.02.019

20. Prince SA, Cardilli L, Reed JL, Saunders TJ, Kite C, Douillette K, et al. A comparison of self-reported and device measured sedentary behaviour in adults: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act*. (2020) 17 (1):1–17. doi: 10.1186/S12966-020-00938-3

21. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med*. (2003) 37(3):197–206. doi: 10.1136/BJSM.37.3.197

22. Sims J, Smith F, Duffy A, Hilton S. The vagaries of self-reports of physical activity: a problem revisited and addressed in a study of exercise promotion in the over 65s in general practice. *Fam Pract*. (1999) 16(2):152–7. doi: 10.1093/FAMPRA/16.2.152

23. Mandolesi L, Polverino A, Montuori S, Foti F, Ferraioli G, Sorrentino P, et al. Effects of physical exercise on cognitive functioning and wellbeing: biological and psychological benefits. *Front Psychol*. (2018) 9:509. doi: 10.3389/FPSYG.2018.00509

24. Rikli RE, Edwards DJ. Effects of a three-year exercise program on motor function and cognitive processing speed in older women. *Res Q Exerc Sport*. (2013) 62(1):61–7. doi: 10.1080/02701367.1991.10607519

25. Kramer JH, Jurik J, Sha SJ, Rankin KP, Rosen HJ, Johnson JK, et al. Distinctive neuropsychological patterns in frontotemporal dementia, semantic dementia, and Alzheimer disease. *Cogn Behav Neurol*. (2003) 16(4):211–8. doi: 10.1097/00146965-200312000-00002

26. Possin KL, Laluz VR, Alcantar OZ, Miller BL, Kramer JH. Distinct neuroanatomical substrates and cognitive mechanisms of figure copy performance in Alzheimer's disease and behavioral variant frontotemporal dementia. *Neuropsychologia*. (2011) 49(1):43–8. doi: 10.1016/J.NEUROPSYCHOLOGIA.2010.10.026

27. Kerchner GA, Racine CA, Hale S, Wilheim R, Laluz V, Miller BL, et al. Cognitive processing speed in older adults: relationship with white matter integrity. *PLoS One*. (2012) 7(11). doi: 10.1371/journal.pone.0050425

28. Kramer JH, Mungas D, Possin KL, Rankin KP, Boxer AL, Rosen HJ, et al. NIH EXAMINER: conceptualization and development of an executive function battery. *J Int Neuropsychol Soc*. (2014) 20(1):11. doi: 10.1017/S1355617713001094

29. American Heart Association. All about heart rate (pulse). Available at: https://www.heart.org/en/health-topics/high-blood-pressure/the-facts-about-high-blood-pressure/all-about-heart-rate-pulse (Accessed October 25, 2021).

30. Fox K, Borer JS, Camm AJ, Danchin N, Ferrari R, Lopez Sendon JL, et al. Resting heart rate in cardiovascular disease. *J Am Coll Cardiol*. (2007) 50 (9):823–30. doi: 10.1016/J.JACC.2007.04.079

31. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in mri data. *IEEE Trans Med Imaging*. (1998) 17(1):87–97. doi: 10.1109/42.668698

32. Ashburner J, Friston KJ. Unified segmentation. *Neuroimage*. (2005) 26 (3):839–51. doi: 10.1016/J.NEUROIMAGE.2005.02.018

33. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage*. (2007) 38(1):95–113. doi: 10.1016/J.NEUROIMAGE.2007.07.007

34. Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. A probabilistic atlas of the human brain: theory and rationale for its development: the international consortium for brain mapping (ICBM). *Neuroimage*. (1995) 2(2):89–101. doi: 10.1006/NIMG.1995.1012

35. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. (2006) 31(3):968–80. doi: 10.1016/J.NEUROIMAGE.2006.01.021

36. Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, et al. Exercise training increases size of hippocampus and improves memory. *Proc Natl Acad Sci U S A*. (2011) 108(7):3017–22. doi: 10.1073/PNAS.1015950108

37. Cohen J. *Statistical power analysis for the behavioral sciences*. New York City: Routledge (1988).

38. Colbert LH, Matthews CE, Havighurst TC, Kim K, Schoeller DA. Comparative validity of physical activity measures in older adults. *Med Sci Sports Exerc*. (2011) 43(5):867–76. doi: 10.1249/MSS.0B013E3181FC7162

39. Stahl ST, Insana SP. Caloric expenditure assessment among older adults: criterion validity of a novel accelerometry device. *J Health Psychol*. (2014) 19 (11):1382–7. doi: 10.1177/1359105313490771

40. Snowden M, Steinman L, Mochan K, Grodstein F, Prohaska TR, Thurman DJ, et al. Effect of exercise on cognitive performance in community-dwelling older adults: review of intervention trials and recommendations for public health practice and research. *J Am Geriatr Soc*. (2011) 59(4):704–16. doi: 10.1111/J.1532-5415.2011.03323.X

41. Erickson KI, Raji CA, Lopez OL, Becker JT, Rosano C, Newman AB, et al. Physical activity predicts gray matter volume in late adulthood: the Cardiovascular Health Study. *Neurology*. (2010) 75(16):1415–22. doi: 10.1212/WNL.0B013E3181F88359

42. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Young Yoo J, et al. Accuracy of fitbit devices: systematic review and narrative syntheses of quantitative data. *JMIR mHealth uHealth*. (2018) 6(8):e10527. doi: 10.2196/10527

43. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. *J Sci Med Sport*. (2014) 17(5):496–500. doi: 10.1016/J.JSAMS.2013.10.241

44. Northey JM, Cherbuin N, Pumpa KL, Smee DJ, Rattray B. Exercise interventions for cognitive function in adults older than 50: a systematic review with meta-analysis. *Br J Sports Med*. (2018) 52(3):154–60. doi: 10.1136/BJSPORTS-2016-096587

45. Yaffe K, Barnes D, Nevitt M, Lui LY, Covinsky K. A prospective study of physical activity and cognitive decline in elderly women: women who walk. *Arch Intern Med*. (2001) 161(14):1703–8. doi: 10.1001/ARCHINTE.161.14.1703

46. Pentikainen H, Savonen K, Ngandu T, Solomon A, Komulainen P, Paajanen T, et al. Cardiorespiratory fitness and cognition: longitudinal associations in the FINGER study. *J Alzheimers Dis*. (2019) 68(3):961–8. doi: 10.3233/JAD-180897

47. Vidoni ED, Johnson DK, Morris JK, Van Sciver A, Greer CS, Billinger SA, et al. Dose-response of aerobic exercise on cognition: a community-based, pilot randomized controlled trial. *PLoS One*. (2015) 10(7):e0131647. doi: 10.1371/JOURNAL.PONE.0131647

48. Harada ND, Chiu V, King AC, Stewart AL. An evaluation of three self-report physical activity instruments for older adults. *Med Sci Sports Exerc*. (2001) 33 (6):962–70. doi: 10.1097/00005768-200106000-00016

49. Baranowski T. Validity and reliability of self report measures of physical activity: an information-processing perspective. *Res Q Exerc Sport*. (2013) 59 (4):314–27. doi: 10.1080/02701367.1988.10609379

50. Friedman HS, Martin LR, Tucker JS, Criqui MH, Kern ML, Reynolds CA. Stability of physical activity across the lifespan. *J Health Psychol*. (2008) 13 (8):1092–104. doi: 10.1177/1359105308095963

frontiers | Frontiers in Digital Health

TYPE Original Research
PUBLISHED 04 November 2022
DOI 10.3389/fdgth.2022.999981

OPEN ACCESS

EDITED BY
OraLee Branch,
Lark Health (United States), United States

REVIEWED BY
Sarah Anne Graham,
Lark Health, United States
Rochelle Henderson,
Washington University in St. Louis, United States

*CORRESPONDENCE
Lisa Gualtieri
lisa.gualtieri@tufts.edu

SPECIALTY SECTION
This article was submitted to Human Factors and Digital Health, a section of the journal Frontiers in Digital Health

RECEIVED 21 July 2022
ACCEPTED 11 October 2022
PUBLISHED 04 November 2022

CITATION
Gualtieri L, Shaveet E, Estime B and Patel A (2022) The role of home medication storage location in increasing medication adherence for middle-aged and older adults.
Front. Digit. Health 4:999981.
doi: 10.3389/fdgth.2022.999981

The role of home medication storage location in increasing medication adherence for middle-aged and older adults

Lisa Gualtieri*, Eden Shaveet, Brandon Estime and Avi Patel

Department of Public Health and Community Medicine, Tufts University School of Medicine, Boston, MA, United States

Background: Over 50% of US adults do not take their prescriptions as prescribed, which is responsible for 33%–69% of hospital admissions and 125,000 deaths annually. Given the higher prevalence of prescription drug use among middle-aged and older adult populations, promoting medication adherence is of particular importance with these age groups. Two speculated facilitators of medication adherence are home medication storage location and the use of digital health devices.

Objective: Our objective was to use survey data to investigate the associations between medication storage location and medication adherence among adults 40 years and older. Additionally, we aimed to report preliminary findings about the associations between use of devices and medication adherence in this same population.

Methods: We conducted primary analysis of data sampled from a home medication management survey deployed in November 2021 ($n = 580$). We conducted exploratory analyses by way of chi^2 tests and creation of bivariate logistic regression models.

Results: The most commonly used storage locations by our sample were nightstand drawers (27%), kitchen cabinets (25%), and atop bedroom nightstands (23%). Several medication storage locations were significantly associated with decreased odds of having ever forgotten to take a medication, including kitchen drawers, in refrigerators, atop bedroom nightstands, in nightstand drawers, and backpacks, purses, or bags. Two home medication storage locations were significantly associated with increased odds of having ever forgotten to take a medication: kitchen cabinets and bathroom vanities. Further, most (94%) survey respondents indicated they would be receptive to guidance about where to store their medications.

Conclusions: Given that some home medication storage locations are associated with adherence, an intervention to guide storage location selection may support increased adherence, especially with high receptivity expressed for such guidance. Increased adherence may also accrue from device usage paired with optimized home medication storage location. We plan to investigate that further, as well as how new device designs can incorporate contextual cues related to location to promote medication adherence more effectively in middle aged and older adults.

KEYWORDS

medication adherence, older adults, digital health, medication management, medication storage, adherence devices

Introduction

Definition and importance

Medication adherence, the extent to which a patient follows a stipulatory medication treatment plan, is crucial to the success of patient care and is indispensable in reaching clinical goals (1). However, only 50% of people adhere to medication guidelines (2), posing substantial risk to patient health and safety. Medication nonadherence is responsible for as many as 33%–69% of hospital admissions and 125,000 deaths annually (3). Other consequences of medication nonadherence include waste of medication, disease progression, and a lower quality of life overall (4). There is a rising trend in medication nonadherence across all age, sex, and racial groups in the United States (5), which is particularly concerning for middle-aged and older adults as the likelihood to be prescribed long term medications increases with age (6, 7).

Medication nonadherence, while well-studied, continues to be a vexing issue for clinicians and researchers (2, 8). In 2003, the World Health Organization (WHO) indicated that improved medication adherence interventions may have greater impact on population health than specific medical treatments (9). Known barriers to medication adherence include out-of-pocket costs of medications, physical difficulties obtaining a medication supply, and difficulty following intake guidance (10, 11). Patients facing multiple, co-occurring barriers are even less likely to adhere to prescribed medications (12).

Home medication management

Given the importance of contextual cues and mental associations in developing routines (13), medication storage locations may be a vital component of improving medication adherence across populations. In the process of habit formation, settings associated with certain behaviors may influence an individual's actions to deviate from conscious motivation (14). External contextual cues in a person's environment trigger an automatic response that promotes habit formation (15). Previous work exploring strategies for medication adherence found that reliance on multiple cues may have a beneficial impact on adherence to medication regimens (16). Locations where medications are frequently stored may have the capacity for a plethora of contextual cues.

Understanding a patient's home medication management routine, including home medication storage location selection, may be needed for more effective medication management and strategic placement of devices that serve to dispense or remind to take medications. For example, devices that rely on auditory or visual cues need to be in locations where patients can hear or see the alerts and notifications. While home medication management is well-studied in relation to patient safety (17), less is known about the role of home medication storage location in medication adherence.

Patients and locations of interest

Behavior change theories such as the Rubicon model of action phases suggest that the level of intent in improving medication adherence should be considered when developing medication adherence interventions (18). According to this model, interventions are most effective when an individual intends to be adherent to their medication but only needs additional support and guidance to do so (19). As a result of exploring this theory, our primary interest is in patients who do not face well-documented barriers to medication adherence, such as cost and access (11), and intend to be adherent, yet still struggle to achieve adherence (19).

Our population of interest is middle-aged and older adults as this demographic is more likely to take one or more prescription medications compared to other age groups (20). For some health conditions, such as hypertension and diabetes mellitus, a patient's older age may also impact their level of adherence (21).

Our location of interest is in patients' homes as more medications are taken at home than in hospitals and clinics combined (1, 22). Our research focuses on understanding the barriers to medication adherence in the home to design aids that increase adherent behavior. We explore this by first understanding where patients store their medications, and then by investigating how patients select these locations. We then investigate whether the locations themselves or the determinants of selection correlate with adherence. Finally, we examine factors related to home medication management including the use of digital aids to help with adherence.

Home medication storage locations

Prior research into home medication storage has examined location with respect to safety, climatic conditions, and routines (23–25); however, the impact of home storage locations on medication adherence is understudied in literature. A study of medication storage conditions found that, of 170 participants aged 65 and older, 76% complied with drug product label recommendations for temperature, light, and humidity (26). A study on medication disposal found that 81.5% of respondents in 445 telephone interviews had prescription medications and almost all the respondents indicated that there were excess and leftover medications in their homes (23). Another study on safe medication storage surveyed 1,074 people aged 50–80 with grandchildren aged 0–17 (27). Their findings indicated that 89% of respondents had prescription medications in their homes and 84% reported that they kept their medications in

the same place they typically store them when their grandchildren visited (27).

MedlinePlus, an online information service produced by the United States National Library of Medicine, suggests storing medications in the dresser drawer, kitchen cabinet, storage box, shelf, or closet (28). Although they do not reference medication storage location relative to medication adherence, they state that location plays a role in medication effectiveness and safety. With 418 million users in 2021 (29), MedlinePlus, and services like it, could influence medication storage decisions.

A few studies have addressed the role of strategies, routines, and habit formation in establishing a successful medication regimen (16, 30). One study found that older adults employ both internal (e.g., use of mental associations) and external (e.g., use of physical objects and/or locations) strategies to remember to take their medications (16). Another study considered the use of contextual cues as an aid to develop a new behavior, reporting that patients who store their medications in locations that are conducive to their routines were more likely to be adherent to their medication regimens (13). Behavioral interventions that provided counseling on adherence strategies considered the storage of medications only as a cue to remind individuals to maintain adherence (31, 32), yet their effect was modest. Survey respondents who used visible cues throughout the day were unable to remember whether they had taken their medication (13).

Several efforts have been dedicated to the development and use of devices to improve medication adherence, including designing digital health devices and apps to dispense medication and/or remind users that it is time to take medication (13). However, simple, low-cost devices have yet to produce clinically impactful outcomes (33). There is an increasing number of devices aimed at improving adherence, yet sub-optimal placement of these devices in the home may lead to less efficacious changes in adherence.

No study has evaluated the relationship between medication storage locations and adherence nor the relationship between storage location and device use. Given these gaps in the literature, we designed and deployed a survey to learn more about medication management in the home. Namely, we aimed to assess whether there exists a bivariate relationship between home medication storage location and self-reported medication adherence. Additionally, we aimed to assess the bivariate relationships between use of digital devices for medication intake reminders and self-reported medication adherence. Our first hypothesis is that middle-aged and older adults store their medications in multiple locations while our second hypothesis is there is an association between home medication storage locations and self-reported medication adherence.

Methods

Study sample

We collected data *via* deployment of the Home Medication Management Survey, designed by members of the Digital Health Research Group at Tufts University, and fielded between November 18 and December 14, 2021. We deployed the survey in English *via* Google Forms, an online, cloud-native survey-development platform that encrypts files in transit and at rest (34). We recruited participants *via* informational posts on social media platforms, including Twitter, Facebook, Instagram, and LinkedIn, and *via* the Osher Lifelong Learning Institute at Tufts University electronic mailing list, which reaches an audience of around 2,000 older adults, primarily in the Greater Boston area. Eligible participants for the survey were 18 years of age or older with access to an internet-enabled device. Upon completion of the survey, participants could elect to enter a drawing for one of five Amazon gift cards valued at $25. All study protocols were reviewed and approved by the Tufts University Health Sciences Institutional Review Board in Boston, MA.

Procedures

We recorded a total of 1,966 survey responses at the close of the survey and deemed 1,673 (85%) of responses to be valid after excluding responses on suspect of fraudulence. Dropped responses met any of the following exclusion criteria: responses that were not in English (49 responses dropped); consecutive sets of identical responses posted at the same time (132 responses dropped); responses containing an abnormal email address for the drawing that included a long string of numbers we suspected to be fraudulently generated (19 responses dropped); or responses that included suspicious identical open-text responses within the same day (83 cases dropped).

Of the 1,673 remaining eligible responses, our study sample for this paper was middle-aged and older adults meaning respondents indicating an age of 40 years or older ($n = 580$). We selected this age range to adhere with the centers for disease control and prevention's inclusion of 40–79 years in their report of prescription drug use (35). However, we did not exclude respondents 80 years of age and older for sample size considerations.

Measures

We designed survey questions to learn about respondents' experiences with medication management in the home.

Questions assessed use of aids to adherence, perceived importance of adherence, self-reported adherence, as well as demographic items.

Demographic variables

Demographic items included in our analysis were age in years by decade (40–49, 50–59, 60–69, 70–79, 80–89, 90+), race and ethnicity (multi-select categorical item), sex, and highest level of education completed. **Table 1** displays the demographic distribution of our sample.

Digital medication reminders

We derived our measures indicating use of digital and non-digital medication reminder methods from two survey items from which respondents could select all that applied: non-digital methods ("Written notes," "Post-it notes," "Calendar," "Chart," "Pillbox") and digital methods ("Smartphone app," "Smartphone alarm," "Siri," "Alexa," "Smartwatch," "Electronic pill dispenser," "GlowCap or attachment to pill bottle").

Perceived importance of medication adherence

We measured perceived importance of medication adherence on a five-point Likert scale ranging from "not at all important" to "very important" in response to a survey item asking, "How important is it for you to take your medication as prescribed?".

Medication adherent behavior

We derived our variables assessing medication adherence from questions which used memory of intaking a medication as a proxy for adherence (36). Our variables were dichotomous ("yes", "no") in response to survey items asking, "Have you ever forgotten to take a medication?" and "In the past two weeks, have you forgotten to take a medication?".

Medication storage location

We derived our variable for storage locations of medications taken regularly from a survey item ("Where in your home do you store your prescriptions that you take on a regular basis?") to which respondents could select all that applied: "Kitchen table," "Kitchen cabinet," "Kitchen counter," "Kitchen drawer," "In the refrigerator," "On the bathroom vanity," "In the vanity drawer or cabinet," "Bathroom medicine cabinet," "On top of the bedroom nightstand," "In the nightstand drawer," "Desk," "Dining room table," "Backpack, purse, or bag," "Closet," and an open-text selection for unlisted locations. We categorized open-text responses indicating use of already existing values as those values accordingly *via* consensus coding to promote interrater reliability. **Table 2** displays the storage selection distribution for our sample.

Receptivity to storage location guidance

We measured receptivity to storage guidance using a survey item ("If you received a new prescription, would you be open to receiving guidance on where to store the medication?") to which respondents could select all that applied ("Yes, from my physician," "Yes, from my pharmacist," "Yes, on an app or website," "Yes, in a brochure," "No"). We consolidated all "yes" responses into a binary variable indicating guidance receptivity to one or more listed sources (physician, pharmacist, mobile app or website, or brochure). We cleaned "No" responses by excluding any response for whom a "Yes" response was also given for the item.

Analysis

We assessed relative frequencies for all sample characteristics and variables listed in our *Measures* subsection. We conducted bivariate analyses for digital medication reminder method variables and variables indicating adherent behavior to medications taken regularly, receptivity to medication storage guidance, and medication storage locations currently in use by respondents. Analyses included chi^2 tests of homogeneity of proportions and bivariate logistic regression models.

Results

Sample characteristics

Respondents in our middle-aged and older adult sample were 59% female and 40% male. Less than 1% of respondents identified with a non-listed sex. Respondents ranged from 40 to over 90 years old with most (64%) between the ages of 40 and 59 years. Most respondents self-identified as white (76%). Of white respondents ($n = 439$), the vast majority (95%) selected no other race or ethnicity. Just over 12% of the sample identified as American Indian or Alaskan Native, 7% identified as Black or African American, 5% identified as Asian, 5% identified as Hispanic or Latino, 2% identified as Native Hawaiian or Pacific Islander, and less than 1% identified with an unlisted race or ethnicity. Respondents' education levels varied, though most respondents (60%) identified as having obtained a bachelor's, master's, or professional degree (**Table 1**).

Digital medication reminders

Over half of respondents (56%) indicated that they use digital methods to remember to take their medication. The most common digital methods reported included use of

TABLE 1 Sample characteristics by self-reported adherence ($n = 580$).

Characteristic	Overall N (%)	Ever forgot to take a medication (row %)				Recently forgot to take a medication[a] (row %)			
		Yes	No	chi^2*	p	Yes	No	chi^2*	p
Overall	580 (100)	72	28			35	65		
Age by decade									
40–49	244 (42)	59	41	38.76	<0.001	39	61	8.36	0.079
50–59	127 (22)	75	25			35	65		
60–69	82 (14)	85	15			37	63		
70–79	84 (14)	86	14			23	77		
≥ 80[b]	43 (7)	81	19			28	72		
Race/Ethnicity[c]									
White	439 (76)	75	25	10.20	0.001	32	68	5.37	0.02
American Indian or Alaskan Native	70 (12)	49	51	20.65	<0.001	36	64	0.05	0.817
Black or African American	41 (7)	63	37	1.44	0.231	49	51	3.99	0.046
Asian	27 (5)	85	15	2.59	0.108	59	41	7.69	0.006
Hispanic or Latino	27 (5)	74	26	0.09	0.766	44	56	1.24	0.265
Native Hawaiian or Other Pacific Islander	12 (2)	75	25	0.07	0.789	67	33	5.62	0.018
Sex[d]									
Female	345 (59)	77	23	12.87	0.002	34	66	3.96	0.138
Male	233 (40)	64	36			35	65		
Highest Education Completed									
<High school diploma or equivalency	17 (3)	82	18	34.72	<0.001	59	41	12.30	0.056
High school diploma or equivalency	45 (8)	53	47			38	62		
Some college	82 (14)	51	49			32	68		
Associate's degree	47 (8)	68	32			49	51		
Bachelor's degree	160 (28)	74	26			29	71		
Master's or professional degree[b]	183 (32)	80	20			33	67		
Doctoral degree	46 (8)	83	17			39	61		
Digital Reminder Methods Used[c]									
Any digital method[e]	325 (56)	60	40	48.46	<0.001	36	64	0.48	0.489
Smartphone app or alarm[b]	224 (39)	63	37	14.69	<0.001	37	63	1.07	0.301
Apple Siri	86 (15)	45	55	34.06	<0.001	35	65	0.01	0.932
Amazon Alexa	42 (7)	60	40	3.22	0.073	50	50	4.83	0.028
Smartwatch	86 (15)	59	41	7.44	0.006	40	60	1.14	0.285
Electronic pill dispenser	73 (13)	58	42	8.06	0.005	45	55	4.25	0.039
GlowCap or attachment to pill bottle	22 (4)	55	45	3.25	0.071	41	59	0.42	0.518
No digital methods	255 (44)	86	14	48.46	<0.001	33	67	0.48	0.489
Perceived importance of medication adherence									
Not at all or not very important[b]	11 (2)	64	36	28.45	0.274	36	64	35.40	<0.001
Neutral	52 (9)	83	17			62	38		
Important	197 (34)	72	28			43	57		
Very important	320 (55)	70	30			25	75		
Receptivity to storage guidance[c]									
Receptive to guidance from at least one listed source[e]	545 (94)	71	29	0.14	0.712	35	65	0.15	0.695
From physician	317 (55)	73	27	0.35	0.556	34	66	0.003	0.957
From pharmacist	383 (66)	73	27	1.34	0.247	31	69	4.96	0.026
From mobile application or website	111 (19)	79	21	4.03	0.045	32	68	0.53	0.467
From printed material (brochures)	130 (22)	88	12	23.54	<0.001	34	66	0.03	0.862

(*continued*)

TABLE 1 Continued

Characteristic	Overall N (%)	Ever forgot to take a medication (row %)				Recently forgot to take a medication[a] (row %)			
		Yes	No	*chi*^{2}*	*p*	Yes	No	*chi*^{2}*	*p*
Not receptive to guidance from any listed source[f]	35 (6)	74	26	0.14	0.712	31	69	0.15	0.695

[a]Defined as forgetting to take a medication in the last two weeks.
[b]Values combined due to small sample size or similarity.
[c]Non-mutually exclusive; single respondent may fall into several categories, except for composite values.
[d] < 1% identified with an unlisted sex.
[e]Composite variable or value.
[f]Derived from absence of value selection.
*Pearson chi^{2}; *P*-value denotes likelihood of observed chi^{2} assuming homogeneity of proportions. For non-mutually-exclusive items, chi^{2} is calculated for each selection as a binary value by the binary adherence proxy.

TABLE 2 Home medication storage location by self-reported adherence (*n* = 580).

Home medication storage location	Overall N (%)	Ever forgot to take a medication (%)				Recently forgot to take a medication** (%)			
		Yes	No	*chi*^{2}*	*P*	Yes	No	*chi*^{2}*	*P*
Overall	580 (100)	72	28			35	65		
Kitchen table	58 (10)	67	33	0.59	0.443	48	52	5.43	0.02
Kitchen cabinet	146 (25)	78	22	4.09	0.043	44	56	7.55	0.006
Kitchen counter	90 (16)	69	31	0.37	0.542	38	62	0.51	0.474
Kitchen drawer	80 (14)	59	41	7.47	0.006	40	60	1.25	0.26
In refrigerator	82 (14)	61	39	5.25	0.022	38	62	0.47	0.495
On bathroom vanity	80 (14)	90	10	15.52	<0.001	36	64	0.13	0.72
In vanity drawer or cabinet	81 (14)	67	33	1.10	0.293	31	69	0.55	0.46
Bathroom medicine cabinet	112 (19)	71	29	0.07	0.791	33	67	0.13	0.72
Atop bedroom nightstand	135 (23)	61	39	10.1	0.001	38	62	0.85	0.358
In the nightstand drawer	158 (27)	56	44	24.72	<0.001	25	75	8.08	0.004
Desk	80 (14)	50	50	21.18	<0.001	33	67	0.16	0.688
Dining room table	46 (8)	61	39	2.8	0.094	37	63	0.14	0.713
Backpack, purse, or bag	64 (11)	58	42	6.67	0.01	42	58	1.89	0.169
Closet	24 (4)	79	21	0.71	0.398	29	71	0.31	0.576
Unlisted location (alone)***	14 (2)	100	0	5.7	0.017	36	64	0.01	0.922

*Pearson chi^{2}; *P*-value denotes likelihood of observed chi^{2} assuming homogeneity of proportions. For all values, chi^{2} is calculated for each selection as a binary value (only affirmative displayed) by the binary adherence proxy.
**Defined as forgetting to take a medication in the last two weeks.
***Inclusive of uncategorizable responses for which a categorizable response was not also provided by the respondent.

smartphone applications or alarms (39%). The least common digital method reported was the use of pill bottle attachments, such as GlowCaps (4%).

Perceived importance of medication adherence and adherent behavior

Most respondents (89%) indicated that they felt taking medication as prescribed was "important" or "very important," and 65% had not forgotten to take a medication in the two weeks prior to survey. Over 71%, however, indicated that they had forgotten to take a medication at least once in their lives. Of those who had ever forgotten to take a medication (n = 415), almost half (48%) had forgotten to take a medication in the two weeks prior to responding to the survey.

Home medication storage locations

The most popular home medication storage locations for medications taken regularly by our sample included

nightstand drawers (27%), kitchen cabinets (25%), and atop bedroom nightstands (23%). Despite the name, only 19% of our sample stored medication in medicine cabinets. Other locations included kitchen counters (16%), in refrigerators (14%), in vanity drawers or cabinets (14%), inside of bathroom vanities (14%), in kitchen drawers (14%), inside or on top of desks (14%), in backpacks, purses, or bags (11%), on kitchen tables (10%), on dining room tables (8%), and in closets (4%). Less than 3% of respondents stored medications principally in unlisted locations (Table 2).

Receptivity to storage guidance

Most respondents (94%) indicated that they would be receptive to guidance about where to store their medications, with most preferring guidance from pharmacists (66%) followed by guidance from physicians (55%). Fewer (19%) indicated that they would be receptive to guidance delivered on a digital platform, such as websites or mobile applications.

Bivariate analyses

Medication storage locations and medication adherence

Several home medication storage locations were significantly associated with decreased odds of having ever forgotten to take a medication, including kitchen drawers (OR: 0.51, 95% CI: 0.31–0.83), in refrigerators (OR: 0.57, 95% CI: 0.35–0.93), atop bedroom nightstands (OR: 0.52, 95% CI: 0.35–0.78), in nightstand drawers (OR: 0.38, 95% CI: 0.26–0.56), desks (OR: 0.33, 95% CI: 0.21–0.54), and backpacks, purses, or bags (OR: 0.5, 95% CI: 0.29–0.85). Two home medication storage locations were significantly associated with increased odds of having ever forgotten to take a medication: kitchen cabinets (OR: 1.57, 95% CI: 1–2.45) and bathroom vanities (OR: 4.12, 95% CI: 1.94–8.76). All remaining home medication storage locations were not associated with having ever forgotten to take a medication.

Far fewer home medication storage locations were significantly associated with having forgotten to take a medication in the two weeks prior to survey. Only one location was significantly associated with decreased odds of forgetting to take a medication in the two weeks prior to survey: nightstand drawers (OR: 0.56, 95% CI: 0.37–0.84). Two locations were significantly associated with increased odds of forgetting to take a medication in the two weeks prior to survey: kitchen tables (OR: 1.9, 95% CI: 1.09–3.28) and kitchen cabinets (OR: 1.71, 95% CI: 1.16–2.51) (Table 3).

TABLE 3 Bivariate associations between home medication storage location and self-reported medication adherence.

Home Medication Storage location	Ever forgot to take a medication	Recently forgot to take a medication**
	OR (95% CI)	OR (95% CI)
Kitchen table	0.8 (0.45–1.42)	1.9 (1.1–3.28)*
Kitchen cabinet	1.57 (1.01–2.45)*	1.71 (1.16–2.51)*
Kitchen counter	0.86 (0.53–1.4)	1.19 (0.74–1.89)
Kitchen drawer	0.51 (0.31–0.83)*	1.32 (0.81–2.14)
In refrigerator	0.57 (0.35–0.93)*	1.18 (0.73–1.92)
On bathroom vanity	4.12 (1.94–8.76)*	1.09 (0.67–1.79)
In vanity drawer or cabinet	0.76 (0.46–1.26)	0.83 (0.5–1.37)
Bathroom medicine cabinet	0.94 (0.6–1.48)	0.92 (0.6–1.43)
Atop bedroom nightstand	0.52 (0.35–0.78)*	1.21 (0.81–1.8)
In the nightstand drawer	0.38 (0.26–0.56)*	0.56 (0.37- 0.84)*
Desk	0.33 (0.21- 0.54)*	0.9 (0.55–1.49)
Dining room table	0.59 (0.32–1.1)	1.12 (0.6–2.1)
Backpack, purse, or bag	0.5 (0.29–0.85)*	1.45 (0.85–2.45)
Closet	1.53 (0.56–4.18)	0.77 (0.32–1.9)

*$P < 0.05$.
**Defined as forgetting to take a medication in the last two weeks.

Use of digital reminder methods and medication adherence

Bivariate analyses revealed statistically significant associations between use of any digital medication reminder methods (composite variable) and having ever forgotten to take a medication ($p < 0.001$, chi^2). Affirmative indication of having used any digital method was significantly associated with decreased odds of having ever forgotten to take a medication (OR: 0.24, 95% CI: 0.16–0.36). Significant associations were not observed between use of digital methods (composite variable) and forgetting to take a medication in the two weeks prior to survey.

Digital medication reminder methods significantly associated with decreased odds of having ever forgotten to take a medication included use of smartphone applications and alarms (OR: 0.49, 95% CI: 0.34–0.71), Apple Siri (OR: 0.26, 95% CI: 0.16–0.42), smartwatches (unadjusted OR: 0.52, 95% CI: 0.32–0.84), and electronic pill dispensers (OR: 0.49, 95% CI: 0.29–0.81). Use of other listed digital methods (Amazon Alexa and electronic attachments to pill bottles) were not significantly associated decreased odds of with having ever forgotten to take a medication.

Interestingly, while no digital reminder methods were significantly associated with decreased odds of having forgotten to take a medication in the two weeks prior to survey, some digital methods were associated with increased odds, including Amazon Alexa (OR: 2, 95% CI: 1.07–3.77) and electronic pill dispensers (OR: 1.68, 95% CI: 1.02–2.76) (Table 4).

TABLE 4 Bivariate associations between Use of digital reminders and self-reported medication adherence.

Digital reminder method	Ever forgot to take a medication	Recently forgot to take a medication**
	OR (95% CI)	OR (95% CI)
Any digital reminders***	0.24 (0.16–0.36)*	1.9 (1.1–3.28)
Smartphone app or alarm	0.49 (0.34–0.71)*	1.2 (0.85–1.71)
Apple Siri	0.26 (0.16–0.42)*	1.02 (0.63–1.65)
Amazon Alexa	0.56 (0.29–1.06)	2 (1.07–3.77)*
Smartwatch	0.52 (0.32–0.84)*	1.29 (0.81–2.07)
Electronic pill dispenser	0.49 (0.29–0.81)*	1.68 (1.02–2.76)*
GlowCap or attachment to pill bottle	0.46 (0.2–1.09)	1.33 (0.56–3.17)
No digital methods	4.19 (2.75–6.38)*	0.89 (0.63–1.25)

*$P < 0.05$.
**Defined as forgetting to take a medication in the last two weeks.
***Composite variable.

Discussion

Importance of and self-reported adherence

Most respondents (89%) indicated that they felt taking medication as prescribed was "important" or "very important," indicating that they may be more likely to be receptive to medication adherence interventions. As previously stated, adherence in this study was defined as having ever or recently forgotten to take a routine medication (36). Most respondents (65%) had remembered to take all medications in the two weeks prior to survey. Over 70%, however, indicated that they had forgotten to take a medication at least once in their lives. It is crucial that we explore how these findings relate to other studies that have explored both medication adherence and nonadherence (37).

Home medication storage locations

We found that the most commonly used storage locations for medications taken regularly by our sample were nightstand drawers (27%), kitchen cabinets (25%), and atop bedroom nightstands (23%). These exploratory results offer support to our first hypothesis that patients utilize a variety of home storage locations to store their medications. In future work, we will learn more about the reasons and motivations underlying these choices. We suspect that some reasons may be related to home conditions unexplored in this research. For example, bedside nightstands, by virtue of where they are placed in the home, may be ideal for private medication storage over locations in communal areas. Climactic requirements or safety considerations unexplored in this research may also be factors.

The University of Michigan National Poll on Healthy Aging is the only previous study that considered storage location, although their study interest was grandchild safety (27). In total, 1,074 adults aged 50–80 with a grandchild aged 0–17 complete the University of Michigan National Poll, while 580 adults aged 40 and older completed our study analyzed here. In the two studies, researchers gave respondents different lists of location options to choose from and the naming was also different for some of the locations (e.g., "kitchen cabinet" compared to "cupboard or cabinet"). The locations reported from the Tufts University School of Medicine and the University of Michigan National Poll on Healthy Aging surveys differ greatly, which we attribute to the options available to respondents. Neither survey asked about the characteristics of the locations nor the determinants of location selection.

Our exploratory analyses suggest significant bivariate associations between several home medication storage locations and decreased odds of having ever forgotten to take a medication, including atop nightstands and inside nightstand drawers. Other locations, including kitchen cabinets, were associated with increased odds of having ever forgotten to take a medication. Fewer locations were associated with decreased odds of having forgotten to take a medication in the two weeks prior to survey. These findings offer support to our second hypothesis speculating that some storage locations are associated with greater adherence and some locations with worse adherence.

Since these unadjusted, exploratory analyses suggest that some home medication storage locations may be associated with adherence, the next step in our research is to learn what factors influence location selection, understand the implications for adherent behavior, and apply more sophisticated multivariable approaches to our analysis.

Most survey respondents (94%) indicated that they would be receptive to guidance regarding where to store their medications, with more preferring guidance from a pharmacist (66%) followed by guidance from physicians (55%). In subsequent research, we hope to learn either if the characteristics of the locations or the determinants of location selection are associated with adherence. Knowing respondents' receptivity to guidance and knowing more about the selection of locations associated with adherence, in our next study we plan to design interventions to guide storage selection.

Medication reminders

Several devices and apps offer medication intake reminders; however, in prior research, these devices have not been effective

(33, 37, 38). There may be many reasons for this including the lack of incorporation of location as well as the focus on time-based notifications.

No known medication reminder or dispensing devices offer recommendations for optimal device storage locations, yet, since many devices use auditory and/or visual cues, we suspect that location may be a consideration for cues to be heard or seen. Further, no known devices are designed for specific locations in the home. Over half of respondents (56%) indicated that they use digital methods to remember to take their medication, the most common being the use of smartphone applications or alarms (39%). The sheer number of respondents using reminders is indicative of a perceived need, yet the lack of effectiveness in studies is concerning.

Our analysis found no significant associations between use of digital methods and forgetting to take a medication in the two weeks prior to survey. However, digital medication reminder methods that were significantly associated with decreased odds of having ever forgotten to take a medication included use of smartphone applications and alarms, Apple Siri, smartwatches, and electronic pill dispensers. Interestingly, while no digital reminder methods were significantly associated with decreased odds of having forgotten to take a medication in the two weeks prior to survey, some digital methods were associated with increased odds, including Amazon Alexa and electronic pill dispensers.

Future directions

In subsequent analyses, we will examine other survey responses to better understand other aspects of home medication management. One question asked about the frequency with which respondents check for unused or expired medications, which may be a proxy for conscientious medication management; a reduction of clutter in a home storage location may increase adherence with fewer prescription bottles to choose from. Another question asked about the impact of the Coronavirus (COVID-19) pandemic to learn if more time at home eased medication management and thus increased adherence; alternatively, other changes, such as decreased time with pharmacists, may have reduced adherence.

Finally, in planned qualitative studies we hope to better understand how the selection of medication storage locations relate to daily routines and cues to engage in those routines (13). Through these interviews, we hope to understand the relationship between medication adherence and factors we have not previously explored, two being how adherence changes with travel and the role of patient activation, which comprises the degree of knowledge, confidence, and skills that patients have to manage their overall health (39).

Limitations

As our survey recruitment was primarily achieved *via* an electronic mailing list compiled from Osher Lifelong Learning Institute at Tufts, our sample is not representative of the US population. This resulted in our sample appearing to be more educated and more digitally literate than the US census (40). Few studies explore the relationship directly between digital literacy and medication adherence however increased digital literacy has been correlated with increased health literacy which has been correlated to medication adherence.

As our analyses utilized self-reported adherence as measured *via* a digital survey rather than an actual measurement of adherence directly *via* observation or prescription records, response bias may have affected our results. Self-reported medication adherence has been shown to overestimate adherence behavior compared with other assessment methods and generally have high specificity but low sensitivity (41).

Additionally, our exploratory analyses were bivariate and thus did not adjust for other additional variables.

Conclusions

Our long-term goal is to aid middle-aged and older adults in making informed choices about where to store their medications in their homes to increase adherence. Strong evidence to guide optimization of storage locations could play a crucial role in improving adherence and, thereby, the health and safety of middle-aged and older adults living independently. Our future research will investigate the relationship between medication storage locations and adherence by more fully understanding the characteristics of storage locations, the determinants of location selection, and their role in routines. With an understanding of which factors relate to storage location and impact adherence, we hope to develop best practice guidelines that can be used by pharmacists, by physician, and in innovative digital health solutions to counsel patients on optimal selection of home medication storage locations to improve their medication adherence. Our future research will also investigate how new device designs can incorporate contextual cues related to location to promote adherence more effectively. With a rising number of middle-aged and older adults and a commensurate increase in the number of patients taking prescription medications, interventions to increase adherence will lead to greater health and longevity.

Data availability statement

Raw data that support the findings of this study are available from the corresponding author, upon reasonable request.

Ethics statement

The studies involving human participants were reviewed and approved by Tufts University Health Sciences Institutional Review Board in Boston, MA. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

The authors are responsible for developing and deploying the survey and all analysis of data reported here. LG initiated this research project, was responsible for developing and deploying the survey, and contributed to writing this paper. ES performed the primary analysis of the survey and contributed to writing this paper. BE and AP contributed to writing this paper. All authors contributed to the article and approved the submitted version.

Funding

This research was funded in part by Tufts University through the Springboard Program, and in part by the National Center for Advancing Translational Sciences, National Institutes of Health, Award Number UL1TR002544. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Acknowledgments

The authors offer their appreciation to the graduate students at Tufts University School of Medicine who contributed to this work, particularly, Justin Barton, Cheryl Croll, Meera Singhal, and Ricardo Boschetti.

Conflict 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.

Publisher's note

References

1. Sabaté E. *Adherence to long-term therapies: Evidence for action*. Geneva, Switzerland: World Health Organization (2003).

2. Bouwman L, Eeltink CM, Visser O, Janssen JJ, Maaskant JM. Prevalence and associated factors of medication non-adherence in hematological-oncological patients in their home situation. *BMC cancer*. (2017) 17(1):1–8. doi: 10.1186/s12885-017-3735-1

3. Benjamin RM. Medication adherence: helping patients take their medicines as directed. *Public Health Rep*. (2012) 127(1):2–3. doi: 10.1177/003335491212700102

4. Jimmy B, Jose J. Patient medication adherence: measures in daily practice. *Oman Med J*. (2011) 26(3):155. doi: 10.5001/omj.2011.38

5. Desai R, Fong HK, Shah K, Kaur VP, Savani S, Gangani K, et al. Rising Trends in hospitalizations for cardiovascular events among young cannabis users (18-39 Years) without other substance abuse. *Medicina*. (2019) 55(8):438. doi: 10.3390/medicina55080438

6. Charlesworth CJ, Smit E, Lee DS, Alramadhan F, Odden MC. Polypharmacy among adults aged 65 years and older in the United States: 1988–2010. *J Gerontol - Biol Sci Med Sci*. (2015) 70(8):989–95. doi: 10.1093/gerona/glv013

7. Ihara E. Prescription drugs georgetown university's institute for health care research and policy Available at: https://hpi.georgetown.edu/rxdrugs/.

8. Stirratt MJ, Curtis JR, Danila MI, Hansen R, Miller MJ, Gakumo CA. Advancing the science and practice of medication adherence. *J Gen Intern Med*. (2018) 33(2):216–22. doi: 10.1007/s11606-017-4198-4

9. Brown MT, Bussell JK. eds. *Medication adherence: WHO cares? Mayo clinic proceedings*. Rochester, MN: Elsevier (2011).

10. Chan AHY, Cooper V, Lycett H, Horne R. Practical barriers to medication adherence: what do current self-or observer-reported instruments assess? *Front Pharmacol*. (2020) 11:572. doi: 10.3389/fphar.2020.00572

11. Gellad WF, Grenard JL, Marcum ZA. A systematic review of barriers to medication adherence in the elderly: looking beyond cost and regimen complexity. *Am J Geriatr Pharmacother*. (2011) 9(1):11–23. doi: 10.1016/j.amjopharm.2011.02.004

12. Newman-Casey PA, Robin AL, Blachley T, Farris K, Heisler M, Resnicow K, et al. The most common barriers to glaucoma medication adherence: a cross-sectional survey. *Ophthalmol*. (2015) 122(7):1308–16. doi: 10.1016/j.ophtha.2015.03.026

13. Stawarz K, Gardner B, Cox A, Blandford A. What influences the selection of contextual cues when starting a new routine behaviour? An exploratory study. *BMC Psychol*. (2020) 8(1):1–11. doi: 10.1186/s40359-020-0394-9

14. Gardner B, Lally P. Modelling habit formation and its determinants. *Psychol Habit*. (2018):207–29. doi: 10.1007/978-3-319-97529-0_12

15. Vortac O, Edwards MB, Manning CA. Functions of external cues in prospective memory. *Memory*. (1995) 3(2):201–19. doi: 10.1080/09658219508258966

16. Boron JB, Rogers WA, Fisk AD. Everyday memory strategies for medication adherence. *Geriatr Nurs (Minneap)*. (2013) 34(5):395–401. doi: 10.1016/j.gerinurse.2013.05.010

17. Marck PB, Lang A, Macdonald M, Griffin M, Easty A, Corsini-Munt S. Safety in home care: a research protocol for studying medication management. *Implement Sci*. (2010) 5(1):1–9. doi: 10.1186/1748-5908-5-43

18. Keller L, Gollwitzer PM, Sheeran P. Changing behavior using the model of action phases. *The Handbook of Behavior Change*. (2020) 2:77–88. doi: 10.1017/9781108677318.006

19. Pakpour AH, Gellert P, Asefzadeh S, Updegraff JA, Molloy GJ, Sniehotta FF. Intention and planning predicting medication adherence following coronary artery bypass graft surgery. *J Psychosom Res*. (2014) 77(4):287–95. doi: 10.1016/j.jpsychores.2014.07.001

20. Rathore S, Mehta S, Boyko Jr W, Schulman K. Prescription medication use in older Americans: a national report card on prescribing. *Fam Med*. (1998) 30 (10):733–9.

21. Park DC, Morrell RW, Frieske D, Kincaid D. Medication adherence behaviors in older adults: effects of external cognitive supports. *Psychol Aging*. (1992) 7(2):252. doi: 10.1037/0882-7974.7.2.252

22. Aspden P. *Preventing medication errors*. Washington, DC: National Academies Press (2007).

23. Wieczorkiewicz SM, Kassamali Z, Danziger LH. Behind closed doors: medication storage and disposal in the home. *Ann Pharmacother*. (2013) 47 (4):482–9. doi: 10.1345/aph.1R706

24. Funk OG, Yung R, Arrighi S, Lee S. Medication storage appropriateness in US households. *Inov Pharm*. (2021) 12(2). doi: 10.24926/iip.v12i2.3822

25. Sanders MJ, Van Oss T. Using daily routines to promote medication adherence in older adults. *Am J Occup Ther*. (2013) 67(1):91–9. doi: 10.5014/ajot.2013.005033

26. Vlieland ND, van Den Bemt BJ, Bekker CL, Bouvy ML, Egberts TC, Gardarsdottir H. Older patients' compliance with drug storage recommendations. *Drugs Aging*. (2018) 35(3):233–41. doi: 10.1007/s40266-018-0524-8

27. Singer D, Solway E, Kirch M, Kullgren J, Malani P. *Safely storing medication around grandchildren*. Ann Arbor, MI: University of Michigan National Poll on Healthy Aging (2019).

28. Storing your medicines: A. D. A. M. (2020) Available at: https://medlineplus.gov/ency/article/003247.htm.

29. Learn about MedlinePlus: MedlinePlus. (2022) Available at: https://medlineplus.gov/about/general/aboutmedlineplus/.

30. Badawy SM, Shah R, Beg U, Heneghan MB. Habit strength, medication adherence, and habit-based mobile health interventions across chronic medical conditions: systematic review. *J Med Internet Res*. (2020) 22(4):e17883. doi: 10.2196/17883

31. Smaje A, Weston-Clark M, Raj R, Orlu M, Davis D, Rawle M. Factors associated with medication adherence in older patients: a systematic review. *Aging Med*. (2018) 1(3):254–66. doi: 10.1002/agm2.12045

32. Safren SA, Otto MW, Worth JL, Salomon E, Johnson W, Mayer K, et al. Two strategies to increase adherence to HIV antiretroviral medication: life-steps and medication monitoring. *Behav Res Ther*. (2001) 39(10):1151–62. doi: 10.1016/S0005-7967(00)00091-7

33. Choudhry NK, Krumme AA, Ercole PM, Girdish C, Tong AY, Khan NF, et al. Effect of reminder devices on medication adherence: the REMIND randomized clinical trial. *JAMA Intern Med*. (2017) 177(5):624–31. doi: 10.1001/jamainternmed.2016.9627

34. Google Forms: Google. Available at: https://workspace.google.com/products/forms/.

35. Hales CM, Servais J, Martin CB, Kohen D. Prescription drug use among adults aged 40–79 in the United States and Canada. (2019).

36. Tan JP, Cheng KKF, Siah RCJ. A systematic review and meta-analysis on the effectiveness of education on medication adherence for patients with hypertension, hyperlipidaemia and diabetes. *J Adv Nurs*. (2019) 75(11):2478–94. doi: 10.1111/jan.14025

37. Kini V, Ho PM. Interventions to improve medication adherence: a review. *Jama*. (2018) 320(23):2461–73. doi: 10.1001/jama.2018.19271

38. Checchi KD, Huybrechts KF, Avorn J, Kesselheim AS. Electronic medication packaging devices and medication adherence: a systematic review. *Jama*. (2014) 312(12):1237–47. doi: 10.1001/jama.2014.10059

39. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the patient activation measure (PAM): conceptualizing and measuring activation in patients and consumers. *Health Serv Res*. (2004) 39(4p1):1005–26. doi: 10.1111/j.1475-6773.2004.00269.x

40. A Description of U.S. Adults Who Are Not Digitally Literate [press release]. National Center for Education Statistics. (2018).

41. Stirratt MJ, Dunbar-Jacob J, Crane HM, Simoni JM, Czajkowski S, Hilliard ME, et al. Self-report measures of medication adherence behavior: recommendations on optimal use. *Transl Behav Med*. (2015) 5(4):470–82. doi: 10.1007/s13142-015-0315-2

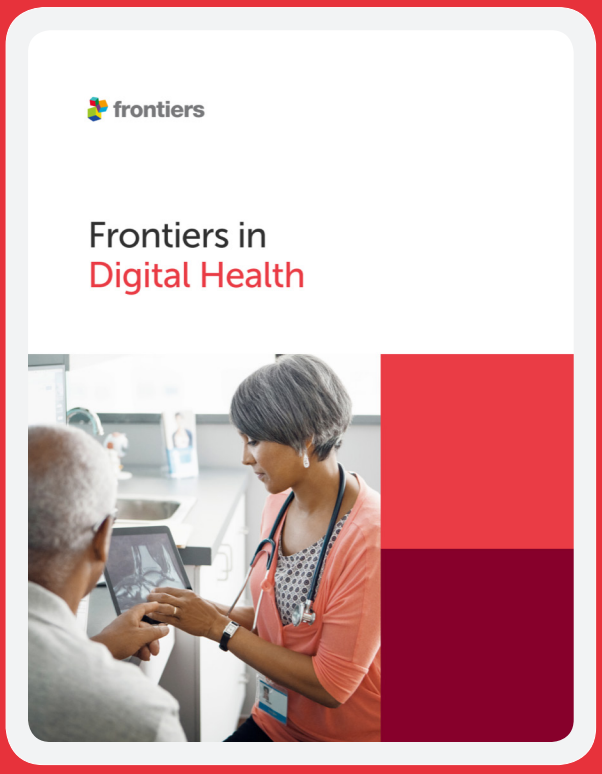
Frontiers in
Digital Health

Explores digital innovation to transform modern healthcare

A multidisciplinary journal that focuses on how we can transform healthcare with innovative digital tools. It provides a forum for an era of health service marked by increased prediction and prevention.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact