

# Smartphone App Designed to Collect Health Information in Older Adults: Usability Study

Joanne M Murabito, Jamie M Faro, Yuankai Zhang, Angelo DeMalia, Alexander Hamel, Nakesha Agyapong, Hongshan Liu, Eric Schramm, David D McManus, Belinda Borrelli

Submitted to: JMIR Human Factors  
on: January 22, 2024

**Disclaimer:** © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 36

    Figures ..... 37

        Figure 1..... 38

        Figure 2..... 39

        Figure 3..... 40

    Multimedia Appendixes ..... 41

        Multimedia Appendix 0..... 42

        Multimedia Appendix 1..... 42

# Smartphone App Designed to Collect Health Information in Older Adults: Usability Study

Joanne M Murabito<sup>1,2</sup> MD, ScM; Jamie M Faro<sup>3</sup> PhD; Yuankai Zhang<sup>4</sup> MA; Angelo DeMalia<sup>1</sup> BA; Alexander Hamel<sup>3</sup> BA; Nakesha Agyapong<sup>3</sup> BA; Hongshan Liu<sup>1</sup> MA; Eric Schramm<sup>5</sup> PharmD; David D McManus<sup>3,6</sup> MD, ScM; Belinda Borrelli<sup>7</sup> PhD

<sup>1</sup>Framingham Heart Study Boston University Framingham US

<sup>2</sup>Section of General Internal Medicine Department of Medicine Boston University Chobanian & Avedisian School of Medicine Boston US

<sup>3</sup>Department of Medicine University of Massachusetts Chan Medical School Worcester US

<sup>4</sup>Department of Biostatistics Boston University School of Public Health Boston US

<sup>5</sup>CareEvolution, Inc Ann Arbor US

<sup>6</sup>Cardiology Division, Department of Medicine University of Massachusetts Chan Medical School Worcester US

<sup>7</sup>Center for Behavioral Science Research Henry M. Goldman School of Dental Medicine, Boston University Boston US

## Corresponding Author:

Joanne M Murabito MD, ScM

Section of General Internal Medicine

Department of Medicine

Boston University Chobanian & Avedisian School of Medicine

72 E Concord St Suite L-516

Boston

US

## Abstract

**Background:** Few studies evaluate the usability of mobile-phone assessments in older adults.

**Objective:** The aim of this study was to identify design-based barriers and facilitators to mobile app survey completion among two samples of older adults; those in the Framingham Heart Study and a more diverse sample from a hospital-based setting.

**Methods:** We used mixed methods to identify challenging and beneficial features of the mobile app in participants from the electronic Framingham Heart Study (eFHS, n=15, mean age 72 years, 40% women, 100% non-Hispanic and White) and among participants recruited from a hospital-based setting (n=15, mean age 71 years, 47% women, 20% Hispanic and 53% non-White). A variety of app-based measures with different response formats were tested including self-reported surveys, pictorial assessments (to indicate body pain sites), and cognitive testing tasks (e.g., Trail Making Test (TMT) and Stroop). Participants completed each measure using a think-aloud protocol, while being audio- and video-recorded with a qualitative interview conducted at the end of the session. Recordings were coded for participant usability errors by two pairs of coders. Participants completed the Mobile App Rating Scale (MARS) to assess the app (response range 1=inadequate to 5= excellent).

**Results:** In eFHS participants, the average total MARS score was 7.6 (SD=1.1), with no significant differences in the hospital-based sample. In general participants were pleased with the app and found it easy to use. A large minority had at least one navigational issue, most committed only once. Most older adults did not have difficulty completing the self-reported multiple choice measures unless it included lengthy instructions but participants had usability issues with the Stroop and TMT.

**Conclusions:** Our methods and results help guide app development and app-based survey construction for older adults, while also giving consideration to sociodemographic differences.

(JMIR Preprints 22/01/2024:56653)

DOI: <https://doi.org/10.2196/preprints.56653>

## Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.  
Only make the preprint title and abstract visible.

✓ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>



## Original Manuscript

## Smartphone App Designed to Collect Health Information in Older Adults: Usability Study

Joanne M. Murabito, MD ScM<sup>1, 2</sup>; Jamie M. Faro, PhD<sup>3</sup>; Yuankai Zhang, MA<sup>4</sup>; Angelo De Malia, BA<sup>2</sup>; Alexander Hamel, BA<sup>3</sup>; Nakesha Agyapong, BA<sup>3</sup>; Hongshan Liu, MA<sup>2</sup>; Eric Schramm, PharmD<sup>5</sup>; David D. McManus MD, ScM<sup>3,6</sup>; Belinda Borrelli PhD<sup>7</sup>

1. Section of General Internal Medicine, Department of Medicine, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, USA

2. Boston University's and National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA

3. Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, USA

4. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

5. CareEvolution Inc, Ann Arbor, Michigan, USA

6. Cardiology Division, Department of Medicine, University of Massachusetts Chan Medical School, Worcester, MA, USA

7. Center for Behavioral Science Research, Boston University, Henry M. Goldman School of Dental Medicine, Boston, MA, USA

**Author email addresses:** murabito@bu.edu; Jamie.Faro@umassmed.edu; yukiz@bu.edu; ademalia@bu.edu; Alexander.Hamel@umassmed.edu;

Nakesha.Agyapong@umassmed.edu; hongslu@bu.edu; schramm@careevolution.com; David.McManus@umassmed.edu; belindab@bu.edu

**Corresponding author:** Joanne M. Murabito, 73 Mount Wayte Ave, Framingham, MA 01702; (508)935-3400; fax (508) 626-1262; murabito@bu.edu

Total word count: 5223 (without conclusion or figure legends)

**Funding.** This work was support by the National Institutes of Health, National Institute of Aging R01AG076757 to JMM.

**Conflict of Interest.** DDM has received research support from Fitbit, Apple Inc, Bristol–Myers Squibb, Boehringer–Ingelheim, Pfizer, Flexcon, Samsung, Philips Healthcare, and Biotronik, and he has received consultancy fees from Heart Rhythm Society, Bristol–Myers Squibb, Pfizer, Fitbit, Flexcon, Boston Biomedical Associates, VentureWell, Avania and Rose Consulting. DDM also declares financial support for serving on the Steering Committee for the GUARD-AF study (NCT04126486) and the advisory committee for the Fitbit Heart study (NCT04176926). Eric Schramm is an employee of CareEvolution, Inc., a healthcare technology company. The remaining authors have no conflicts of interest to declare.

## Abstract

**Background:** Studies evaluating the usability of mobile-phone assessments in older adults are limited.

**Objective:** The aim of this study was to identify design-based barriers and facilitators to mobile app survey completion among two samples of older adults; those in the Framingham Heart Study and a more diverse sample from a hospital-based setting.

**Methods:** We used mixed methods to identify challenging and beneficial features of the mobile app in participants from the electronic Framingham Heart Study (eFHS,  $n=15$ , mean age 72 years, 40% women, 100% non-Hispanic and White) and among participants recruited from a hospital-based setting ( $n=15$ , mean age 71 years, 47% women, 20% Hispanic and 53% non-White). A variety of app-based measures with different response formats were tested including self-reported surveys, pictorial assessments (to indicate body pain sites), and cognitive testing tasks (e.g., Trail Making Test (TMT) and Stroop). Participants completed each measure using a think-aloud protocol, while being audio- and video-recorded with a qualitative interview conducted at the end of the session. Recordings were coded for participant usability errors by two pairs of coders. Participants completed the Mobile App Rating Scale (MARS) to assess the app (response range 1=inadequate to 5= excellent).

**Results:** In eFHS participants, the average total MARS score was 7.6 ( $SD=1.1$ ), with no significant differences in the hospital-based sample. In general participants were pleased with the app and found it easy to use. A large minority had at least one navigational issue, most committed only once. Most older adults did not have difficulty completing the self-reported multiple choice measures unless it included lengthy instructions but participants had usability issues with the Stroop and TMT.

**Conclusion:** Our methods and results help guide app development and app-based survey



construction for older adults, while also giving consideration to sociodemographic differences.

Word count 291

Key words: mobile application surveys; mixed methods; electronic data collection



## Introduction

Adults aged 65 years and older are increasingly using tablets and smartphones and engaging with a range of technologies.[1] Technology use can reduce social isolation[2] and enhance communication with family members and healthcare providers thereby increasing well-being. Further, digital technologies have the potential to improve the health of older adults by facilitating symptom monitoring and self-care management as well as monitoring cognitive and mobility decline. However, older adults often lack confidence in their ability to use technology[3] and report needing assistance with new electronic devices and mobile applications.[4] They face unique challenges with technology including poor eyesight, hearing loss, fine motor skill and sensory limitations, and cognitive decline. These challenges make it essential to understand how technologies can be made more useful to older adults. Perceived value, usefulness, and impact on quality of life are important predictors of technology adoption in this age group.[5] [6] In addition, design that minimizes user frustration will enhance use and lower the risk of leaving older users out of the technology revolution.

Older adults are often not well represented in user testing of technology[7] due to the restricted age range of research studies, physical or sensory impairments, or because technology studies may be less appealing to them. There are a growing number of smartphone applications (apps) that include opportunities for self-management of specific diseases and cognitive self-assessment but the quality and usability of the apps are often unknown especially among groups of older adults and adults from diverse race and ethnic populations.[8, 9] In addition, health care providers and hospital systems are increasingly requesting that patients complete pre-visit health questionnaires electronically, which help with care efficiency and are preferred by providers.[10, 11] However, older adults are less likely to access and use patient portals and may have unique needs influencing use.[12-14]

Usability information provides practical recommendations that can help increase patient responsiveness to electronically collected data. For example, studies that have evaluated the usability of mobile apps that assess fall risk demonstrated the importance of simple instructions and clear feedback such as a color change to indicate task completion. [6, 15, 16] A mobile app designed for older adults with heart failure to report patient reported outcomes measurement information system (PROMIS) measures demonstrated that these adults successfully returned the PROMIS data and an additional survey indicated high levels of usability.[17]

We designed a smartphone app for use by community dwelling older adults who are participants of the Framingham Heart Study (FHS). The smartphone app consists of surveys with different response formats (e.g., multiple choice, pictorial, tasks). The aim of the current study was to identify design-based barriers and facilitators to mobile app navigation and survey completion through usability testing. We also sought to understand whether participant feedback differed depending upon the response format of each measure. Because Framingham Heart Study participants were White, we enrolled a diverse sample of older adults at a second site to understand if any additional barriers to mobile app survey completion were observed given that digital literacy and preferences for using technologies can vary across older adults of different race and ethnic backgrounds. [12, 18] Importantly, little is known about the usability of mobile apps for racially/ethnically diverse populations.[8]

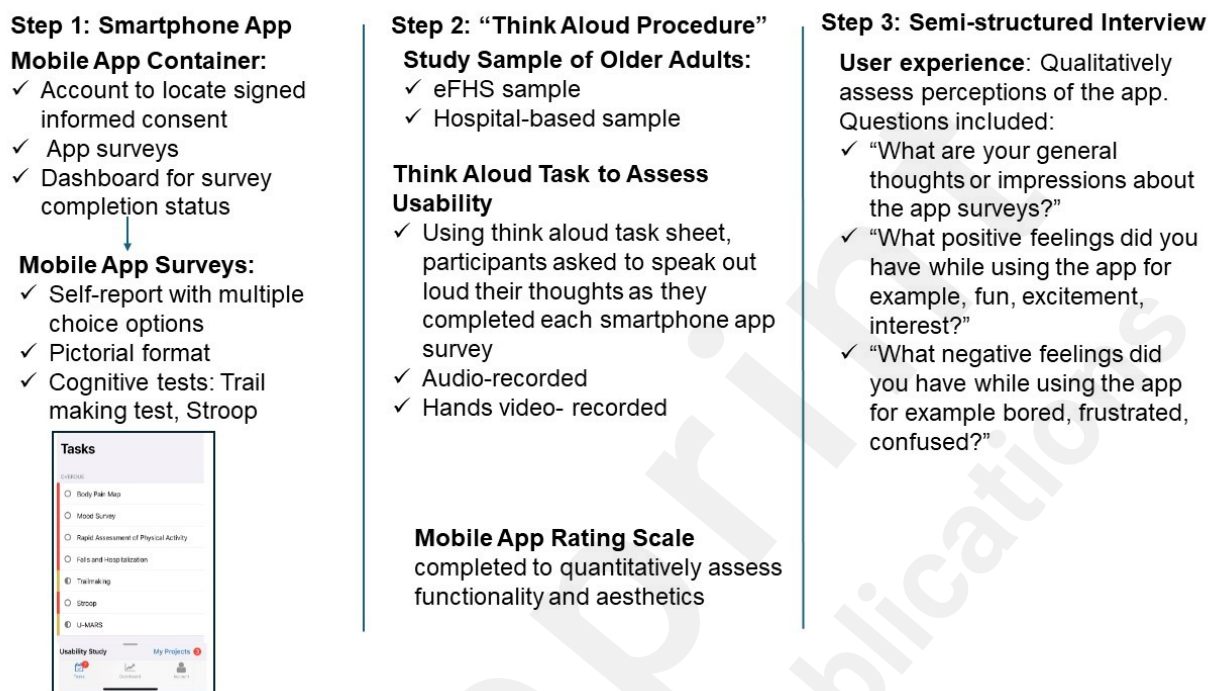
## Methods

### Study Design

The study used mixed methods to conduct one usability testing session followed by a post-session interview with enrolled participants (**Figure 1**). Usability testing methods included using the “Think Aloud” protocol while conducting a series of surveys and tasks on the

smartphone app. This was followed by a semi-structured interview using an interview guide, to solicit information on barriers encountered in the session.

**Figure 1. Overall Study Design**



### Study sample

The study sample was drawn from two sites: the FHS Offspring study and a hospital-based site. The FHS Offspring participants were recruited in 1971 and are invited back to the research center for examination every 6 to 8 years.[19] The tenth examination of the FHS Offspring participants occurred from 2019 to 2022 and included a mobile health component called the electronic Framingham Heart Study (eFHS). For the current study, we enrolled eFHS participants who were English speaking, owned a smartphone (iPhone or Android), attended Offspring exam 10 before the eFHS began, and enrolled in eFHS between July and September 2022. The eFHS research technician assisted participants with registration, informed consent, and app download. Because most of the eFHS sample

had iPhones and were less racially and ethnically diverse, we recruited participants from a second site who were not part of eFHS. The second site was a hospital in an urban area with a racially and ethnically diverse patient population. At the second site, inclusion criteria were age 60 years and older, English speaking, and able to attend a study session between December 2022 and April 2023. Participants were recruited through flyers placed in clinic waiting rooms and community centers, participation in prior research studies, and through patient registry lists. Non-white and Hispanic/Latino patients were oversampled from patient registry lists. As with eFHS, the research technician assisted the participant with registration, informed consent, and app download on a study iPhone. Participants at both sites were using the app for the first time. We enrolled fifteen participants at each study site to ensure representation of men and women, iPhone and Android users, and older people below and above 75 years. In addition the number of unique challenges identified with the study design proposed appears to asymptote at fifteen participants.[20] While eFHS participants were not compensated for their time in the study (because FHS participants have not been compensated for participation in the parent study), participants from the hospital-based cohort were provided a \$100 card for participating upon completion of the session.

The Institutional Review Board at Boston University Medical Campus approved the eFHS study (H-36586) and the current study (H-42659). The Institutional Review Board at the University of Massachusetts Chan Medical School approved the hospital-based study (#00000567).

## *Measures*

### *Study app*

A mobile app hosted a compendium of different types of survey assessments and

tasks that users could click on to complete. CareEvolution's MyDataHelps™ Designer platform ([https:// mydatahelps.org](https://mydatahelps.org)) was used to build the smartphone app surveys for iPhone (iOS 10.0 or higher) or Android (version 7.0 or higher) devices. The MyDataHelps mobile app container includes an account where participants can locate their signed consent form, tasks (app surveys), and a dashboard. The dashboard was created in the app to provide the participant with survey completion status and encouragement with a thank you message. Investigators and CareEvolution industry partners internally tested the app surveys and tasks with attention to consistency inspection, and user centered design principles to ensure clear instructions, easy navigation, and simple words and sentences. [21]

Because the goal of the study was to assess usability, we included a variety of app-based surveys and tasks with different response formats (Supplement 1). First, we tested several self-report surveys with multiple choice options including: 1) the short form of the PROMIS measure of mood (anxiety and depression, 8-items)[22] and cognitive function (4-items)[23] that included five response options that ranged from “never” to “always” for the mood assessment and “not at all” to “very much” for the cognitive function assessment; 2) falls and hospitalizations with yes/no response options and a calendar wheel to ascertain the date of a fall occurrence; and 3) Rapid Assessment of Physical Activity[24] to report frequency of physical activity and level of intensity (light, moderate and vigorous), assessing frequency, intensity, and time, with yes or no response choices. The hospital-based site did not complete the PROMIS measure of mood. Second, we tested a pictorial format measure to collect data on chronic pain, based on a modified version of the Michigan Body Pain Map measure.[25] Participants were shown an outline of a human body and were asked to click on the different places of the body where they are currently

experiencing chronic pain, defined as pain lasting three months or longer. The app first displays the image of the front of the body and next the back of the body. If the participant is not experiencing chronic pain, a box was provided so that the user can indicate 'no chronic pain.' The third and final smartphone tasks tested were two commonly used cognitive tests. The Trail making test[26, 27] is a timed assessment that requires the user to consecutively taps dots in alternating order between numbers and letters by first tapping the number "1" followed by tapping the letter "A" and then "2" followed by "B" until the user reaches the number "7". Correct answers result in the appearance of lines between the dots, resulting in a 'trail.' Finally, we assessed the Stroop [28] on the smartphone that requires the user to complete a series of four increasingly more difficult tasks responding as quickly and accurately as possible to changes in color and instructions. Because the Stroop requires the user to be able to see colors (yellow, green, red, and blue), persons with color blindness are not eligible for this task. A practice session was provided for each set of the four tasks with the ability to repeat the practice session should the user desire to do so. [29] At the end of the testing session participants completed the mobile app rating scale (MARS) on the smartphone to assess app functionality and aesthetics including ease of use, navigation, visual appeal, performance, graphics, and layout.[30] Items were rated by the participant using a 5-point Likert scale from 1= inadequate to 5= excellent.

Demographic data were not collected within the study app. For eFHS participants this data was collected as part of Offspring exam 10. Sociodemographics such as age, gender, and employment were assessed via the research study coordinators and entered into a secure web-based software platform, REDCap at the hospital-based site.

### *Procedure*

Study sessions were conducted in person in the FHS Research Center. eFHS

participants were asked to complete the above assessments using their own smartphone, but hospital based participants used a study smartphone (iPhone 7). In addition, they were also asked to navigate to different areas of the MyDataHelps app container (account, tasks, dashboard). While doing these tasks, participants were asked to “think your thoughts” out loud, including feelings (positive like ‘fun’ and negative like ‘frustrating’). Prior to beginning the usability testing, the research technician demonstrated the “think aloud” procedure using the text app on a smartphone and then asked the participant to demonstrate the think aloud procedure using the same app verbalizing every movement, feeling, and decision. Once the participant was ready to begin, the research technician encouraged the participant to use the Think Aloud Task Sheet (Supplement 1) that included a list of app-based tasks. Participants were asked to speak out loud their thoughts, feelings and actions as they completed each task. The participant was audio-recorded, and the participant’s hands were video recorded throughout the Think Aloud procedure. The participant also completed the MARS[30] on the smartphone after the Think Aloud procedure. The technician was present to audio and video-record the Think Aloud procedure and to encourage participants to speak their thoughts out loud. The technician was explicitly trained not to assist the participant with the app unless the participant was irrevocably stuck.

After completion of the Think Aloud procedure, the research technician conducted a 15- 20 minute interview using open-ended questions and reflective listening to obtain participant feedback on their experiences using the app. Interview questions are available in the Supplement and included questions such as “What are your general thoughts or impressions about the app surveys?”; “What positive feelings did you have while using the app for example, fun, excitement, interest?”; and “What negative feelings did you have while using the app for example bored, frustrated, confused?” In addition the interviewer



asked the participant to what extent family, friends, people of their own age would be able to use and enjoy the app and thoughts on how to ensure the app would be acceptable to people of different cultures. The research technicians (AD, NA, AH) are co-authors of this work.

### *Research technician training*

Four interviewers were trained (by BB and JF), three were bachelor level and one was PhD level. Training consisted of learning the Think Aloud procedure and also how to conduct the post-think aloud interview. Building rapport and communication skills (open-ended questions, reflective listening) were part of the training. Training included didactics and role plays and trainees were required to complete three sessions with 'mock' participants, supervised by one or both trainers before being cleared to do the protocol with study participants. Feedback on study participants was provided to research technicians on an ongoing basis, by viewing the audio and video tapes together. The audiotaped portion of the Think Aloud procedure and interview was professionally transcribed (Daily Transcription, Los Angeles, CA).

### *Process of Coding the Sessions*

Investigators developed a coding sheet and accompanying coding manual for use by teams of coders when coding the video- and audio- recordings of the Think Aloud procedure and post interviews. The coding sheet included general items (e.g., navigation between surveys, tapping in the wrong area to advance to the next task, unclear instructions for surveys) as well as assessment-specific variables (clarity of instruction, understanding concepts, navigating within a survey, and "look and feel" e.g. font size, line spacing, color). For training purposes, all coders reviewed three participant recordings together and resolved any coding discrepancies before separating into the two coding

teams to code in pairs. In order to maintain coding reliability over time, the two teams (team one: JF, JM; team 2: BB, DDM) independently coded the same participant recordings on five occasions throughout the coding process and came together to review the coding sheets for any discrepancies across teams to ensure all coders were following the coding guidelines. The average percent agreement between coders ranged from 80.5% to 98% across the five recordings that were coded in common by the two teams.

### *Statistical Analysis*

Descriptive statistics of the two study samples used mean  $\pm$  standard deviation for continuous variables, numbers and percentages for nominal variables. T-tests were applied for the comparison of continuous variables, and chi-square tests or Fisher's exact tests were used to compare nominal variables between the samples. The percent agreement between coding teams was calculated by the ratio of the number of discrepancies divided by the number of items in the coding sheet. For the MARS, all items used a 5-point Likert scale and the mean score for each domain (functionality, aesthetics) was calculated separately and an overall MARS score was computed for each of the two study samples. In addition, we calculated the mean score of each item within a domain (e.g. layout, graphics, visual appeal).

## Results

### *Study Sample*

In eFHS, 15 participants signed informed consent (mean age 72 years, SD=4.2, age range 64 to 80 years, 6 of 15 (40% ) women, 100% non-Hispanic, White). All eFHS participants owned a smartphone with 9 of 15 (60%) of the eFHS study sample participants owning an iPhone (Table 1). In the

Table 1. Participant Characteristics
--------------------------------------

	eFHS n=15	Hospital-based sample n=15
Age, mean (SD) and range (years)	72 (4.2) range: 64, 80	70.6 (6.2) range: 62, 79
Women (%)	6 (40%)	7 (46.7%)
Non-White (%)*	0 (0 %)	8 (53.3%)
Hispanic (%)	0 (0 %)	3 (20%)
Smartphone owner, iPhone	9 (60%)	6 (42.9%)
Smartphone, Android	6 (40%)	8 (53.3%)
Bachelor's degree and higher	10 (66.7%)	9 (60%)
Marital status- married, living as married	12 (80%)	8 (53.3%)
Income, < \$55,000	4 (31%)	5 (33%)
Subjective health, very good or excellent*	13 (86.7%)	5 (33.3%)
Characteristics were presented as mean $\pm$ standard deviation for continuous variables, numbers and percentages for nominal variables. All Offspring participants owned a smartphone (iPhone or Android); One hospital-based participant did not own a smartphone; # n=2 in eFHS sample income was unknown  * $P<.01$		

hospital-based site, 77 were contacted, 19 declined to participate, 3 deferred enrollment to a later date, and 15 participants signed inform consent (mean age 70.6 years, SD=6.2, age range 62 to 79 years, 7 of 15 (47%) women, 3 of 15 (20%) Hispanic/Latino, and 8 of 15 (53%) non-White). In the hospital-based sample, one participant did not own a smartphone and among smartphone owners six of 14 (42.9%) owned an iPhone. More than half of participants at both sites had a college education or advanced degree. While nearly 90% (13 of 15 participants) of eFHS participants reported their health to be very good to excellent, only one-third (5 of 15 participants) of the participants at the hospital-based site did ( $p=0.003$ ).

In the eFHS sample the average length of time of the Think Aloud procedure was 25.5 minutes (range 12.4 to 44.0 minutes) and the post procedure interview time on average was 18.7 minutes (range 12.4 to 20.9 minutes). Two participants declined the

interview (Think Aloud times were 28.5 and 33.35 minutes respectively). At the hospital-based site the average length of time of the Think Aloud procedure was about 11 minutes longer (mean 36.5 minutes, range 24.1 to 55.1) while the post procedure interview time was similar to eFHS with an average of 18.5 minutes (range 10.6 to 35.1). No participant at the hospital-based site declined the post-Think Aloud interview.

#### *Barriers to App Use identified during Think Aloud and Post Procedure Interview*

Table 2 presents the themes identified from the Think Aloud task and post procedure interview along with sample participant quotes.

**General Structure of the app.** In general, participants were pleased with the general structure of the app. However, a few participants had not used an app previously and they said it was “not intuitive” but “once you run through it a few times it isn’t a problem.” Participants generally thought that the app had the right amount of surveys (and that additional surveys would make them bored). Participants suggested that a progress indicator be inserted to let users know the status of survey completion. One participant also raised the importance of conveying information regarding the security of the app.

**Ease of Use.** Overall, participants found the app ‘easy to use,’ “simple” and “fun.” One participant said that it looked and functioned similarly to apps he was already familiar with, such as a social security app. A small minority of participants, however, wanted increased font size and more user friendly colors. One participant was frustrated with the functionality of the app, but he blamed it on using an old phone. There were some suggestions for modifications, such as tailoring the app by age group and disability, so it is not so ‘one size fits all.’ Another participant suggested removing the free text space (which was needed for one question) or adding a digital keyboard so it is more intuitive that the person is supposed to type. For example, upon completion of the Stroop task, there was a

query (“did you encounter any issues during the task?”) that permitted the participant to freely type in a response.

Navigation. A large minority of participants reported at least one navigational issue, including navigating within and between surveys and also clicking the incorrect area in the app

Table 2. Barriers and facilitators identified during the Think Aloud and Post Procedure Interview

Domain/Survey type	Theme	Sample Quotes
General Structure, satisfaction	<ul style="list-style-type: none"> <li>Information about security of app</li> <li>Use to raise awareness of health</li> <li>Everyone could learn from it</li> </ul>	<ul style="list-style-type: none"> <li>“The more I do, the easier it gets”</li> <li>“I just think it's gonna help me monitor my health and go from there. I like technology that helps me and doesn't just amuse me or keep going and if this helps me stay healthy, stay fit, it makes sense.”</li> </ul>
Ease of use	<ul style="list-style-type: none"> <li>Easy, simple, fun</li> <li>Questions easy to answer</li> <li>Confusion related to not paying attention and not reading the app instructions carefully enough</li> <li>Font color yellow difficult to see</li> <li>Font could be bigger</li> </ul>	<ul style="list-style-type: none"> <li>“The app spells everything perfectly clear after you read it a couple times to get it”</li> <li>“I wasn't really paying attention to what I was reading. That was mostly my problem.”</li> </ul>
Navigation	<ul style="list-style-type: none"> <li>Navigating within and between surveys was easy</li> <li>More detail on getting back to the homescreen and using “next”, “cancel” and “done”</li> </ul>	<ul style="list-style-type: none"> <li>“The more I did it, the more I was able to figure out how to get from one place to another.”</li> <li>“You might want to give a little more detail about some of the sections, like what happens when you cancel, what happens when you hit done.”</li> </ul>
Multiple choice survey formats	<ul style="list-style-type: none"> <li>Surveys were interesting and simple except physical activity survey with lengthy definitions and confusing response choices</li> </ul>	<ul style="list-style-type: none"> <li>“It was easy the questions were simple and it was easy to find the answers.”</li> </ul>
Pictorial format	<ul style="list-style-type: none"> <li>Body pain map did not include all areas that can be a real problem</li> <li>Checkbox worked differently than response choices for other surveys</li> <li>Hard to find the “no pain” box –</li> </ul>	<ul style="list-style-type: none"> <li>“Well, on the pain thing, I think they should have something near the anus. Because that can be a real headache. And then, of course, for women, it</li> </ul>

	small font	would also include, uh, the uterus.”
Cognitive tasks: Trail Making Test, Stroop	<ul style="list-style-type: none"> <li>• Stroop was confusing</li> <li>• Difficulty understanding practice session vs testing session</li> <li>• Confusion over 4 increasingly challenging sets of tests within the Stroop task</li> <li>• TMT had difficulties with instructions and navigation but the eFHS sample did not</li> </ul>	<ul style="list-style-type: none"> <li>• “The all underling thing was a little confusing to me. I think because I was trying to get through it quick. I lost track of whether underline the word or is it the color or is it this color or the word?”</li> <li>• “The trail-making one I liked”</li> </ul>
Friends/family	<ul style="list-style-type: none"> <li>• Fear of technology;</li> <li>• Little interest in learning how to use technology</li> <li>• Learning curve for older people</li> </ul>	<ul style="list-style-type: none"> <li>• “I think they would find it very useful I think it's very useful, just for like, the cognitive part of it.”</li> </ul>
Different cultural backgrounds	<ul style="list-style-type: none"> <li>• Need the app to be available in different languages</li> </ul>	<ul style="list-style-type: none"> <li>• “I come from the Indian community, and we place a lot of value on education, you know, I think this is the kind of thing that, you know -- I would -- we do -- I would like to be quizzed, and this is a quiz.”</li> <li>• “For example, I came from Burma. Burma used to be very underdeveloped country, and, also, still lots of problem. But they are very good at technology.”</li> </ul>

to navigate to the next screen (e.g., navigating to ‘done’). The vast majority of these errors were committed only once. Participants suggested making the buttons look more like buttons or have them flash when you are supposed to press the button, “...making it extremely obvious what you are supposed to hit.” Participants also suggested that more detail should be added, such as what will happen after you hit a button (e.g, after “cancel” or ‘done’ is pressed), and more information about what the “icons at the bottom of the app” mean (e.g., dashboard). A few participants mentioned that the app should avoid having to scroll to see more information, and instead put all of the information on one page if possible or continue to the next page.

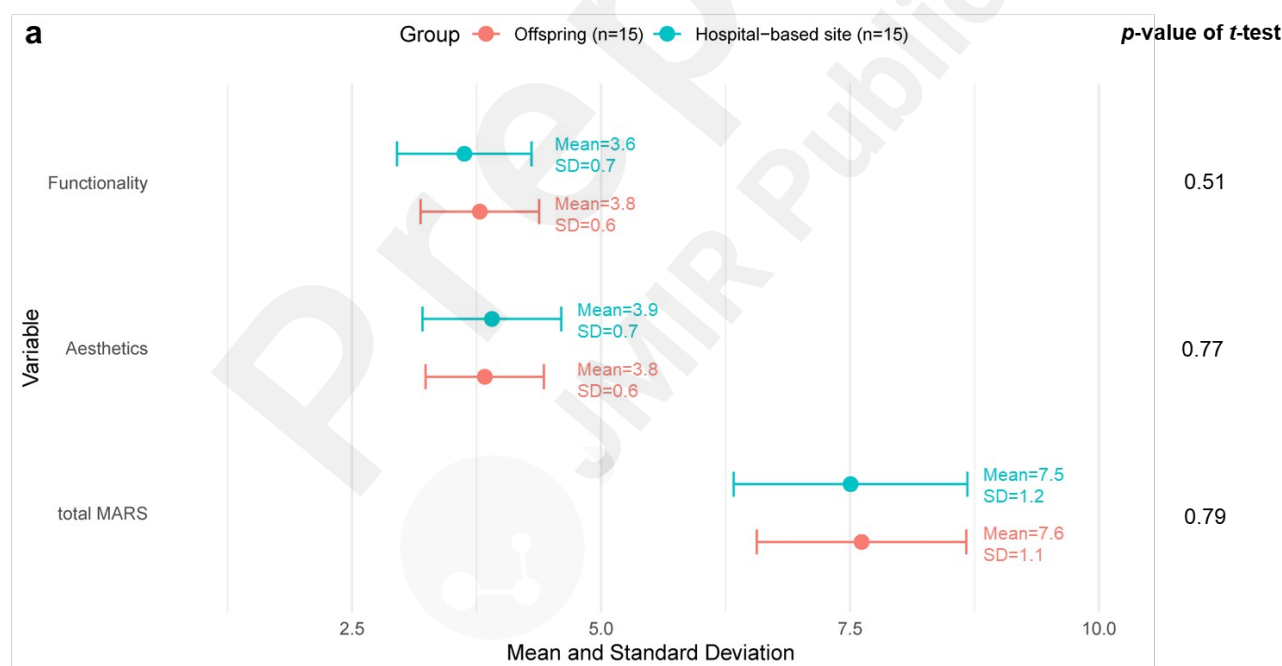
### *Response Choice Formats*

Multiple Choice formats. Participants reported that most of the multiple choice surveys were easy and interesting to complete, with only a few sporadic usability errors by a few participants. For the most part, participants reported that the instructions were easy to understand, with one exception, which was "The Rapid Assessment of Physical Activity" survey, which included lengthy definitions of physical activity levels (e.g., mild, moderate and vigorous activity) which were needed to answer the subsequent questions. Participants were confused by the content and format of the definitions. With regard to the latter, both physical activity intensity level and physical activity frequency in a single question (e.g., *I do some light or moderate physical activities but not every week*, yes or no).

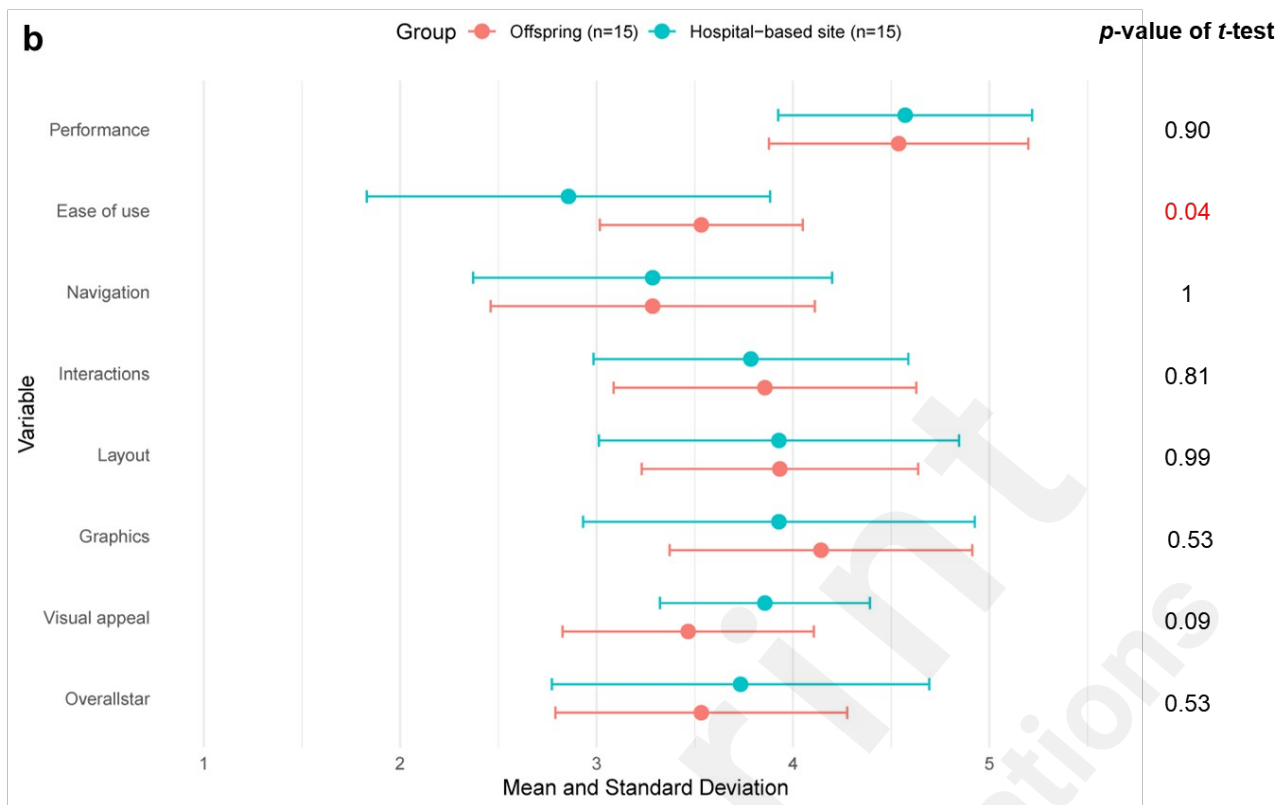
Pictorial Format. For the pictorial format assessment, (modified version of the Michigan Body Pain map), a small minority of participants noted the checkbox for response worked differently than the other surveys; instead of seeing a checkmark in the response box, the response box if selected was highlighted in color and was confusing for some. Some participants also noted that the map did not include all body areas where pain is present. For some participants, the "no chronic pain" response box was difficult to find due to small font, and the small font used to label the body parts was also difficult for several participants (but necessary to eliminate scrolling to see the entire body outline). There were no participants who reported difficulty understanding the instructions or concept. Only a couple of participants did not realize that they had to press 'next' to go to the next page to view the back of the body to indicate pain areas there. One participant said that bowels and reproductive areas give people his age a lot of issues, and that these areas should be added to the body pain map.

Cognitive Testing Formats. In terms of the formats used for cognitive testing, most older participants had some difficulty with the process of completing these measures. In terms of the Stroop, some participants had difficulty understanding that they were in the practice session vs. the test session. Participants also thought they needed to do the testing quickly and were confused if they did not read the instructions for each of the 4 test sets, as each had different instructions. Most participants did not take advantage of the opportunity to repeat the practice session when asked by the app, even when they were confused by the task. The yellow font color for the Stroop was also problematic for a few participants. One participant suggested that

Figure 2: Mobile App Rating Scale Scores by Study Sample: (A) Overall MARS Functionality and Aesthetics Scores; (B) Individual items within the Functionality and Aesthetics domains







voice and animation be used to explain the instructions for the Stroop. For the TMT, the eFHS sample did not report any difficulties with instructions or navigating within the task, and seemed to have a good understanding of the concepts included in the task. However, a large minority of the hospital-based sample had difficulties in all three of these areas. There were no difficulties reported on the look and feel of the task (e.g., font, line spacing, color).

Relevance for Friends/Family/Different Cultures. Participants noted they knew older people who engaged with technology and others who were not interested in the “electronic age”, did not use a computer or smartphone and had no interest in learning. One participant said that she had a friend in her 90’s who would easily be able to interact with the app and another friend of the same age who would have more difficulty. Participants noted older adults may have a fear of technology, be less confident using technology, and need assistance or a training session given that, for older people who have not used computers or technology, using the app would “be like a foreign language to them.” Participants provided their thoughts for using the app in older adults of different cultural backgrounds. Many noted the current version of the app is available only in the English language and would need translation to other languages. One participant was from a country that they felt embraced technology, and another participant reported her culture valued education and felt like the app included an educational like component like a “quiz” which would be viewed positively by her culture.

Satisfaction with the app. The mean total MARS score (7.6, SD 1.1), mean functionality score (3.8, SD 0.6), and mean aesthetics score (3.8, SD 0.6) in the eFHS sample did not significantly differ from the hospital-based sample (Figure 2a). With the exception of ease of use, the individual items of the functionality and aesthetics scores (performance, navigation, interactions such as taps, swipes, layout, graphics, and visual

appeal) did not significantly differ between the two samples (Figure 2b). Ease of use may have differed as participants at the hospital-based site used a study iPhone whereas eFHS participants used their own smartphone. The performance item was rated the highest with a mean of 4.5 in both samples. The mean overall star rating was 3.5 (SD 0.7) in the eFHS sample and 3.7 (SD 0.96) in the hospital-based sample indicating participants rated the app above average. In addition, during the interview, participants noted that the app could be used to raise awareness of health and people could learn from it, even though that was not the original purpose of the app. Some participants liked the app because it enabled one to 'express yourself' through the surveys. . Finally, a few participants suggested to make the app "more entertaining' by adding narration.

## Discussion

We tested the usability of a smartphone app designed to collect health information using complementary approaches in a community sample of older adult participants of the Framingham Offspring Study and to understand generalizability we also tested older adults from a more diverse hospital-based sample. In general participants liked the structure of the app and found the app was simple, fun, and easy to use. However, a large minority reported navigation issues that mostly occurred once with the ability to learn and figure out how to move within and between app-based surveys. A small minority of participants verbalized a preference for larger font size or more user-friendly colors. We observed that most older adults did not have difficulty with the multiple-choice app-based surveys unless the survey included lengthy instructions. Finally, most older adults experienced challenges with the app-based cognitive tasks especially the Stroop that required participants to read and understand a series of four increasingly more challenging tests with the Stroop task. Of importance, some participants noted that the app could be used to raise awareness of health and one could learn from it. Our observations confirm those of others that

involvement of older users can result in positive feelings among older adults, dispel stereotypes associated with older users, and the insights gained from older users can be used to enhance the quality of the design.[31] For example, our participants suggested using voice instructions and animated tutorials. To enhance usability app-designers and investigators, should consider training that includes tutorials within the app provided by an older adult guide to boost confidence, when designing smartphone apps for use by older adults.

This work has several implications for tailoring technology for older adult users. First, a guide within the app explaining the purpose of the app and highlighting key app functions including such functions as “next”, “back”, and “done” would enhance usability. Older adults are more likely to engage with technology that they perceive as useful, [5] therefore having a clear understanding of the goals of technology is critical. Further, the addition of basic training in smartphone use in older adults less proficient with technology was associated with less errors and less cueing during a smartphone app-based health prevention program and may result in improved engagement.[32] In our sample older adults who had not used an app before did not find it intuitive but after running through it a few times did not find it a problem. Our results support recently published app design guidelines for older adults advocating for initial training, if possible face-to-face, along with video instructions that are contextualized and provide step-by-step instructions to support older users .[33] Training may boost confidence and make the experience as frictionless as possible lowering the potential for abandonment. Second, streamlining and simplifying instructions may enhance understanding by inviting older adults to read them attentively. Participants in our study noted confusion that they attributed to not paying attention or related to the need to slow down and read directions more than once. Consistent with our observation, others have noted the need for clear and simple instructions when designing mHealth apps for older

adults.[6] Some participants also requested features they enjoyed in their use of other apps. Gamification of functions where possible such as a flashing done button or a countdown to the start of next task may improve engagement. Finally, consider voice narration and animated guides throughout the app surveys where features other than straightforward multiple choice questions and responses will be encountered.

Our study may have important insights to help address the continued digital disparities observed in older adults. Older adults are increasingly using the internet and smartphones,[34] however connection to the internet decreases across age with nearly half of young adults almost constantly connected vs. eight percent of adults age 65 years and older.[4] Other key digital health behaviors are also lower in a nationally representative sample of older adults including using health apps, using a digital device to track health or a health-related goal, and digitally communicating with a healthcare provider.[35] Digital technologies were lifesaving during the COVID pandemic as the rapid transformation from in-person visits to televisits permitted access to healthcare in a setting that provided social distance and did not expose vulnerable older adults to the virus. Similarly, the ability to participate in digital interventions may provide several benefits, such as improved memory and independent living,[36] physical functioning, physical activity,[37] depression, and anxiety.[38] Including older adults in technology design and conducting usability testing may address digital health inequities by addressing digital health literacy and creating programs that are user-friendly to this population.[39] They may also improve implementation beyond pilot studies and achieve the needed sustainability of technology solutions[40] for chronic disease management and home care options for older adults and at the same time may be one step in addressing digital disparities. We plan to use the smartphone app more widely in the Framingham cohorts as a tool to monitor health. We will be able to provide critical information on the characteristics of who enrolls and uses the technology and who chooses

not to.

Our study had several strengths. There is no “best” method to assess usability,[21] therefore we used both qualitative and quantitative methods. We tested older users with mean age over 70 years in both samples often not included in studies testing technology and included older adults from diverse race and ethnic backgrounds. The study sample included a range of older users. Participants without a smartphone or experience with app use and both iPhone and Android users were included. This strategy allowed us to uncover errors with the app beyond what would be observed with “regular” users and permit greater guidance in app re-design to benefit older users.

In addition, participants with health issues were included. Some older adults with health conditions or geriatric syndromes such as frailty have higher levels of nonuse of information communication technologies and more negative views on usefulness and usability.[41] Our study also had some limitations that merit comment. Participants at the hospital-based site used an iPhone only. This may have been a limitation if participants were Android owners or had a different iPhone version, however, this is also a strength as we were able to include participants who were not smartphone owners. Our observations focus on first interactions with the app-based surveys. It is beyond the scope of the study to examine other aspects of use such as efficiency (how quickly the survey/task is completed once the design is learned) and memorability (how easy to use after a period of not using the app) that may have important implications to research study designs. Continued engagement with technology changes over time in older adults but the factors related to continued technology use are unclear and require further investigation.[42] Our study took place in Massachusetts and therefore may not be generalizable to other geographic areas. Participants with color blindness were not eligible for the Stroop task and were excluded from testing. Therefore results may not be generalizable to this group of older adults. FHS

participants that enroll in eFHS are healthier and have higher levels of education than participants who chose not to enroll.

Our study of a diverse sample of older adults testing several different smartphone app survey types and response formats provides a guide to investigators and clinicians that can be used for future app development and app-based survey construction for older adults. Many older users are able to interact with and enjoy technology. Further work to enhance engagement among older users and diminish digital disparities in this group is needed if the potential of technology to improve well-being, functioning, and health in older adults is to be realized.

## Figure Legends

### Figure 1: **Overall Study Design**

Figure 2: Mobile App Rating Scale Scores by Study Sample: (A) Overall MARS Functionality and Aesthetics Scores; (B) Individual items within the Functionality and Aesthetics domains

**Multimedia appendix 1: Screenshots of app-based surveys and tasks, Think Aloud Task sheet, and post procedure interview questions**



## References

1. <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 April 11, 2022.
2. Umoh ME, Prichett L, Boyd CM, Cudjoe TKM. Impact of technology on social isolation: Longitudinal analysis from the National Health Aging Trends Study. *Journal of the American Geriatrics Society*. 2023 Apr;71(4):1117-23. PMID: 36519748. doi: 10.1111/jgs.18179.
3. Lee CC, Czaja SJ, Moxley JH, Sharit J, Boot WR, Charness N, et al. Attitudes Toward Computers Across Adulthood From 1994 to 2013. *Gerontologist*. 2018 Jul 5. PMID: 29982458. doi: 10.1093/geront/gny081.
4. <https://www.pewinternet.org/2017/05/17/barriers-to-adoption-and-attitudes-towards-technology/>. accessed 9/14/19.
5. Berkowsky RW, Sharit J, Czaja SJ. Factors Predicting Decisions About Technology Adoption Among Older Adults. *Innov Aging*. 2018 Jan;2(1):igy002. PMID: 30480129. doi: 10.1093/geroni/igy002.
6. Hsieh KL, Fanning JT, Rogers WA, Wood TA, Sosnoff JJ. A Fall Risk mHealth App for Older Adults: Development and Usability Study. *JMIR Aging*. 2018 Nov 20;1(2):e11569. PMID: 31518234. doi: 10.2196/11569.
7. Schumacher RM. Gran Got tech: Inclusivity and Older Adults. *Journal of User Experience*. 2023;18(2):62-7.
8. Sarkar U, Gourley GI, Lyles CR, Tieu L, Clarity C, Newmark L, et al. Usability of Commercially Available Mobile Applications for Diverse Patients. *J Gen Intern Med*. 2016 Dec;31(12):1417-26. PMID: 27418347. doi: 10.1007/s11606-016-3771-6.
9. 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. *Journal of medical Internet research*. 2020 Jan 17;22(1):e14551. PMID: 31951218. doi: 10.2196/14551.
10. Wald JS, Businger A, Gandhi TK, Grant RW, Poon EG, Schnipper JL, et al. Implementing practice-linked pre-visit electronic journals in primary care: patient and physician use and satisfaction. *J Am Med Inform Assoc*. 2010 Sep-Oct;17(5):502-6. PMID: 20819852. doi: 10.1136/jamia.2009.001362.
11. Cusatis R, Holt JM, Williams J, Nukuna S, Asan O, Flynn KE, et al. The impact of patient-generated contextual data on communication in clinical practice: A qualitative assessment of patient and clinician perspectives. *Patient Educ Couns*. 2020 Apr;103(4):734-40. PMID: 31744702. doi: 10.1016/j.pec.2019.10.020.
12. Gordon NP, Hornbrook MC. Differences in Access to and Preferences for Using Patient Portals and Other eHealth Technologies Based on Race, Ethnicity, and Age: A Database and Survey Study of Seniors in a Large Health Plan. *Journal of medical Internet research*. 2016 Mar 04;18(3):e50. PMID: 26944212. doi: 10.2196/jmir.5105.
13. Crouch E, Gordon NP. Prevalence and Factors Influencing Use of Internet and Electronic Health Resources by Middle-Aged and Older Adults in a US Health Plan Population: Cross-Sectional Survey Study. *JMIR Aging*. 2019 Mar 26;2(1):e11451. PMID: 31518256. doi: 10.2196/11451.
14. Irizarry T, Shoemaker J, Nilsen ML, Czaja S, Beach S, DeVito Dabbs A. Patient Portals as a Tool for Health Care Engagement: A Mixed-Method Study of Older Adults With Varying Levels of Health Literacy and Prior Patient Portal Use. *Journal of medical Internet research*. 2017 Mar 30;19(3):e99. PMID: 28360022. doi: 10.2196/jmir.7099.
15. Harte R, Quinlan LR, Glynn L, Rodriguez-Molinero A, Baker PM, Scharf T, et al. Human-

Centered Design Study: Enhancing the Usability of a Mobile Phone App in an Integrated Falls Risk Detection System for Use by Older Adult Users. *JMIR Mhealth Uhealth*. 2017 May 30;5(5):e71. PMID: 28559227. doi: 10.2196/mhealth.7046.

16. Mansson L, Wiklund M, Ohberg F, Danielsson K, Sandlund M. Co-Creation with Older Adults to Improve User-Experience of a Smartphone Self-Test Application to Assess Balance Function. *Int J Environ Res Public Health*. 2020 May 26;17(11). PMID: 32466484. doi: 10.3390/ijerph17113768.

17. Reading Turchioe M, Grossman LV, Baik D, Lee CS, Maurer MS, Goyal P, et al. Older Adults Can Successfully Monitor Symptoms Using an Inclusively Designed Mobile Application. *Journal of the American Geriatrics Society*. 2020 Jun;68(6):1313-8. PMID: 32157679. doi: 10.1111/jgs.16403.

18. Greer DB, Abel WM. Exploring Feasibility of mHealth to Manage Hypertension in Rural Black Older Adults: A Convergent Parallel Mixed Method Study. *Patient Prefer Adherence*. 2022;16:2135-48. PMID: 35999840. doi: 10.2147/PPA.S361032.

19. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Preventive medicine*. 1975 Dec;4(4):518-25. PMID: 1208363.

20. Blair J, Conrad, F., Ackermann, A.C., Claxton, G. The effect of sample size on interview findings. . AAPOR-American Statistical Association, Section on Survey Research Methods. 2005.

21. Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *Int J Med Inform*. 2009 May;78(5):340-53. PMID: 19046928. doi: 10.1016/j.ijmedinf.2008.10.002.

22. Cella D, Choi SW, Condon DM, Schalet B, Hays RD, Rothrock NE, et al. PROMIS((R)) Adult Health Profiles: Efficient Short-Form Measures of Seven Health Domains. *Value Health*. 2019 May;22(5):537-44. PMID: 31104731. doi: 10.1016/j.jval.2019.02.004.

23. Lai JS, Wagner LI, Jacobsen PB, Cella D. Self-reported cognitive concerns and abilities: two sides of one coin? *Psychooncology*. 2014 Oct;23(10):1133-41. PMID: 24700645. doi: 10.1002/pon.3522.

24. Rapid Assessment of Physical Activity. <http://deptswashingtonedu/hprc/rapa>. 2006.

25. Brummett CM, Bakshi RR, Goesling J, Leung D, Moser SE, Zollars JW, et al. Preliminary validation of the Michigan Body Map. *Pain*. 2016 Jun;157(6):1205-12. PMID: 26835782. doi: 10.1097/j.pain.0000000000000506.

26. Dahmen J, Cook D, Fellows R, Schmitter-Edgecombe M. An analysis of a digital variant of the Trail Making Test using machine learning techniques. *Technol Health Care*. 2017;25(2):251-64. PMID: 27886019. doi: 10.3233/THC-161274.

27. Ashendorf L, Jefferson AL, O'Connor MK, Chaisson C, Green RC, Stern RA. Trail Making Test errors in normal aging, mild cognitive impairment, and dementia. *Arch Clin Neuropsychol*. 2008 Mar;23(2):129-37. PMID: 18178372. doi: 10.1016/j.acn.2007.11.005.

28. Hanke LD, Preis SR, Piers RJ, Beiser AS, Devine SA, Liu Y, et al. Population Normative Data for the CERAD Word List and Victoria Stroop Test in Younger- and Middle-Aged Adults: Cross-Sectional Analyses from the Framingham Heart Study. *Experimental aging research*. 2016 Jul-Sep;42(4):315-28. PMID: 27410241. doi: 10.1080/0361073X.2016.1191838.

29. Jongstra S, Wijsman LW, Cachucho R, Hoevenaer-Blom MP, Mooijaart SP, Richard E. Cognitive Testing in People at Increased Risk of Dementia Using a Smartphone App: The iVitality Proof-of-Principle Study. *JMIR Mhealth Uhealth*. 2017 May 25;5(5):e68. PMID: 28546139. doi: 10.2196/mhealth.6939.

30. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth*. 2015 Mar 11;3(1):e27. PMID: 25760773. doi: 10.2196/mhealth.3422.

31. Fischer B, Peine A, Ostlund B. The Importance of User Involvement: A Systematic Review of

Involving Older Users in Technology Design. *Gerontologist*. 2020 Sep 15;60(7):e513-e23. PMID: 31773145. doi: 10.1093/geront/gnz163.

32. Harte R, Hall T, Glynn L, Rodriguez-Molinero A, Scharf T, Quinlan LR, et al. Enhancing Home Health Mobile Phone App Usability Through General Smartphone Training: Usability and Learnability Case Study. *JMIR Hum Factors*. 2018 Apr 26;5(2):e18. PMID: 29699969. doi: 10.2196/humanfactors.7718.

33. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design Guidelines of Mobile Apps for Older Adults: Systematic Review and Thematic Analysis. *JMIR Mhealth Uhealth*. 2023 Sep 21;11:e43186. PMID: 37733401. doi: 10.2196/43186.

34. <http://www.pewinternet.org/2017/05/17/tech-adoption-climbs-among-older-adults/>.

35. Ratcliff CL, Krakow M, Greenberg-Worisek A, Hesse BW. Digital Health Engagement in the US Population: Insights From the 2018 Health Information National Trends Survey. *Am J Public Health*. 2021 Jul;111(7):1348-51. PMID: 34014759. doi: 10.2105/AJPH.2021.306282.

36. Scullin MK, Jones WE, Phenis R, Beevers S, Rosen S, Dinh K, et al. Using smartphone technology to improve prospective memory functioning: A randomized controlled trial. *Journal of the American Geriatrics Society*. 2022 Feb;70(2):459-69. PMID: 34786698. doi: 10.1111/jgs.17551.

37. Stockwell S, Schofield P, Fisher A, Firth J, Jackson SE, Stubbs B, et al. Digital behavior change interventions to promote physical activity and/or reduce sedentary behavior in older adults: A systematic review and meta-analysis. *Experimental gerontology*. 2019 Jun;120:68-87. PMID: 30836130. doi: 10.1016/j.exger.2019.02.020.

38. Riadi I, Kervin L, Dhillon S, Teo K, Churchill R, Card KG, et al. Digital interventions for depression and anxiety in older adults: a systematic review of randomised controlled trials. *Lancet Healthy Longev*. 2022 Aug;3(8):e558-e71. PMID: 36102765. doi: 10.1016/S2666-7568(22)00121-0.

39. 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. PMID: 31447716. doi: 10.3389/fpsy.2019.00568.

40. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, Implementing, and Evaluating Mobile Health Technologies for Managing Chronic Conditions in Older Adults: A Scoping Review. *JMIR Mhealth Uhealth*. 2016 Jun 9;4(2):e29. PMID: 27282195. doi: 10.2196/mhealth.5127.

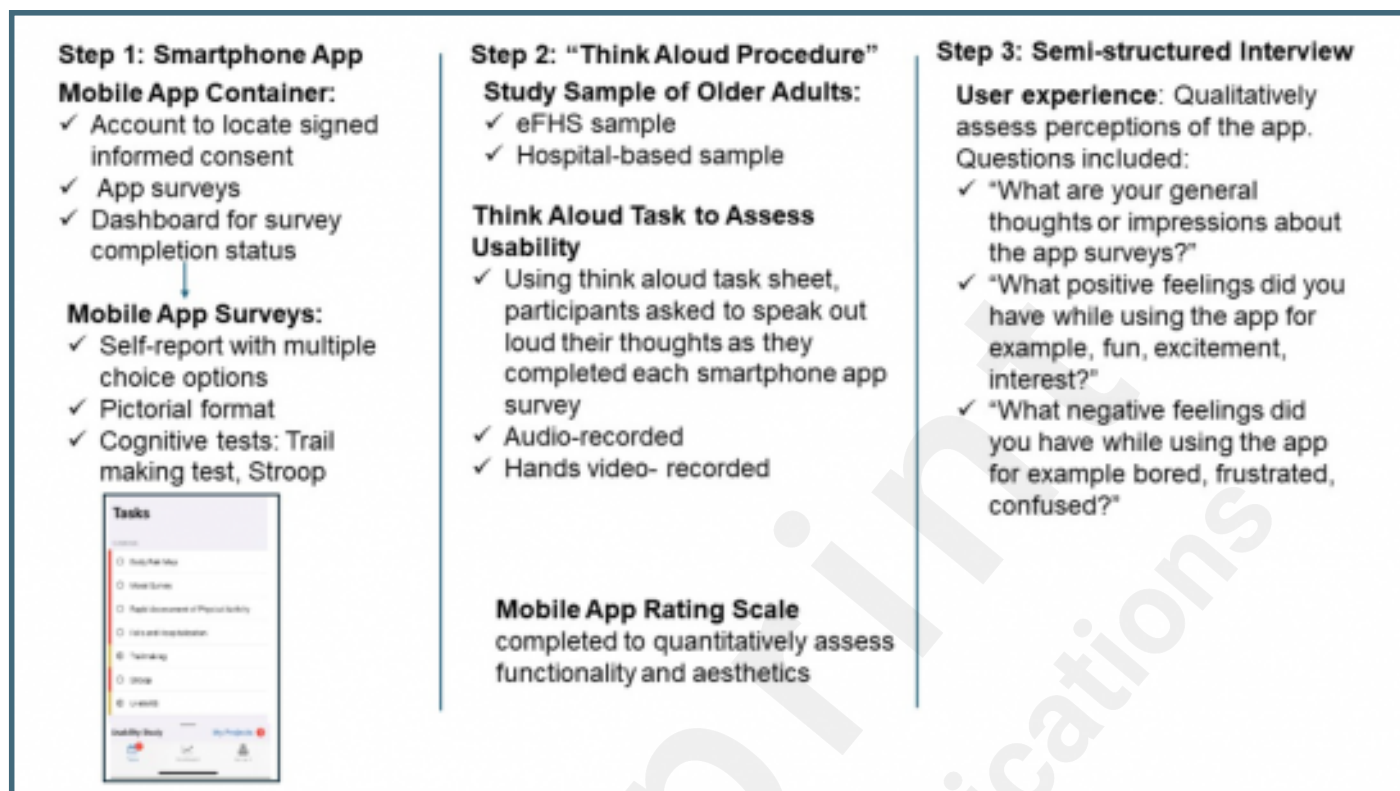
41. Keranen NS, Kangas M, Immonen M, Simila H, Enwald H, Korpelainen R, et al. Use of Information and Communication Technologies Among Older People With and Without Frailty: A Population-Based Survey. *Journal of medical Internet research*. 2017 Feb 14;19(2):e29. PMID: 28196791. doi: 10.2196/jmir.5507.

42. Zhang S, Boot WR. Predicting Older Adults' Continued Computer Use After Initial Adoption. *Innov Aging*. 2023;7(4):igad029. PMID: 37197443. doi: 10.1093/geroni/igad029.

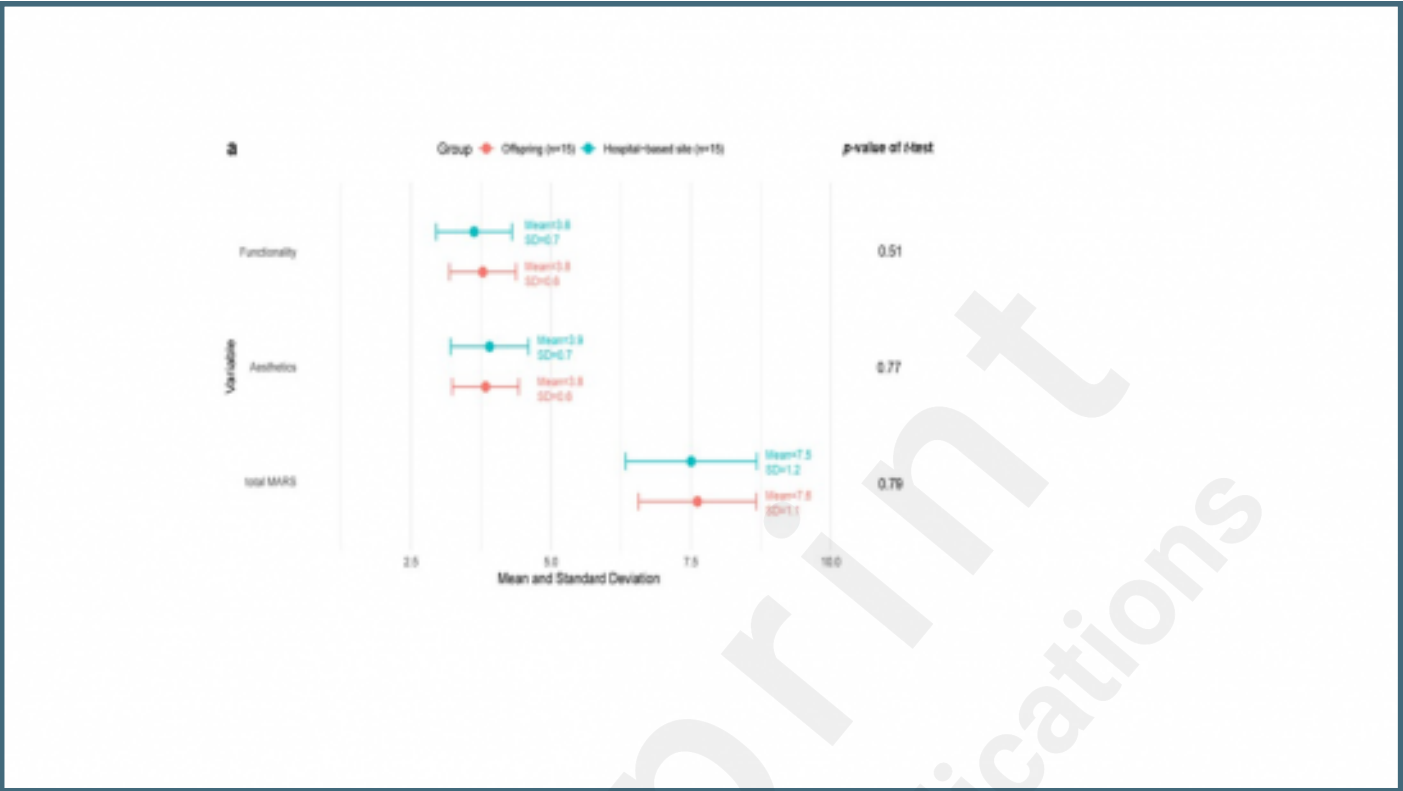
## Supplementary Files

## Figures

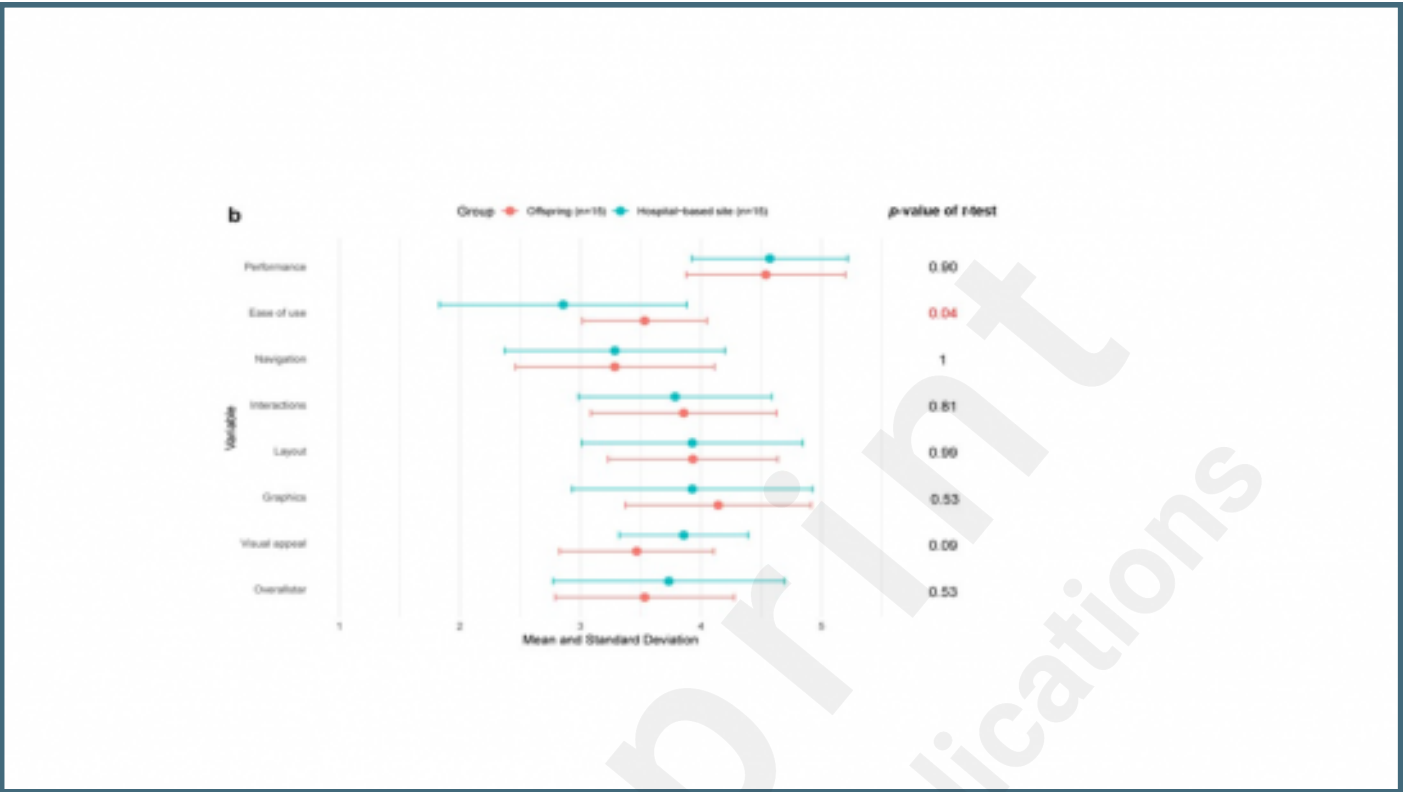
## Overall Study Design.



Mobile App Rating Scale Scores by Study Sample: (A) Overall MARS Functionality.



Mobile App Rating Scale Scores by Study Sample (B) Individual items within the Functionality and Aesthetics domains.





## **Multimedia Appendixes**

Manuscript revision tracked.

URL: <http://asset.jmir.pub/assets/f16e467c2b1f50a28a48731bf02c3c4d.docx>

Untitled.

URL: <http://asset.jmir.pub/assets/a49116072f1a85f54c62cfcee9be92de.docx>

