Mobility-Based Smartphone Digital Phenotypes Unobtrusively Capture Everyday Cognition, Mood, and Community Life-Space in Older Adults: A Pilot Study

Katherine Hackett, Shiyun Xu, Moira McKniff, Lido Paglia, Ian Barnett, Tania Giovannetti

Submitted to: JMIR Human Factors
on: April 29, 2024

Table of Contents

Original Manuscript 5
Supplementary Files 52
Figures 53
Figure 1 54
Figure 2 55
Multimedia Appendixes 56
Multimedia Appendix 1 57
Multimedia Appendix 2 57
Multimedia Appendix 3 57
Multimedia Appendix 4 57

Mobility-Based Smartphone Digital Phenotypes Unobtrusively Capture Everyday Cognition, Mood, and Community Life-Space in Older Adults: A Pilot Study

Katherine Hackett[1] PhD; Shiyun Xu[2] BS; Moira McKniff[1] BA; Lido Paglia[3] BS; Ian Barnett[2] PhD; Tania Giovannetti[1] PhD

[1]Department of Psychology and Neuroscience Temple University Philadelphia US
[2]Department of Biostatistics, Epidemiology, and Informatics University of Pennsylvania Philadelphia US
[3]Information Technology College of Science & Technology Temple University Philadelphia US

Corresponding Author:
Tania Giovannetti PhD
Department of Psychology and Neuroscience
Temple University
Weiss Hall, 6th Floor
1701 N 13th St
Philadelphia
US

Abstract

Background: Current methods of monitoring cognition in older adults are insufficient to address the growing burden of Alzheimer's disease and related dementias (AD/ADRD). New approaches that are sensitive, scalable, objective, and reflective of meaningful functional outcomes are direly needed. Mobility trajectories and geospatial life space patterns reflect many aspects of cognitive and functional integrity and may be useful proxies of age-related cognitive decline.

Objective: We investigated the feasibility, acceptability, and preliminary validity of a 1-month smartphone digital phenotyping protocol to infer everyday cognition, function, and mood in older adults based on passively-obtained GPS data. We also sought to clarify the intrinsic and extrinsic factors that are associated with mobility phenotypes for consideration in future studies.

Methods: A total of 37 older adults aged 63-85 with healthy cognition (n=28), Mild Cognitive Impairment (n=5) and mild dementia (n=1) used an open-source smartphone application (mindLAMP) to unobtrusively capture GPS trajectories for 4 weeks. Raw GPS data were processed into interpretable features across categories including activity, inactivity, routine, and location diversity. Monthly average and monthly intraindividual variability (IIV) metrics were calculated for each feature to test a priori hypotheses informed by a neuropsychological model. Comprehensive validation measures collected at baseline were compared alongside monthly GPS features to examine construct validity. Feasibility and acceptability outcomes included participant retention, comprehension of study procedures, technical difficulties, and satisfaction ratings at study debriefing.

Results: All (100%) participants completed the 4-week monitoring period without major adverse events, 100% reported satisfaction with the explanation of study procedures, and 97% reported no feelings of discomfort. Performance on the comprehension of consent quiz was 97% on average and associated with education and race. Technical issues requiring troubleshooting were relatively infrequent, though 41% of participants reported battery drain. Moderate to strong correlations (r>.3) were identified between GPS features and validators. Specifically, individuals with greater mobility activity, less routine, and more location diversity demonstrated better cognition, less functional impairment, less depression, more community participation, and more geospatial life-space on objective and subjective validation measures. Contrary to predictions, greater IIV in mobility was also associated with positive validation outcomes. While many demographic and technology-related factors were not associated with GPS features, income, native English language, season of study participation and occupational status were related.

Conclusions: Theoretically-informed digital phenotypes of mobility are feasibly captured from older adults' personal smartphone devices and relate to a variety of clinically meaningful measures including cognitive test performance, reported functional decline, mood, and community activity. Future studies should consider the impact of intrinsic and extrinsic factors when interpreting mobility phenotypes. Overall, smartphone digital phenotyping is a promising method to unobtrusively capture

relevant risk and resilience factors in the context of aging and AD/ADRD and should continue to be investigated in larger and more diverse samples.

(JMIR Preprints 29/04/2024:59974)
DOI: https://doi.org/10.2196/preprints.59974

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 v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Paper type: Original Paper

Manuscript title: Mobility-Based Smartphone Digital Phenotypes Unobtrusively Capture Everyday Cognition, Mood, and Community Life-Space in Older Adults: A Pilot Study

Authors: Katherine Hackett[1a], Shiyun Xu[2], Moira McKniff[1], Lido Paglia[3], Ian Barnett[2], Tania Giovannetti[1]

1. Department of Psychology and Neuroscience, Temple University, Philadelphia, PA, United States
2. Department of Biostatistics Epidemiology and Informatics, University of Pennsylvania, Philadelphia, PA, United States

3. Information Technology, Temple University, Philadelphia, PA, United States

[a] current affiliation: Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, United States

Corresponding Author:
Tania Giovannetti, PhD
Department of Psychology and Neuroscience, Temple University
1701 N 13th St, Philadelphia, PA 19122, United States
Phone: (484) 843-1321
Email: tania.giovannetti@temple.edu

Abstract

Background: Current methods of monitoring cognition in older adults are insufficient to address the growing burden of Alzheimer’s disease and related dementias (AD/ADRD). New approaches that are sensitive, scalable, objective, and reflective of meaningful functional outcomes are direly needed. Mobility trajectories and geospatial life space patterns reflect many aspects of cognitive and functional integrity and may be useful proxies of age-related cognitive decline.

Objective: We investigated the feasibility, acceptability, and preliminary validity of a 1-month smartphone digital phenotyping protocol to infer everyday cognition, function, and mood in older adults based on passively-obtained GPS data. We also sought to clarify the intrinsic and extrinsic factors that are associated with mobility phenotypes for consideration in future studies.

Methods: A total of 37 older adults aged 63-85 with healthy cognition (n=28), Mild Cognitive Impairment (n=5) and mild dementia (n=1) used an open-source smartphone application (mindLAMP) to unobtrusively capture GPS trajectories for 4 weeks. Raw GPS data were processed into interpretable features across categories including activity, inactivity, routine, and location diversity. Monthly average and monthly intraindividual variability (IIV) metrics were calculated for each feature to test *a priori* hypotheses informed by a neuropsychological model. Comprehensive validation measures collected at baseline were compared alongside monthly GPS features to examine construct validity. Feasibility and acceptability outcomes included participant retention, comprehension of study procedures, technical difficulties, and satisfaction ratings at study debriefing.

Results: All (100%) participants completed the 4-week monitoring period without major adverse events, 100% reported satisfaction with the explanation of study procedures, and 97% reported no feelings of discomfort. Performance on the comprehension of consent quiz was 97% on average and associated with education and race. Technical issues requiring troubleshooting were relatively infrequent, though 41% of participants reported battery drain. Moderate to strong correlations ($r \geq .3$) were identified between GPS features and validators. Specifically, individuals with greater mobility activity, less routine, and more location diversity demonstrated better cognition, less functional impairment, less depression, more community participation, and more geospatial life-space on objective and subjective validation measures. Contrary to predictions, greater IIV in mobility was also associated with positive validation outcomes. While many demographic and technology-related factors were not associated with GPS features, income, native English language, season of study participation and occupational status were related.

Conclusions: Theoretically-informed digital phenotypes of mobility are feasibly captured from older adults’ personal smartphone devices and relate to a variety of clinically meaningful measures including cognitive test performance, reported functional decline, mood, and community activity. Future studies should consider the impact of intrinsic and extrinsic factors when interpreting mobility phenotypes. Overall, smartphone digital phenotyping is a promising method to unobtrusively capture relevant risk and resilience factors in the context of aging and AD/ADRD and should continue to be investigated in larger and more diverse samples.

Keywords: digital phenotyping; digital biomarkers; monitoring; mHealth; cognition; mobility; life space; depression; location data; Alzheimer’s disease; aging

Introduction

Alzheimer’s disease and related dementias (AD/ADRD) place immense pressure on our healthcare system. They represent a global issue that is worsening and will be exacerbated by insufficient disease screening methods and lack of ecologically-valid outcome measures for clinical trials [1]. New and innovative approaches for early detection and monitoring are direly needed to address this global crisis. In this paper, we present results from a proof-of-concept study demonstrating the promise of smartphone digital phenotyping to capture clinically-relevant risk factors and outcomes in the context of aging and AD/ADRD.

Decades of clinical trial research indicates that early intervention will be critical for effective AD/ADRD treatment [2–5]. Biomarker testing (e.g., PET neuroimaging, cerebrospinal fluid); traditional neuropsychological evaluation; and clinician- or informant-rated assessments have historically been the gold-standard methods for detection and diagnosis. However, these methods present drawbacks including limited accessibility (e.g., proximity, cost); scalability (i.e., logistical constraints for wide scale clinical implementation); prognostic value (i.e., inconsistent correspondence with clinical progression); and ecological validity (i.e., poor representation of real-world functioning) [6–10]. Traditional neuropsychological measures also are affected by sociocultural factors such as educational quality, socioeconomic status, native language status, and acculturation – rendering them inappropriate for the increasingly diverse global population [11–20]. Although subtle functional difficulties on complex everyday tasks signal early decline, standard functional assessments have a limited ability to detect such subtle changes because they are plagued by recall bias, low spatiotemporal resolution, availability of an informant reporter, bias due to attributes of the informant, outdated items, and poor sensitivity [21–27]. Taken together, healthcare systems are ill-equipped to screen for early signs of cognitive impairment at scale, as evidenced by recent estimates that 92% of cases of MCI remain undiagnosed [28–31]. In addition to poor screening methods, traditional clinical trial outcome measures are not sufficiently meaningful and precise to demonstrate therapeutic benefit at early stages [4,25,32–37].Thus, a growing priority in the field of AD/ADRD is to develop and implement sensitive and functionally meaningful screening and outcome measures as new therapies are evaluated earlier in the disease course [38].

New digital assessment methods have great potential for efficient, accessible, sensitive, and objective assessment of early cognitive and functional changes reflecting risk for AD/ADRD [6,7,39–43]. Digital tools can capture micro-level behavioral data with increased sensitivity and reduced sample size requirements compared with traditional paper-and-pencil neuropsychological measures and functional scales [44]. Gathering this information at home can address accessibility limitations for those in rural environments or who face hardships traveling to/from a clinical site, and may provide a more reliable and ecologically valid representation of real-world functioning compared to traditional evaluation at a single timepoint in a highly controlled setting [45–47]. As global technology use and affordability of personal devices continues to rise [48,49], new technologies can potentially address crucial gaps in the scalability and accessibility of current methods and counteract higher rates of missed diagnosis among populations experiencing low socioeconomic status [31,50–54].

Digital phenotyping is one innovative approach that utilizes the “moment-by-moment quantification of the individual-level human phenotype *in situ”* based on interactions with technology, including smartphones and smarthome devices, to capture social and behavioral data passively, continuously, and with minimal interference [55–57]. It collects high-frequency, fine-grained data reflecting everyday behaviors “in the wild” without relying on user engagement, subjective report, or burdensome procedures. Preliminary support has been demonstrated in psychiatry studies leveraging a host of sensors (e.g., Global Positioning System [GPS], accelerometer, WiFi/Bluetooth signals, ambient sound and light, application use, call and text message metadata, keystroke dynamics) and imputed behavioral features (e.g., time spent at home, sleep cycles, level of socialization, routine/anomalies) to predict clinical outcomes including depression and bipolar disorder symptoms, suicide risk, psychosis relapse, and depression treatment response [58–67]. In the context of neurodegenerative disorders, several sensors - particularly keystroke dynamics and phone/battery use metrics - have shown associations with neuropsychological test performance, diagnostic severity, and even gray matter volume in clinical cohorts including those with MCI, AD, frontotemporal dementia, and multiple sclerosis [68–74].

Basic questions of feasibility, acceptability, and ethical considerations related to data privacy are important to weigh when considering the highly sensitive nature of digital phenotyping data in vulnerable populations [75–77]. Few if any existing studies have proactively addressed these questions in the context of older adults with cognitive decline [62]. Further, many existing studies have used exploratory approaches without *a priori* hypotheses [78]. As described by Hackett and Giovannetti (2022) and Leaning et al. (2024), some of the many interpretive and logistical challenges of digital phenotyping can be mitigated with conceptual models and clinically-informed features to provide context to results and improve reproducibility [7,63].

In this manuscript we present findings from a proof-of-concept study evaluating a smartphone digital phenotyping protocol to assess cognition, everyday function, and mood in a cohort of older adults with and without cognitive decline. Here we focus on smartphone-derived GPS data as the digital phenotyping sensor of interest. Our study design and analytic approach were informed by a conceptual framework proposed by Hackett & Giovannetti (2022) based on established trends in the cognitive neuroscience, neuropsychology, neurology, and computer science literature [7]. The conceptual framework (i.e., the Variability in Everyday Behavior [VIBE] model) posits that pathological cognitive decline is accompanied by a reduction in everyday activities, worsening mood, and lower scores on standardized neuropsychological measures. These declining mean-level trends occur alongside increases in intraindividual variability (IIV) on measures of cognition and everyday function as individuals become more inefficient and work to compensate for underlying disease progression [7]. Trends of decreasing *level* of everyday activity and increasing *variability (IIV)* in everyday activities may be indexed by passively obtained smartphone data such as GPS trajectories, hence the focus of the present manuscript.

The primary aims of the current study were twofold; 1) to examine the feasibility and acceptability of a digital phenotyping protocol among older adult smartphone users, and 2) to examine associations between passively obtained GPS movement trajectories collected over a 1-month study period and traditional validated measures of cognition, everyday functioning, mood, and mobility habits collected at baseline. Data were collected using the Learn, Assess, Manage, and Prevent (LAMP) platform, an open-source platform for research and clinical use, via the mindLAMP app [79–81]. Data for the second aim were analyzed and interpreted according to *a priori* hypotheses based on our conceptual framework (i.e., the VIBE model). We predicted that GPS mean-level activity metrics would show a positive linear relation with global cognition, function, and mood, whereas GPS IIV metrics would show a negative relation with cognition, function, and mood[2]. We also predicted a positive linear relation between GPS mean-level activity metrics and objective and subjective measures of gait speed, life space and community participation that would reflect concurrent validity. The study also included two exploratory aims: i) to examine relations between GPS features and participant intrinsic and extrinsic factors (e.g., sociodemographic, contextual) to inform selection of covariates and/or moderating variables in future studies; and ii) to explore whether *patterns* of mobility – rather than absolute amounts of mobility – also relate to validators. Overall, our results provide preliminary support for the feasibility, acceptability, and validity of digital phenotyping in older adults, along with key insights that can be used to inform future studies.

[2] Hypotheses from the VIBE model were adjusted to consider trends along the continuum from healthy cognition to MCI only (excluding trends from MCI to dementia), given the current study sample only included 1 participant with dementia.

Methods

Recruitment

Participants ages 60+ with healthy cognition, diagnoses of MCI or mild AD were recruited from specialty dementia clinics and the community within the Philadelphia region beginning in January 2022. Eligibility was determined by a phone screen and confirmed by measures administered during the initial testing session. General inclusion/exclusion criteria for all participants included a) age 60 or older; b) fluent in English; c) existing smartphone user (iOS or Android; meeting minimum software version compatibility) for at least 1 year prior to joining the study; d) Wi-Fi connectivity at home; e) phone use on at least a daily basis; f) no plans to purchase or switch to a new smartphone over the next 4 weeks; g) availability of an informant reporter who has knowledge of the participant's daily functioning; h) no history of severe psychiatric or nervous system disorders (other than dementia); i) no current metabolic or systemic disorders; j) no severe sensory or motor deficits precluding smartphone use; k) no intellectual disability; and l) no scheduled surgery or major travel over the 4-week study period. Participants with self- or clinician-reported diagnoses of healthy cognition, MCI, or mild AD were re-evaluated during a baseline visit (as described below) and Jak/Bondi neuropsychological actuarial criteria were used to confirm diagnostic group membership[82]. Follow-up consensus diagnosis was used to account for atypical clinical factors that may impact the accuracy of actuarial diagnosis (e.g., English as a second language [ESL], co-occurring mood/psychiatric concerns).

Study informants were also recruited for each participant. General inclusion/exclusion criteria for all informants included a) 18 years of age or older; b) fluent in English; c) cognitively healthy with no diagnosis of dementia, MCI, or other neurologic/psychiatric disorder; d) available and willing to complete study questionnaires in person, by phone, or online; e) has at least weekly contact with the participant; and f) reports that they are knowledgeable of the participant's daily functioning and smartphone use.

Study Procedures

Study Timeline

Participants meeting eligibility criteria were scheduled for an in-person visit (Session 1) and were enrolled in the study for approximately 4 weeks. As outlined in Figure 1, participants and informants completed two study visits, separated by 4 weeks. Session 1 lasted 2 to 4 hours and included a detailed review of study procedures, informed consent, comprehension of consent quiz, cognitive testing, questionnaires, and configuration of the study app (mindLAMP) on participants' personal smartphones (i.e., downloading mindLAMP, logging in using participant-specific secure credentials, and enabling continuous location services). At the completion of Session 1 participants were instructed to resume their daily routines for 4 weeks. They were asked not to enter low battery or airplane mode, and not to log out of the mindLAMP app which would impact GPS data quality. During the 4-week study period, mindLAMP passively and securely collected GPS data without user engagement. At the end of the study period, participants completed Session 2 (in-person or remotely) which included debriefing questionnaires and compensation. The mindLAMP app was deleted from participants' smartphones at Session 2, halting data collection.

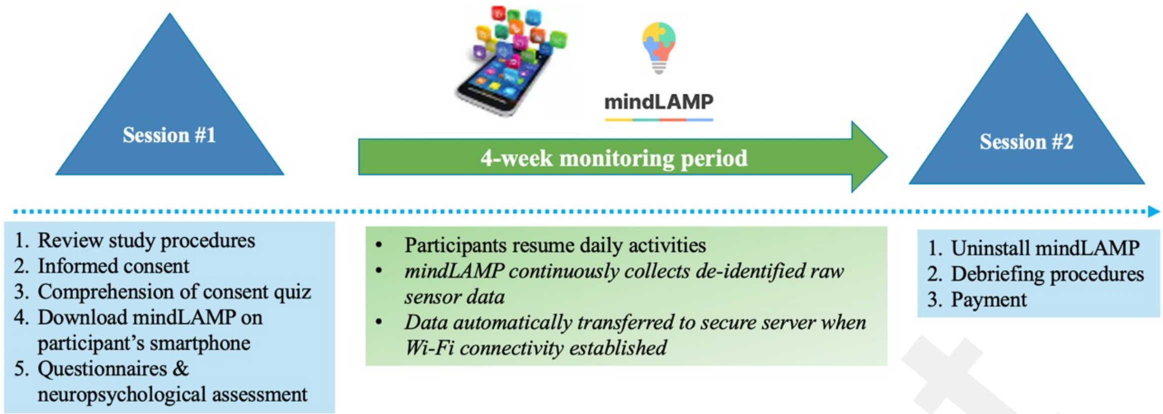


Figure 1. Study Timeline.

Privacy/Security Safeguards

All aspects of the study protocol received ethics approval from the Institutional Review Board (IRB) at Temple University. The LAMP platform was used for collection of digital phenotyping (GPS) data via the mindLAMP smartphone app which is available for free on the App Store and Google Play. LAMP is an open-source platform for research and clinical use developed by the Division of Digital Psychiatry at Beth Israel Deaconess Medical Center[80]. It uses industry standard encryption protocols to render information collected from smartphones unidentifiable, and transmits data to a secure cloud database via HIPAA-compliant methods whenever WiFi connectivity is established. Participants logged in to mindLAMP via a randomly generated 10-digit "UID" generated within the LAMP platform by the study coordinator, which was only used for collection of digital phenotyping data. Therefore, no personally identifiable information is associated with the data collected by mindLAMP. Participants' 6-digit study ID, used for all other clinical data collected within the lab at Temple University, is linked to their mindLAMP UID on a university-approved secure research database only (REDCap).

The above security and privacy information was thoroughly reviewed with participants during informed consent at Session 1. This process included review of written and visual handouts depicting privacy safeguards, examples of the scope of data collected, and results of other published studies that used mindLAMP [83]. Participants could choose to end study participation and have their study data deleted at any time. After reviewing the consent form, a 10-item comprehension of consent quiz with yes/no response options was administered to ensure participants fully understood the information outlined in the consent form. This quiz covered details including the purpose of mindLAMP (e.g., "This study requires downloading the mindLAMP app which collects information from my smartphone sensors"), possible risks/benefits to study participation such as potential battery drain (e.g., "It is possible that I will notice a reduction in my phone's battery life while mindLAMP is running on my phone"), and data security and encryption methods (e.g., "The mindLAMP app uses a secure encryption system called 'hashing' to make all information that it collects unidentifiable and untraceable"). See Figure S1 in Multimedia Appendix 1 for the complete comprehension of consent quiz. Incorrect items were further reviewed until comprehension was established. These procedures were informed by the Digital Health Checklist and other materials from the ReCODE health team, which encourages digital health researchers to proactively identify gaps in the communication of study risks, benefits, and privacy/security details [75,76].

Backend Technical Implementation

Data collection and storage were supported by a self-deployed version of the open-source LAMP Platform. A Temple University-approved secure cloud server (1 Terabyte capacity) was purchased and configured with the LAMP Application Programming Interface (API) to enable research participants to connect to the LAMP platform and access the mindLAMP app. Study data were written to our instance of the mindLAMP database (i.e., our copy of the LAMP platform located on a study-specific cloud server) via CouchDB, ensuring a standardized data format consistent with other studies using the LAMP platform [81]. The functionality of the mindLAMP app itself is continually maintained by the team at mindLAMP at the Beth Israel Deaconess Division of Digital Psychiatry. Ongoing security monitoring and backup of study data were maintained by Temple University IT. To monitor unexpected periods of missing data due to potential technical issues (i.e., phone powered off, mindLAMP log out, permissions reset), we created a code to automate an email alert to the study team (Cronjob) when there were >3 days of missing sensor data; in these instances the email message included the UID and the corresponding missing sensor, and a member of the study team promptly reached out to the participant to troubleshoot.

Measures

Feasibility and Acceptability

To assess feasibility, we tracked the number of participants who completed the 4-week study period after providing consent and completing Session 1, and those who requested to withdraw for any reason. Feasibility was also operationalized by performance on the comprehension of consent quiz, which demonstrates participants' ability to comprehend complex technical information specific to digital phenotyping studies. To gauge acceptability of the informed consent and overall study procedures, we administered a debriefing survey at Session 2 after participants had completed the full study period. Participants were asked to rate their level of satisfaction with the explanation of the study procedures at Session 1 on a scale of 4 (Very satisfied – all components of the study were clearly explained) to 1 (Very unsatisfied – all components of the study were poorly explained).

Participants were also asked if they experienced any major difficulties with their phones, if they experienced any feelings of discomfort or paranoia due to the study application running on their smartphone, or if there were any major changes in how they used their smartphone during the study period (Yes/No). If they answered "Yes" to changes in smartphone use, participants selected all that apply from options including "I used my phone less/more overall," "I charged my phone less/more," "I carried my phone with me less/more," and "Other." Finally, troubleshooting contacts between study staff and participants were tracked and are reported as part of feasibility findings.

Validators

Validation measures administered at Session 1 included neuropsychological tests used widely in the clinical diagnosis of MCI and dementia, self- and informant-report measures of cognitive decline and everyday functioning, questionnaires pertaining to mood, and measures of gait speed, geospatial life space and community participation. These measures have demonstrated strong psychometric properties and were therefore used together as validation comparisons against digital phenotyping data. More details are provided below.

Neuropsychological Tests

The neuropsychological test battery included the Hopkins Reading Test[84] as an estimate of premorbid intellectual ability (IQ); the Mini-Mental State Exam (MMSE [85]) as a global cognitive screener; and tests of attention (Trail Making Test-Part A[86], Wechsler Memory Scale-Revised Digit Span Forward [87]); processing speed (Salthouse Letter and Pattern Comparison [88]); executive function (Trail Making Test-Part B[86], Wechsler Memory Scale-Revised Digit Span Backward [87]); episodic memory (Hopkins Verbal Learning Test-Revised Delayed Recall [89], Brief Visuospatial Memory Test-Revised Delayed Recall [90]); and language (Animal Fluency [91], Boston Naming Test-30 item [92]). Raw scores of each test were transformed into demographically corrected T-scores using the Calibrated Neuropsychological Normative System (CNNS [91]) adjusting for age, sex, education, and estimated premorbid IQ (i.e., Hopkins Reading Test score), which enabled more accurate estimation of cognitive ability within our diverse sample. An average T-score was computed for the two tests within each cognitive domain to generate composite scores for attention, processing speed, executive function, delayed memory recall, and language abilities to streamline presentation of results.

Reported Cognitive Decline and Everyday Functioning

Self-reported cognitive decline was collected using the Everyday Cognition Scale-Short Form (ECog [93]). Self- or informant-reported everyday functioning was captured with the Functional Activities Questionnaire (FAQ [94]). Self-reported FAQ was used for participants with healthy cognition; informant-reported FAQ was used for participants with MCI or dementia. Higher scores on these measures indicate more cognitive decline and more functional impairment, respectively.

Mood

Mood symptoms were indexed using the 15-item Geriatric Depression Scale (GDS [95]), a widely used self-report measure of depressive symptoms among older adults which requires participants to indicate whether they experience a list of common depression symptoms in a “Yes/No” format. Raw scores were transformed into demographically corrected T-scores using the CNNS as above. Higher T-scores indicate higher levels of depression.

Gait Speed, Life Space and Community Participation

The Timed Up and Go Test (TUG) was administered at Session 1 as an objective measure of gait speed. This task measures the time it takes to rise from a chair, walk 10 feet, turn, walk back to the chair and sit down. It is widely used to examine balance, functional mobility, gait speed and fall risk in older adults [96,97]. Participants also completed the University of Alabama at Birmingham Life Space Assessment (LSA), a self-report measure of mobility for community-dwelling older adults [98]. It captures the level of independence and spatial extent of a person’s life over the preceding month, and has shown strong associations with mobility within the home and community and with performance of activities of daily living [99]. Constricted life space has also been associated with risk for MCI and dementia [100,101]. The Australian Community Participation Questionnaire (ACPQ) 15-item version was administered as an additional measure of concurrent validity, and assesses the extent to which someone engages in a range of community activities. Subscales include contact with immediate household, extended family, friends, and neighbors; participating in organized community activities; taking an active interest in current affairs; and religious observance [102]. An index of breadth of participation across the seven domains was derived using a mean-split procedure for each domain, followed by summing these scores to generate an overall index ranging from 1 – 7 (as described by [103]).

Other Participant Features

Demographic

Demographic data included participants' self-reported biological sex assigned at birth, age, race, ethnicity, current living status (alone or with others), current occupational status, educational attainment, and other information related to socioeconomic status (e.g., highest household annual income).

Technology Use

Participants completed a 6-item Habitual Smartphone Behavior subscale [104] to assess smartphone use patterns, providing responses ranging from "strongly agree" to "strongly disagree" to questions such as "Smartphone usage is part of my daily routines." We also asked participants "Do you usually have your phone with you when you leave home?" (Smartphone portability), to which they could reply "Yes- I almost never leave my house without my phone," "In between – I leave my house without my phone about half the time," or "No- I often leave my house without my phone." The Mobile Device Proficiency Questionnaire (MDPQ) was administered as a measure of digital literacy [105]. Participants also indicated their smartphone operating system (Android vs. iOS).

Seasonal/COVID Factors

Dates of study participation were collected and coded as winter/fall/spring/summer to explore potential seasonal effects on mobility habits. Because study participation took place during the COVID-19 pandemic for some participants, we asked about the impact of the pandemic on social participation, routines, and mobility behaviors at the time of study participation. Participants were asked "On a scale of 1-5, how isolated or cut off from family and friends are you feeling due to limited/cancelled social gatherings resulting from COVID-19?"; "On a scale of 1-5, how disruptive has the COVID-19 pandemic been to your daily routines and activities?" and "On a scale of very much limited to very much expanded, how much has the COVID-19 pandemic changed your mobility/your movements outside of the home?"

Self-Reported Health Changes During the Study Period

At the end of the study period during Session 2, participants reported whether there were any changes in their overall health during the study period by responding to a single question on the debriefing survey. Response options included: 1) Yes, significant change; 2) Yes, a little change; or 3) No change. If they endorsed any change in health, they were given the option to elaborate.

Digital Phenotyping (mindLAMP)

Though the mindLAMP app enables collection of a wide array of de-identified passive and active data, the present study focused on passively obtained GPS data. mindLAMP was configured to continuously record the device's GPS coordinates at a maximum frequency of 1 Hz. Raw data outputs include latitude, longitude, altitude, and the coordinates' estimated accuracy. At study completion, these raw data were extracted from the study server and processed into daily summary features (see Table 1) using a publicly available R script developed by Ian Barnett and colleagues [106,107]. GPS data from smartphone devices are prone to large amounts of missing data; therefore, advanced multiple imputation methods were used to account for missingness prior to feature calculation within the aforementioned processing script.

Table 1. Daily GPS features generated from mindLAMP raw data

Category	Abbreviation	Feature Description
Activity	Distance Travelled	Sum of all flight lengths that day (m)
	Radius of gyration	Average distance (m) a person is from their center of mass (average position) on a given day
	Max Diameter	Maximum diameter (m); longest pairwise distance between any two pause locations that occur that day
	Max Distance Home	Maximum distance from home (m); distance between home and farthest pause location from home that day
	Avg Flight Length	Average length of all flights that day (m)
	Avg Flight Duration	Average duration of all flights that day (sec)
Inactivity	Home Time	Time spent at home (min); amount of time that day spent within 200 m of home (the significant location with the largest total amount of time between 9 PM and 6 AM over the course of the study period).
	Probability Paused	Fraction of time a person is stationary (paused) during a day, relative to time spent mobile/in flight.
Routine	Circadian Routine	Physical circadian routine; the fraction of time a person is in the same place (within a 200 m radius) at the same time of each day over the course of the study period. Ranges from 0=completely different routine to 1=identical routine.
	Wknd Circadian Routine	Physical circadian routine weekend/weekday stratified; similar to Circadian Routine except comparisons are stratified by grouping together weekends and grouping together weekdays. Higher scores reflect greater overall routine.
Location Diversity	Significant Location Entropy	Location entropy across a person's significant locations for the day; large values indicate spreading time out across many different locations fairly evenly for that day; small values indicate a concentration at few significant locations.
	Significant Locations Visited	Number of significant locations a person is within 200 m of that day. Determined using K-means on the set of all pause locations with a minimum duration of 10 min (longer pauses given additional weight, no two cluster centers within 400 m of one another).
Other	Minutes Missing	Number of minutes of missing data pre-imputation in a person’s GPS mobility trace that day (min)

Note: See supplementary materials from Barnett and Onnela (2020) for a full description of GPS features and the methods used to calculate each. A flight is defined as a segment of linear movement, pauses are periods of time when a person does not move, and curved movement is approximated by multiple sequential flights.
Abbreviations: min= minutes; m = meters; sec = seconds.

Statistical Analyses

Preliminary Processing

The individual mobility traces derived from GPS data were inspected by examining each participant's de-identified mobility plots averaged across the study period (Figure 2) as well as by ensuring the amount of missing data over the course of follow-up was in the expected/acceptable range. These plots are generated automatically through our processing script and enable visualization of overall mobility habits, amount of time spent at various locations, and time of day. Following visual inspection, daily mobility features (Table 1) were collapsed to monthly overall average and monthly (day-to-day) individual standard deviation (iSD) for each participant to generate estimates of mean mobility and mobility IIV across the study period. For example, 30 days of daily distance travelled estimates for one participant was reduced to an average distance travelled per day, and a standard deviation of daily distance travelled across the 30-day study period. This approach enabled examination of *a priori* hypotheses as an important first step in the validation of GPS data. All variables (GPS and validation measures) with highly skewed distributions were transformed using log(x+c) transformation to reduce the influence of outliers and make parametric analyses more robust. Pearson correlation analyses were used to explore relations among individual GPS features.

Other pre-processing steps were completed in individual cases. Specifically, some participants had unanticipated travel during their scheduled study period (which was an exclusion criterion for validation purposes). In these cases we extended their study duration and excluded days spent traveling from the raw data prior to feature extraction. Three participants reported major unanticipated health changes during the study period (1. eye surgery, 2. major fall with loss of consciousness, 3. fall with head trauma and COVID-19 infection) and were excluded from preliminary validity analyses in the present study because major health events during passive data collection would confound relations with baseline validators.

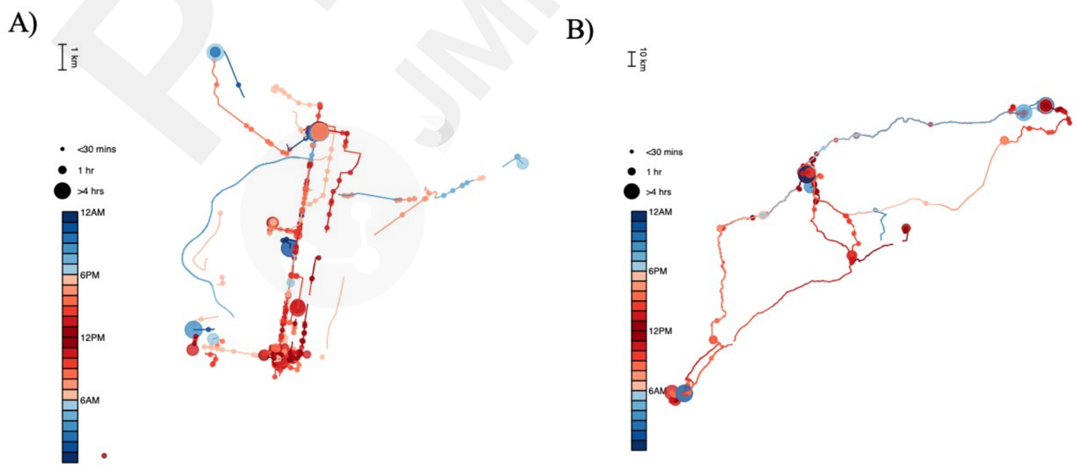


Figure 2. Example monthly mobility plots for two participants: A) 67-year-old female living in northern Philadelphia; B) 70-year-old female living in a suburb of Philadelphia.

Primary Analyses

Pearson or Spearman correlation analyses were used to examine the relations between the mean (monthly average) and IIV (monthly individual standard deviation; iSD) of each mobility feature against validators collected at Session 1 (i.e., cognition, everyday function, mood, mobility). All validity analyses were conducted across the entire sample after excluding 3 participants with major health events (final n for Aim 2 validation analyses = 34). Given the small sample size, we interpreted correlation coefficients with r >.3 as meaningful, regardless of statistical significance. As mentioned, the following *a priori* hypotheses were based on our proposed conceptual framework (i.e., the VIBE model):

- *Cognition/Function:* Average GPS metrics of activity (see Table 1) will show a positive linear relation with measures of cognition and function, whereas GPS IIV for all categories will show a negative relation with cognition and function.
- *Mood:* Average GPS metrics of activity will show a negative linear relation with depression (i.e., greater overall mobility, less depression), whereas GPS IIV for all categories will show a positive relation with depression (i.e., greater mobility variability, more depression).
- *Mobility:* Average GPS metrics of activity will be positively associated with objective and subjective measures of gait speed, life space and community participation.

Exploratory Analyses

A series of exploratory analyses examined relations between GPS metrics and other participant factors such as demographics, technology habits, and smartphone type, as well as environmental factors (e.g., COVID-19 and seasonal impact). Our goal was to better understand what intrinsic and extrinsic factors are associated with GPS trajectories in older adults so that future studies can consider these variables as potential covariates and/or moderators. This is important given the high heterogeneity in individual features and incidental factors that may impact the generalizability of between-group differences in GPS trajectories. Spearman correlation analyses were used for continuous, dichotomous, and ordinal variables and one-way analysis of variance (ANOVA) was used for categorical variables with Welch's test for unequal group variances.

To explore the utility of a GPS composite score, a Gaussian Mixture Modeling (GMM) approach was used to generate a nuanced yet singular representation of mobility trajectories. For each participant's set of daily GPS features (Table 1), we fit a GMM with k=3 mixture distributions to allow for flexible modeling with consideration of differences in the distribution for each person due to various factors (e.g., weekend/weekday differences). The choice of k=3 was determined to maximize clustering complexity subject to our sample size limitations in order to avoid overfitting. This method also allowed us to reduce the full set of GPS features to a single dimension. Next, we created a distance matrix between each pair of participants in the sample by calculating the integral of the squared distance between each GMM density (the larger the distance, the more different the pair's overall mobility patterns). After calculating the distance matrix, we used Multidimensional Scaling (MDS) to extract a 1-dimensional representation of this distance matrix (hereafter termed "MDS1"), akin to a principal component [108]. This MDS1 metric is a relative measure that represents the similarity of overall mobility patterns across participants. For example, if two individuals have similar MDS1 values their overall mobility patterns are similar, whereas two individuals with very different MDS1 values demonstrate different mobility patterns. The MDS1 variable was used in exploratory correlation analyses to identify whether *patterns* of mobility relate to validators, in contrast to total amounts of/variability in individual mobility features.

Results

Participant Characteristics

A total of N=37 individuals participated in our study between April 2022 – January 2024. The full sample was included in analysis of feasibility and acceptability outcomes, whereas a subset (n=34) was included in preliminary validity analyses after excluding those who experienced major unexpected health changes. The n=3 participants excluded due to major health events were on average 71 years old, 66% female, 100% non-Hispanic White, with on average 17 years of education. Two out of three were Android users and all had healthy cognition.

Of the validation subset (n=34), participation was distributed fairly evenly across all four seasons: n=6 in the winter, n=10 in the fall, n=9 in the spring, and n=9 in the summer. Participants ranged in age from 63-85 years old (M=71.6±5.5) and were on average highly educated (M=16.4±2.7 years; range 10-20 years). The majority of participants identified as female (67.6%) and non-Hispanic or Latino/a (97.1%). Fifty-six percent of participants identified as White, 35.3% as Black or African American, and 5.9% as Asian. Most participants lived with others (67.6%) and were retired (79.4%). Most participants were iPhone users (n=26, 76.5%), whereas eight participants were Android users. The majority of participants met diagnostic criteria for healthy cognition (n=28) with a minority meeting criteria for MCI (n=5) or mild dementia (n=1). Scores on the MMSE ranged from 24 to 30 (M=28.3 ±1.3), and according to the FAQ participants on average experienced minimal difficulties with everyday functioning (M=1.5 ± 2.0). Average responses on the GDS revealed low levels of self-reported depression (M=1.5 ± 1.5)[109]. Participant demographic characteristics are detailed in Table 2, and scores on validation measures of cognition, everyday functioning, mood, and mobility are outlined in Table 3.

Table 2. Participant demographics

	Mean (SD)
Age (years)	71.6 (5.5)
Education (years)	16.4 (2.7)
Sex	N (%)
Male	11 (32.4%)
Female	23 (67.6%)
Race	
Black/African American	12 (35.3%)
Asian	2 (5.9%)
White	19 (55.9%)
Not Reported	1 (2.9%)
Ethnicity	
Not Hispanic or Latino/a	33 (97.1%)
Hispanic or Latino/a	1 (2.9%)
English as a Second Language (ESL)	4 (12%)
Living Status	
Live alone	11 (32.4%)
Live with others	23 (67.6%)
Current occupational status	
Full-time employee/volunteer/student	3 (8.8%)
Part-time employee/volunteer/student	4 (11.8%)
Retired	27 (79.4%)
Highest annual household income	
Less than $30,0000	2 (5.8%)
$30,000 to $49,000	4 (11.8%)
$50,000 to $69,000	3 (8.8%)
$70,000 to $89,000	5 (14.7%)
$90,000 to $99,000	3 (8.8%)
$100,000 to $149,000	6 (17.6%)
$150,000 or more	8 (23.5%)
Prefer not to answer	3 (8.8%)
Phone type	
iPhone	26 (76.5%)
Android	8 (23.5%)
Smartphone Portability[a]	
Yes	32 (94.1%)
Half the time	2 (5.9%)
No	0 (0%)
Consensus diagnosis	
Healthy cognition	28 (82.4%)
Mild Cognitive Impairment (MCI)	5 (14.7%)
Mild dementia	1 (2.9%)

Note: Includes data from subset (n=34) included in validation analyses.

[a] Smartphone portability = usually has smartphone when leaves home (no[1]/in between[2]/yes[3])

Table 3. Participant baseline validation measures

Validation Measure	Mean	SD	Minimum – Maximum
Neuropsychological Test (Composite T-scores)			
Global Cognition (MMSE)	52.4	8.1	37 – 67
Attention	52.1	5.8	38 – 66
Processing Speed	54.5	7.9	36 – 74
Executive Function	50.9	7.3	38 – 66
Memory	47.4	11.1	20 – 75
Language	50.0	9.3	28 – 72
Self-and Informant-Reported Functioning			
Functional Activities Questionnaire	1.5	2.0	0 – 6
Everyday Cognition Scale- Short Form	1.3	0.2	1.0 – 1.8
Mood			
Geriatric Depression Scale (raw score)	1.5	1.5	0 – 5
Geriatric Depression Scale (T-score)	52.2	10.4	38 – 76
Mobility			
Timed up and Go Test (TUG; seconds)	11.0	4.0	6 – 29
Life Space Assessment (LSA)	78.5	20.1	36 – 114
Australian Community Participation Questionnaire (ACPQ)	3.8	1.8	1 – 7

Abbreviations: MMSE: Mini Mental Status Exam
Note: Includes data from subset (n=34) included in validation analyses.

Feasibility and Acceptability (Primary Aim 1)

All 37 participants (100%) who began the study completed the 4-week monitoring period and Session 2 without requesting to withdraw. Performance on the comprehension of consent quiz was on average 97% correct. Two participants had an initial score of 80% (eight out of 10 questions correct), seven scored 90%, and 28 scored 100% on their first attempt. The most frequently incorrect item was “Using the mindLAMP app will help improve my cognitive functioning,” to which five participants answered “yes.” Better performance on the comprehension of consent quiz was associated with higher education (r_s=.65, *P*<.001) and differed across racial groups ($F(2,33)=8.4$, $\eta^2=.34$, *p*<.01), with better performance among White versus Black participants according to post-hoc comparisons (*P*=.021). Performance on the quiz was not associated with age, ESL, or cognitive status (all *p*’s >.05).

Satisfaction ratings on the debriefing questionnaire at Session 2 (i.e., responses to the question “How satisfied are you with the study team’s explanation of this study? Did the study team accurately convey what it would be like to participate in this study during the consent process at your first study visit?”) revealed high levels of satisfaction. Specifically, 31 out of 37 participants (84%) reported they were “Very satisfied,” and 6 (16%) reported they were “Satisfied.”

Regarding new issues with or changes to phone use, 34 out of 37 participants (92%) reported they did not experience any new problems using their smartphone during the study period. One participant experienced technical issues that were determined to be unrelated to the study application, a second participant reported that some of their text messages were disrupted (likely unrelated to the study application as we did not collect information from text messages), and a third participant reported that their phone was “a little slow and lack of charge.” Thirty-six out of 37 participants reported they did not experience any feelings of being uncomfortable, suspicious, or paranoid due to the study application running on their smartphone. Although participants were instructed to go about their daily lives and smartphone use as they normally would, 17 out of 37 participants (46%) reported there were major changes in how they used their smartphone during the study period. Specifically, one participant used their phone less overall (3%), 15 charged their phone more (41%), two carried their phone less (5%), and five carried their phone more (14%).

Troubleshooting contacts related to missing GPS data were infrequent. During the entire study period across 37 participants only n=6 incidents were logged affecting n=5 unique participants, with causes including a) Android phone went into “Safe Mode”, b) location permissions reset from “Always” to “Only while using app”, c) outdated version of mindLAMP installed on phone, d) low battery mode enabled, e) mindLAMP app was accidentally deleted. All incidents were promptly resolved by the study coordinator speaking with the participant over the phone to either reconfigure their phone settings or re-download and log into mindLAMP.

Validity Analyses (Primary Aim 2)

Preliminary Analyses

Examining the untransformed GPS data revealed that on average each day, participants spent

1,074 ± 192 minutes at home (i.e., about 18 hours), travelled 42,676 ± 34,694 meters, spent time at 1.56 ± 0.54 unique locations, and had 423 ± 352 minutes of missing GPS data. Correlations among GPS features revealed strong associations among features reflecting activity (distance travelled, radius of gyration, maximum diameter, and maximum distance from home) which were negatively associated with features reflecting inactivity (time spent at home, stationary time) and the two indices of physical circadian routine. Significant locations visited and significant location entropy were intercorrelated suggesting a distinct construct related to location diversity. These associations together support the conceptual GPS feature categories outlined in Table 1 and are used throughout to streamline presentation of results. See Tables S1-S2 in Multimedia Appendix 2 for descriptives of GPS data and intercorrelations among all GPS features.

Relations Between Average Mobility Features and Measures of Cognition, Mood, and Everyday Function

We predicted significant relations between monthly average GPS activity features and baseline neuropsychological measures, mood, and everyday function, such that greater overall mobility would be associated with better performance on neuropsychological tests, less depression, and less reported cognitive and functional decline. Most correlations were in the predicted direction (see Table 4; results for individual neuropsychological tests and function questionnaires are reported in Table S3 in Multimedia Appendix 2). Specifically, greater GPS activity was associated with better scores on the language composite and on individual measures of attention. GPS measures of inactivity and physical circadian routine were associated with lower language scores. Associations between GPS activity features and measures of memory consistently were in the opposite direction, such that greater GPS activity (particularly greater average flight length) associated with worse memory performance. The negative coefficient for the delayed recall memory measure by average flight length was statistically significant (r=-.38, P =.03; see Table S3).

Regarding mood, more inactivity (e.g., more home time), greater routine, and less location diversity were associated with greater depression symptoms (Table 4). Less location diversity was also associated with greater reported functional impairment (r=-.36, P=.039 see Table S3 in Multimedia Appendix 2). Negligible relations were observed between average GPS features and self-reported cognitive decline on the ECog.

Table 4. Bivariate Pearson correlations between monthly average GPS features and neuropsychological measures (T-scores)

GPS Monthly Average Feature	Global Cognition (MMSE)	Attention	Processing Speed	Executive Function	Memory	Language	GDS
Activity							
Distance Travelled	.15	.33	-.01	.07	-.17	.33	-.24
Radius of Gyration	.21	.25	-.03	.11	-.10	.34	-.21
Max Diameter	.19	.33	.01	.10	-.12	.35*	-.20
Max Distance Home	.22	.27	.00	.08	-.10	.43*	-.19
Avg Flight Length	.22	.26	-.05	.03	-.27	.38*	-.33
Avg Flight Duration	-.05	.15	-.10	.07	-.01	.16	.01
Inactivity							
Home Time	-.08	-.11	.22	.16	.17	-.45**	.38*
Probability Paused	-.12	-.28	.13	-.02	.11	-.22	.29
Routine							
Circadian Routine	-.07	-.11	.16	.16	.20	-.51**	.40*
Wknd Circadian Routine	-.08	-.12	.14	.15	.19	-.53**	.38*
Location Diversity							
Sig Location Entropy	.04	.20	-.09	-.06	-.25	.23	-.36*
Sig Locations Visited	-.20	.26	-.02	-.21	-.14	.07	-.35*
Other							
Minutes Missing	.28	.47**	.08	-.01	-.24	.37*	-.39*

Note: Data represent effect size as measured by bivariate Pearson correlation coefficients (*r*), whereby .10, .30, and .50 represent small, moderate, or large effects, respectively. Moderate/large positive coefficients are designated with light/dark green shading, and moderate/large negative coefficients are designated with light/dark blue shading.

All neuropsychological measures reflect T-scores corrected for age, sex, education and estimated premorbid IQ.

Abbreviations: MMSE: Mini Mental Status Exam; GDS: Geriatric Depression Scale – 15 item.

**p*<.05 (2-tailed). ***p*<.01 (2-tailed). See Table S4 in Multimedia Appendix 2 for exact *p*-values.

Relations Between Average GPS Features and Measures of Gait Speed, Life Space and Community Participation

Contrary to our hypotheses, performance on the TUG measure of gait speed was not significantly associated with any of the average GPS features (Table 5). On the other hand, self-reported measures of geospatial life space (LSA) and community participation (ACPQ) were significantly associated with many GPS features. Overall, greater GPS activity was associated with more self-reported life space and more community participation, whereas greater inactivity and physical circadian routine were associated with less geospatial life space and less community participation. Relations with the ACPQ were driven by the domains of extended family and friends (i.e., participants who reported they tend to visit extended family and friends more often demonstrated higher levels of GPS activity and lower physical circadian routine indices; see Table S5 in Multimedia Appendix 2).

Table 5. Bivariate Spearman correlations between average GPS features and measures of gait speed, life space and community participation

GPS Monthly Average Feature	TUG		LSA		ACPQ	
	r	P-value	r	P-value	r	P-value
Activity						
Distance Travelled	-.16	.37	.40	.018	.34	.048
Radius of Gyration	-.17	.35	.52	.002	.39	.023
Max Diameter	-.18	.31	.44	.009	.40	.021
Max Distance Home	-.11	.54	.44	.008	.42	.013
Avg Flight Length	-.05	.77	.23	.18	.29	.10
Avg Flight Duration	.01	.94	.34	.052	-.09	.62
Inactivity						
Home Time	.12	.49	-.42	.014	-.39	.024
Probability Paused	.13	.46	-.48	.004	-.24	.18
Routine						
Circadian Routine	.12	.50	-.45	.007	-.45	.008
Wknd Circadian Routine	.18	.32	-.50	.002	-.48	.004
Location Diversity						
Sig Location Entropy	-.15	.40	.34	.052	.30	.084
Sig Locations Visited	-.18	.31	.29	.093	.21	.232
Other						
Minutes Missing	.17	.34	.16	.35	.32	.067

Note: r represents effect size as measured by bivariate Spearman correlation coefficient, whereby .10, .30, and .50 represent small, moderate, or large effects, respectively. Moderate/large positive coefficients are designated with light/dark green shading, and moderate/large negative coefficients are designated with light/ dark blue shading.

Abbreviations: TUG: Timed Up and Go Test; LSA: Life Space Assessment; ACPQ: Australian Community Participation Questionnaire.

Relations Between GPS Variability and Measures of Cognition, Mood, and Everyday Function

We predicted a negative relation between day-to-day variability in GPS features and baseline measures of cognition, mood, and everyday function, such that greater IIV would be associated with lower neuropsychological test scores, more depression, and more reported cognitive and functional decline. Contrary to our hypotheses, we saw that higher variability in most GPS features (i.e., greater day-to-day iSD in mobility habits) was associated with *better* scores on neuropsychological measures of attention and language, as seen in Table 6 and in Tables S6-S7 in Multimedia Appendix 2. Again, relations with memory were in the opposite direction such that higher GPS IIV associated with worse verbal and visual memory. Regarding relations with mood, greater IIV in location diversity was associated with less depression. This is contrary to expectations but consistent with above results suggesting greater variability in mobility habits is overall beneficial. While GPS IIV was not significantly associated with self-reported cognitive decline or functional impairment, results are directionally consistent such that greater IIV was weakly related to less reported functional decline (see Table S7 in Multimedia Appendix).

Table 6. Bivariate Pearson correlations between monthly IIV GPS features and neuropsychological measures (T-scores)

GPS Monthly IIV Feature	Global Cognition (MMSE)	Attention	Processing Speed	Executive Function	Memory	Language	GDS
Activity							
Distance Travelled	.22	.24	.01	.15	-.08	.37*	-.14
Radius of Gyration	.27	.15	.00	.12	.05	.33	-.07
Max Diameter	.27	.20	.06	.13	.03	.37*	-.06
Max Distance Home	.27	.16	.03	.12	.04	.40*	-.06
Avg Flight Length	.27	.16	.04	.18	-.32	.35*	-.28
Avg Flight Duration	.00	.06	-.08	.10	-.03	.18	.02
Inactivity							
Home Time	.23	.15	-.13	.00	-.18	.64**	-.31
Probability Paused	.17	.27	-.10	.13	-.09	.29	-.10
Routine							
Circadian Routine	.15	.17	.10	.07	-.02	.59**	-.20
Wknd Circadian Routine	.16	.13	-.03	.04	-.02	.55**	-.28
Location Diversity							
Sig Location Entropy	.01	.25	.06	-.09	-.45**	.35*	-.47**
Sig Locations Visited	-.21	.11	-.07	-.33	-.13	.32	-.36*
Other							
Minutes Missing	.29	.06	-.13	.03	.00	.41*	-.22

Note: Data represent effect size as measured by bivariate Pearson correlation coefficients (*r*), whereby .10, .30, and .50 represent small, moderate, or large effects, respectively. Moderate/large positive coefficients are designated with light/dark green shading, and moderate/large negative coefficients are designated with light/dark blue shading. All neuropsychological measures reflect T-scores corrected for age, sex, education and estimated premorbid IQ
Abbreviations: MMSE: Mini Mental Status Exam; GDS: Geriatric Depression Scale – 15 item.
*p<.05 (2-tailed). **p<.01 (2-tailed). See Table S6 in Multimedia Appendix 2 for exact p-values.

Exploratory Analyses

Relations Between GPS Metrics and Participant Intrinsic/Extrinsic Factors

To inform future studies in selection of covariates or moderating variables, we conducted exploratory correlations between GPS average and IIV features and various participant, environmental and contextual factors (see Tables S8-S9 in Multimedia Appendix 2). Average GPS features were unrelated to several sociodemographic factors including age, sex, and cohabitation status (*p's* >.05). Higher education was associated with less location diversity only (*r*=-.35, *P*=.04). On the other hand, lifetime annual income and native language status appeared to be more relevant. Overall, higher lifetime income and native English language were associated with greater levels of GPS activity, less routine, and greater location diversity (.36 ≤ |*r's*| ≤ .55, *p*'s<.05).

Technology factors were largely unrelated to average GPS features. Phone type (iPhone vs. Android) was unrelated to all features except for the overall amount of minutes missing; iPhone users had less missing GPS data than Android users (*r* =-.64, *P*<.001). More habitual smartphone use was associated with less location diversity only (*r*=-.40, *P*=.02). Whether or not participants typically carry their phone when leaving home (smartphone portability) was unrelated to all average GPS features, as was digital literacy as measured by the MDPQ (*p*'s >.05). Regarding the impact of COVID-19, individuals who reported more COVID-related isolation (i.e., due to limited or cancelled social gatherings) demonstrated less location diversity (-.51≤*r's*≤-.42), whereas those who reported a greater expansion in mobility habits due to COVID demonstrated more location diversity (*r*=.34; all *p*'s<.05).

One-way ANOVA of average GPS features were used to explore group differences across categorical demographic variables (race, occupational status, season), with Welch's test for unequal group variances. Group differences for race (Black, Asian, White) were observed only on the minutes missing feature ($F(2,30)=6.32$, $\eta^2=.30$, *P*=.005), with Black participants having greater amounts of missing GPS data than White and Asian participants according to post-hoc comparisons (*P*=.003 and *P*=.023, respectively). A follow-up chi-square test was performed to examine the relationship between race and phone type given prior findings that Android users have greater missing data than iPhones [110]. The relation was significant, such that a greater proportion of Black participants owned Androids, χ^2 (2, N=33) = 6.92, *P*=.031), suggesting that group differences in missing data could be related to phone type.

No group differences in average GPS features were observed according to current occupational status (e.g., retired, working full vs. part-time). On the other hand, several activity features were significantly different across the winter, fall, spring and summer seasons ($4.01 \leq F(3,30) \leq 8.8$, $.18 \leq \eta^2 \leq .28$, *p*'s<.05). According to post-hoc comparisons these differences were driven by significantly greater GPS activity in the summer versus fall months (see Multimedia Appendix 3).

Spearman correlations between GPS IIV features and the participant factors assessed above are detailed in Table S9 in Multimedia Appendix 2. Similar to the results with average GPS metrics, relevant sociodemographic factors appeared to be income and native language status. Higher lifetime income and native English language were associated with more variability in activity, routine, and location diversity features ($.35 \leq |r\text{'s}| \leq .56$, *p*'s<.05). Phone type was more prominently related to IIV measures than average measures; iPhone users showed greater IIV than Android users on most GPS features ($.34 \leq r\text{'s} \leq .52$, *p*'s<.05). Of note, iPhone ownership was also positively correlated with income (*r* =.49, *P*=.005), making it difficult to determine whether higher levels of IIV among iPhone users relates to differences in phone type or to behavioral differences in lifestyle afforded by higher income. Regarding other technology factors, habitual smartphone use and smartphone portability were unrelated to GPS IIV metrics, whereas digital literacy was related to more variability in routine only (*r*=.34, *P*=.047). Individuals who reported more COVID-related social isolation had lower IIV in location diversity (i.e., more consistency in location habits; $.36 \leq |r\text{'s}| \leq .55$), whereas those who reported a greater expansion in mobility habits due to COVID had greater IIV in location diversity (*r*=.37, all *p*'s<.05).

One-way ANOVA of GPS IIV features across categorical demographic variables revealed a significant difference across racial groups on only one measure – IIV for flight length ($F_W(2,18.63)=11.96$, $\eta^2=.12$, *P*<.001). This group difference was driven by lower IIV in Asian participants compared to both Black and White participants according to post-hoc tests. Effects of current occupational status were observed on several IIV metrics ($5.28 \leq F(2,31) \leq 33.90$, $.12 \leq \eta^2 \leq .35$, *p*'s<.05), such that retired participants demonstrated significantly greater IIV in mobility habits compared to those working full- or part-time (see Multimedia Appendix 4). Consistent with results of GPS average metrics, several GPS IIV metrics differed according to season of study participation, with significantly more IIV in activity features observed during summer months versus the fall and spring months ($2.41 \leq F(3,30) \leq 3.78$, $.19 \leq \eta^2 \leq .27$, *p*'s<.05).

Relative Mobility Patterns

A GMM, multidimensional scaling (MDS) approach was used to derive a relative mobility feature (MDS1) including all GPS features simultaneously. This allowed us to explore how similarity in mobility patterns, as reflected through a one-dimensional value, related to validators of cognition, function, and mood. Relative mobility patterns were moderately associated with a cognitive measure of language (|*r*|=.36, *P*=.037) and a global cognitive screener (|*r*|=.31, *P*=.078; see Table S10 in Multimedia Appendix 2). Relative mobility patterns were also associated with depression (|*r*|=.39, *P*=.021), and community participation (|*r*|=.51, *P*=.002). While the MDS1 feature represents a linear contrast combining multiple individual GPS features that excels at differentiating mobility patterns across the sample, this comes at the cost of interpretability of the associations related to MDS1 (e.g., positive versus negative correlation coefficients are not meaningful); however, these results provide additional converging evidence that individuals with similar mobility profiles have similar levels of underlying cognitive ability, depression, and community participation.

Discussion

This study investigated a 4-week smartphone digital phenotyping protocol as a novel method to assess everyday cognition, function and mood in a cohort of 37 older adults. Our preliminary results suggest that theoretically-informed digital phenotypes of mobility are feasibly captured from older adults' personal smartphones and associate with clinically relevant data pertinent to cognitive aging and AD/ADRD. Findings and implications provide key insights to inform the design and interpretation of future studies using this method in larger, more diverse cohorts.

One of our primary aims was to examine the feasibility and acceptability of smartphone digital phenotyping among older adults. All participants completed the 4-week monitoring period without dropping out, and 97% reported no feelings of discomfort during debrief procedures. On the other hand, almost half reported changes in how they used their smartphone, with 41% endorsing charging their phone more frequently. Battery drain was a communicated risk and is common in high frequency continuous data collection including GPS [111]. Increased phone charging behavior may limit the naturalistic aspect of this approach and will be important to address in future designs – particularly if battery power is considered as a digital biomarker in and of itself [73]. Future studies should explore whether lower sampling frequencies are sufficient, as this can be adjusted via the LAMP platform and would lead to less battery drain. It is also likely that advances in smartphone battery lifespan will ultimately circumvent this issue; in the interim, participants may be provided with portable batteries for daily outings.

We were also interested in how participants would respond to and comprehend complex technical details of this study. Participants demonstrated good comprehension of study procedures as demonstrated by an average score of 97% on the comprehension of consent quiz, and all reported satisfaction with the study team's initial review of study procedures. These findings are encouraging given the dearth of studies investigating older adults' attitudes and concerns about passive sensing technologies [75], and suggest that our upfront efforts to enhance privacy, security, transparency, and comprehension were effective [76]. Despite high overall comprehension, however, we observed that higher education and identifying as White vs. Black was associated with better performance on the comprehension of consent quiz – suggesting the language used in our materials may not be culture-free and should be revised using co-design or focus group approaches with increasingly diverse perspectives. Importantly, the most frequently incorrect item on the quiz pertained to potential benefits of study participation ("Using the mindLAMP app will help improve my cognitive functioning," yes/no). Clear communication of potential benefits in research is a core ethical requirement and future study materials should clarify expectations for potential benefits to facilitate trust, particularly among historically marginalized groups [112–115].

Regarding our second primary aim of establishing preliminary validity, we observed converging support that unobtrusively obtained movement trajectories from smartphones are related to established clinically-relevant variables including cognition, function, and mood. This is not surprising but is highly encouraging. Movement trajectories reflect many facets of everyday cognition including the ability and motivation to travel outside the home, the degree of someone's spatial routine, how many unique locations someone can visit, and the total amount and duration of movement – behaviors that require abundant cognitive and psychological resources. Here, individuals with better cognition, less functional impairment, and less depression did in fact demonstrate significantly greater overall mobility according to several GPS features. Specifically, they travelled farther, spent less time at home, and had greater diversity in the locations they visited. These findings are consistent with our hypotheses and with prior studies showing greater physical activity, less time at home, greater life space (the extent of movement through the environment during daily functioning), and engagement in a variety of activities (environmental complexity) are associated with better cognition, less depression, and reduced risk for MCI and dementia/ADRD [63,100,101,116–124].

Inconsistent with our hypotheses, individuals with better cognition and less depression also demonstrated *greater* day-to-day variability in GPS features and *less* physical circadian routine. Regarding depression, it is conceptually reasonable that more varied mobility habits could be protective against depression, and this has been demonstrated in at least one study of younger adults where lower location diversity associated with more depression symptoms [125]. With respect to the negative association between cognition and routine, it is possible that the older adults in our cohort with more cognitive difficulties intentionally engaged in more predictable and less demanding daily activities to compensate for underlying mild difficulties, leading to more consistent physical circadian routines and lower day-to-day variability in mobility habits. This is a pattern identified in the literature and is often recommended as an intervention in clinical practice [126–130].

The VIBE model predictions regarding variability were informed by observations that individuals with cognitive impairment demonstrate increased IIV compared to healthy controls while performing single, standardized tasks in the clinic or laboratory where task demands are the same for everyone. Greater IIV on constrained tasks reflects an inability to maintain consistent levels of performance [7,131,132]. The present study instead involved unconstrained mobility habits, which individuals may modify to compensate for cognitive difficulties. Therefore, the observation of GPS IIV as indicative of positive rather than negative outcomes may be specific to unconstrained geospatial routines, or to our relatively functionally healthy sample. We may still observe that greater IIV associates with worse outcomes when considering more fine-grained digital biomarkers (e.g., diurnal phone use patterns, accelerometer-based sleep estimates) that are relatively more constrained, as has been identified in other studies examining IIV in gait speed, medication-taking routine, and computer use [133–135]. We may also see that greater IIV in broad everyday behaviors is a marker of resilience early on, but becomes maladaptive in later stages of neurocognitive decline. Larger samples with greater heterogeneity in cognitive ability are needed to answer this question, in addition to longitudinal designs that would enable monitoring within-person IIV trends over time. In the meantime, our relatively cognitively healthy sample provides a view into how variability behaves in early stages and unconstrained settings. Consistent with past reports of significant task and timescale effects on IIV [136], IIV in continuous mobility trajectories may be mechanistically distinct from IIV in constrained settings.

Concurrent validity was supported by strong and significant associations between mean-level GPS features and baseline scores on the LSA and the ACPQ – self-report measures of geospatial life space and community activity participation, respectively. These two constructs are highly relevant in the context of aging and AD/ADRD and represent key measures of risk. The LSA measures the extent, frequency, and independence of movement within and outside the home. Constricted life-space has been associated with increased risk of AD, MCI, and cognitive decline in racially diverse groups and may represent an early functional marker of prodromal decline as individuals compensate for early subtle difficulties and limit their range of movement/activity complexity [100,101,126]. Social engagement – particularly leaving the home to visit extended friends and family – protects against social isolation, promotes cognitive reserve, and represents a complex activity requiring cognitive flexibility [137–141]. Thus, the ability to unobtrusively and longitudinally measure these key risk and resilience factors without the burden or bias of self- or informant-report represents a noteworthy application for smartphone-derived mobility trajectories. Minimal associations were identified between GPS features and the TUG measure of gait speed. It is possible that low heterogeneity in TUG scores or small variations in administration of the TUG played a role. It is also likely that gait speed and coordination are lower-level features of mobility that are independent of broader mobility habits, at least within this sample of functionally independent older adults.

We examined many individual and contextual influences on mobility features to inform future studies in selecting covariates or moderating factors. This is a critical open question in the field of digital phenotyping [63], and our preliminary results provide important insights about which sociodemographic, contextual, and technology factors should be considered when interpreting this data. Age, education, sex, race, and cohabitation status appeared to be minimally associated with most mean-level and variability metrics, providing partial support for the objectivity of digital phenotyping features. Nonetheless, other sociodemographic factors including higher lifetime income and native English language were moderately associated with several mobility metrics that appear to be advantageous (i.e., greater activity and greater IIV). This could reflect an association between social determinants of health and access to transportation, opportunities for socialization outside of the home, or an ability to engage in spontaneous activities – and may suggest mechanisms for income and acculturation as resilience factors.

Other extrinsic factors with relatively clear and interpretable effects included season, COVID-related effects, and occupational status. Participants demonstrated greater activity, less home time, and more mobility variability in summer months which is reasonable given increased leisure activities that typically occur in the summer. Seasonal differences are therefore relevant if interpreting data in pre-post designs, suggesting investigators should aim to control for season or re-evaluate during the same season if possible. Participants reporting more COVID-related isolation visited fewer locations whereas those reporting more COVID-related mobility expansion visited more. Individuals who were retired demonstrated more variable mobility habits than those working full- or part-time, which aligns with differential degrees of consistency depending on occupational status. In addition to shedding light on which factors are relevant when interpreting digital phenotyping data, these associations provide additional validation for the mobility features in the present study.

Phone type was unrelated to all mean-level GPS metrics except for the number of minutes of missing data, which was higher for Androids. This finding is consistent with a previous study that identified lower rates of missing GPS data among iOS users [110], and overall bodes well for the generalizability of mean-level metrics across different phone types and operating systems. Missing data was also higher among Black participants, which could be due to a higher proportion of Android ownership among Black participants in our study. Thus, controlling for phone type may be important when interpreting the minutes of missing data feature. iPhone users also demonstrated more variability in most GPS features. Given iPhone ownership was positively associated with income, it is difficult to interpret whether increased variability among iPhone users is related to operating system factors or to aspects of resilience associated with income. Future studies with larger sample sizes should work to clarify these questions.

In considering the neuropsychological correlates of our GPS features, relations with cognition were strongest and most consistent for measures of language. This was somewhat unexpected given language abilities are typically not as critical to the completion of everyday activities compared to domains such as executive functioning [142,143]. Nonetheless, intact language functioning (specifically semantic access) relies on left temporal and prefrontal integrity and connectivity, which are highly relevant neuroanatomical regions in the pathogenesis of AD. Interestingly, relations between GPS features and measures of memory were in the opposite direction compared to other cognitive domains; longer average flight length and more variability in flight length were associated with worse delayed memory performance. This finding may reflect inefficiencies in planning resulting in back-tracking on a particular trip (e.g., to the grocery store) due to forgotten items. In general, differential relations between GPS features and measures of memory and language – two cognitive domains implicated in AD pathology - suggests relations between GPS features and cognitive abilities may not be as straightforward as our model predicted. Future studies should continue to investigate whether mobility phenotypes are uniquely related to specific neuropsychological and neuroanatomical correlates, rather than focus on a global cognitive composite. This may involve developing even more fine-grained GPS features to capture distinct functional difficulties in everyday life (e.g., backtracking, repetitive motions).

Limitations and Strengths

Our study has several limitations worth noting. Our relatively small sample size of 37 limits generalizability and precluded investigation of diagnostic group differences. Digital phenotyping is a new research area with relatively few established standards, yet preliminary validation of digital phenotypes has been reported in samples of under 50 participants [58,70,144,145] for at or under 30 days of data collection [70,146–148]. Another major limitation is the restricted diversity of our participant cohort, in terms of cognition/diagnostic group, demographics (ethnicity, education, socioeconomic status), and severity of depressive symptoms. Further, all participants lived within driving distance of Temple University and therefore reflect an urban/suburban cohort. In addition, individual data processing exclusions were required to account for unanticipated travel and health events. Given our aim of establishing preliminary validity, there was a need to ensure participants met strict inclusion criteria which included relative stability in their health and physical location. This level of control and oversight may be unrealistic in larger, longitudinal studies and may need to be replaced by advanced statistical methods in the future [149]. Finally, it is worth mentioning that mobility traces from smartphone

GPS sensors are a proxy for actual movement trajectories and depend upon the participant having their smartphone on their body, requiring multiple imputation and inference for missing GPS data prior to feature calculation during preprocessing [150]. Sensors worn on the body may provide more accurate and reliable measures of mobility patterns, but present other drawbacks (e.g., may be perceived as more intrusive by participants, may disrupt existing habits, typically costly).

The present study has several strengths. Our cohort was well characterized compared to many prior digital phenotyping studies, including a comprehensive neuropsychological battery with 10 individual neurocognitive tests, objective and subjective validation measures, and application of a combined actuarial and consensus diagnosis criteria to accurately classify participants. Our study design and hypotheses were theoretically informed by a conceptual model, which improves the interpretability and replicability of our findings. Toward this aim, we examined a range of interpretable mobility features (i.e., 13 monthly average features, 13 monthly iSD features) prior to reducing features into a principal component. Although many GPS features were intercorrelated, the presence of differential and unique correlations suggests individual features may be useful in clarifying specific behaviors and should be preserved as we continue to learn more about what these features signal. With regard to our technical protocol, use of an open-source platform and a publicly available data processing script facilitates replicability which is key to ongoing validation efforts [151]. Attempts at minimizing missing data using an automated checking script represent another strength and should be incorporated in future studies given the impact of missing data on subsequent findings [63]. Finally, our device-agnostic approach leveraging personally-owned smartphones versus a study issued device represents a notable strength as it promotes the naturalistic, unobtrusive nature of this method and affords increased scalability.

Future Directions

Future studies with larger and more diverse cohorts will be critical to replicate the present findings and address open questions. Given the lack of accessible and unbiased diagnostics available to individuals from low-income and marginalized groups [50–52,152], increased diversity in terms of race, ethnicity, educational attainment, and socioeconomic status is a priority for subsequent studies. Next steps will also investigate relations between validation measures and other digital phenotyping sensors (accelerometer, device state, steps) to extend the present findings and test our theoretical framework in behavioral features other than mobility trajectories. Machine learning approaches integrating multiple sensor streams may be useful to determine clinically useful digital phenotypes, thereby reducing the analytic burden and narrowing the focus on clinically relevant, non-redundant features. Additional open questions include the test-retest reliability of digital phenotypes; the incremental utility of ecological momentary assessment (EMA) which can be provide context to passive data [7,46,153]; determining the minimum necessary sensor sampling rate and duration to reduce battery drain; and evaluating within-day variability, diurnal patterns, and time of day effects on GPS features [26,46,72,134,154].

As the field of digital phenotyping continues to evolve, it is likely that longitudinal monitoring of individual anomalies will be the ideal approach to capture individually-relevant changes signaling risk of subsequent decline, rather than population- or group-based approaches [7,60,151,155]. However, with longitudinal designs and increased sample sizes it will be important to consider whether the data storage, processing, and interpretation burdens justify breaking from the status quo. Workflows that enhance the efficiency and scalability of this method will be critical, not just for research participants and patients but also for those collecting and interpreting the data [7,77,111].

Conclusions

As the population of older adults continues to rise, efforts to identify new tools to detect risk for future cognitive decline and measure treatment response are critical, particularly as new pharmacologic interventions gain approval. Current gold-standard methods are costly, burdensome, not widely accessible and/or not part of routine clinical care. Measures that are sensitive to cognitive decline that can be passively and affordably integrated into everyday life without burden, disruption or self-report bias are needed to serve as a first-line approach. Our study demonstrates that unobtrusively obtained GPS movement trajectories from personal smartphones may be one such first-line approach, enabling clinicians and/or researchers to efficiently assess cognitive status, mood, and dementia risk on a broader scale. Individuals with at-risk data profiles may ultimately be referred for more comprehensive evaluation and directed toward appropriate intervention or research settings, leading to cost savings, reduced burden, and faster access to care. Much work remains before we determine how smartphone digital phenotyping can be integrated into our current healthcare system; however, preliminary results suggest that it is a worthwhile endeavor and serve to inform follow-up studies that are necessary to answer important outstanding questions.

Acknowledgements

The research reported in this publication was supported by the National Institute on Aging (awards F31AG069444 and T32AG066598 [KH]; R01AG062503 and R21AG066771 [TG]; R01MH116884 [IB]). We would also like to acknowledge Emma Pinsky, Sophia Holmqvist, and Marina Kaplan for assistance with data collection and entry.

Authors' Contributions

KH, TG, IB contributed to conceptualization. IB, SX, LP, KH, TG contributed to the methodology. KH, MM contributed to data collection and implementation. KH, SX, IB, TG contributed to formal analysis. KH and TG contributed to writing (original draft preparation). KH, SX, MM, LP, IB, TG contributed to writing (review and editing. TG contributed to project administration and funding acquisition. All authors read and agreed to the version of the manuscript intended for publication.

Conflicts of Interest

The authors report no conflicts of interest. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Abbreviations

ACPQ: Australian Community Participation Questionnaire

AD: Alzheimer's disease

AD/ADRD: Alzheimer's disease and related dementias

CNNS: Calibrated Neuropsychological Normative System

ECog: Everyday Cognition Scale- Short Form

FAQ: Functional Activities Questionnaire

GPS: Global positioning system

HVLT-DR: Hopkins Verbal Learning Test – Revised Delayed Recall

IIV: Intraindividual variability

LSA: Life Space Assessment

MCI: Mild Cognitive Impairment

MDPQ: Mobile Device Proficiency Questionnaire

MDS1: Relative mobility feature derived from Multidimensional Scaling approach

MMSE: Mini Mental State Exam

TUG: Timed Up and Go Test

VIBE: Variability in Everyday Behavior Model

References

1. Branca C, Oddo S. Paving the Way for New Clinical Trials for Alzheimer's Disease. Biological Psychiatry. 2017;81(2):88-89. doi:10.1016/j.biopsych.2016.10.016

2. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia: The Journal of the Alzheimer's Association. 2011;7(3):280-292. doi:10.1016/j.jalz.2011.03.003

3. Sperling RA, Rentz DM, Johnson KA, et al. The A4 Study: Stopping AD Before Symptoms Begin? Science Translational Medicine. 2014;6(228):228fs13-228fs13. doi:10.1126/scitranslmed.3007941

4. Petersen RC, Aisen PS, Andrews JS, et al. Expectations and clinical meaningfulness of randomized controlled trials. Alzheimer's & Dementia. 2023;n/a(n/a). doi:10.1002/alz.12959

5. Cummings J, Apostolova L, Rabinovici GD, et al. Lecanemab: Appropriate Use Recommendations. J Prev Alzheimers Dis. 2023;10(3):362-377. doi:10.14283/jpad.2023.30

6. Giovannetti T, Hackett K, Tassoni M, Mis R, Simone S. APA Handbook of Neuropsychology. In: Assessing and Predicting Everyday Function. Vol 2. American Psychological Association; 2022. https://www.apa.org/pubs/books/apa-handbook-of-neuropsychology?tab=2

7. Hackett K, Giovannetti T. Capturing Cognitive Aging in Vivo: Application of a Neuropsychological Framework for Emerging Digital Tools. JMIR Aging. 2022;5(3):e38130. doi:10.2196/38130

8. Jutten RJ, Papp KV, Hendrix S, et al. Why a clinical trial is as good as its outcome measure: A framework for the selection and use of cognitive outcome measures for clinical trials of Alzheimer's disease. Alzheimer's & Dementia. 2023;19(2):708-720. doi:10.1002/alz.12773

9. Rabin LA, Paolillo E, Barr WB. Stability in Test-Usage Practices of Clinical Neuropsychologists in the United States and Canada Over a 10-Year Period: A Follow-Up Survey of INS and NAN Members. Archives of Clinical Neuropsychology. 2016;31(3):206-230. doi:10.1093/arclin/acw007

10. Teunissen CE, Verberk IMW, Thijssen EH, et al. Blood-based biomarkers for

Alzheimer’s disease: towards clinical implementation. The Lancet Neurology. 2022;21(1):66-77. doi:10.1016/S1474-4422(21)00361-6

11. Brickman AM, Cabo R, Manly JJ. Ethical Issues in Cross-Cultural Neuropsychology. Applied Neuropsychology. 2006;13(2):91-100. doi:10.1207/s15324826an1302_4

12. Byrd DA, Rivera-Mindt MG. Neuropsychology’s race problem does not begin or end with demographically adjusted norms. Nat Rev Neurol. 2022;18(3):125-126. doi:10.1038/s41582-021-00607-4

13. Chan KCG, Barnes LL, Saykin AJ, et al. Racial-ethnic differences in baseline and longitudinal change in neuropsychological test scores in the NACC Uniform Data Set 3.0. Alzheimer’s & Dementia. 2021;17(S6):e054653. doi:10.1002/alz.054653

14. Fyffe DC, Mukherjee S, Barnes LL, Manly JJ, Bennett DA, Crane PK. Explaining Differences in Episodic Memory Performance among Older African Americans and Whites: The Roles of Factors Related to Cognitive Reserve and Test Bias. Journal of the International Neuropsychological Society. 2011;17(4):625-638. doi:10.1017/S1355617711000476

15. Heaton R, Miller S, Taylor MJ, Grant I. Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults. Lutz, FL: Psychological Assessment Resources. Published online 2004.

16. Manly JJ, Jacobs DM, Touradji P, Small SA, Stern Y. Reading level attenuates differences in neuropsychological test performance between African American and White elders. Journal of the International Neuropsychological Society : JINS. 2002;8(3):341-348. doi:10.1017/S1355617702813157

17. Manly JJ, Byrd DA, Touradji P, Stern Y. Acculturation, Reading Level, and Neuropsychological Test Performance Among African American Elders. Applied Neuropsychology. 2004;11(1):37-46. doi:10.1207/s15324826an1101_5

18. Postal K. President’s Annual State of the Academy Report. The Clinical Neuropsychologist. 2018;32(1):1-9. doi:10.1080/13854046.2017.1406993

19. Sachs BC, Steenland K, Zhao L, et al. Expanded Demographic Norms for Version 3 of the Alzheimer Disease Centers’ Neuropsychological Test Battery in the Uniform Data Set. Alzheimer Dis Assoc Disord. 2020;34(3):191-197. doi:10.1097/WAD.0000000000000388

20. Tan YW, Burgess GH, Green RJ. The effects of acculturation on neuropsychological test performance: A systematic literature review. The Clinical Neuropsychologist. 2021;35(3):541-571. doi:10.1080/13854046.2020.1714740

21. Benge JF, Aguirre A, Scullin MK, et al. Digital Methods for Performing Daily Tasks Among Older Adults: An Initial Report of Frequency of Use and Perceived Utility.

Experimental Aging Research. 2023;0(0):1-22. doi:10.1080/0361073X.2023.2172950

22. Graves LV, Beacham D, Sawitzky N, et al. Informant sex/gender moderates associations between reported functioning and memory performance in Mexican American adults. The Clinical Neuropsychologist. 2023;0(0):1-23. doi:10.1080/13854046.2023.2246669

23. Graves LV, Hamill S, Larry M, Williams D. Informant Characteristics Influence Reports of Participant Functioning and their Associations with Neuropsychological Performance in Non-Hispanic Black Adults. Archives of Clinical Neuropsychology. 2023;38(7):1054-1067. doi:10.1093/arclin/acad022

24. Hackett K, Mis R, Drabick DAG, Giovannetti T. Informant Reporting in Mild Cognitive Impairment: Sources of Discrepancy on the Functional Activities Questionnaire. J Int Neuropsychol Soc. 2020;26(5):503-514. doi:10.1017/S1355617719001449

25. Siemers E, Holdridge KC, Sundell KL, Liu-Seifert H. Function and clinical meaningfulness of treatments for mild Alzheimer's disease. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2016;2(1):105-112. doi:10.1016/j.dadm.2016.02.006

26. Wu CY, Dodge HH, Gothard S, et al. Unobtrusive Sensing Technology Detects Ecologically Valid Spatiotemporal Patterns of Daily Routines Distinctive to Persons With Mild Cognitive Impairment. The Journals of Gerontology: Series A. 2022;77(10):2077-2084. doi:10.1093/gerona/glab293

27. Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE, Initiative for the ADN. Utility of Combinations of Biomarkers, Cognitive Markers, and Risk Factors to Predict Conversion From Mild Cognitive Impairment to Alzheimer Disease in Patients in the Alzheimer's Disease Neuroimaging Initiative. Arch Gen Psychiatry. 2011;68(9):961-969. doi:10.1001/archgenpsychiatry.2011.96

28. Perneczky R, Jessen F, Grimmer T, et al. Anti-amyloid antibody therapies in Alzheimer's disease. Brain. 2023;146(3):842-849. doi:10.1093/brain/awad005

29. Federman AD, Becker JH, Mindt MR, Cho D, Curtis L, Wisnivesky J. Rates of Undiagnosed Cognitive Impairment and Performance on the Montreal Cognitive Assessment Among Older Adults in Primary Care. J GEN INTERN MED. 2023;38(11):2511-2518. doi:10.1007/s11606-023-08102-w

30. Liu Y, Jun H, Becker A, Wallick C, Mattke S. Detection Rates of Mild Cognitive Impairment in Primary Care for the United States Medicare Population. J Prev Alzheimers Dis. Published online October 24, 2023. doi:10.14283/jpad.2023.131

31. Mattke S, Jun H, Chen E, Liu Y, Becker A, Wallick C. Expected and diagnosed rates

of mild cognitive impairment and dementia in the U.S. Medicare population: observational analysis. Alzheimer's Research & Therapy. 2023;15(1):128. doi:10.1186/s13195-023-01272-z

32. DiBenedetti DB, Slota C, Wronski SL, et al. Assessing what matters most to patients with or at risk for Alzheimer's and care partners: a qualitative study evaluating symptoms, impacts, and outcomes. Alzheimer's Research & Therapy. 2020;12(1):90. doi:10.1186/s13195-020-00659-6

33. Andrews JS, Desai U, Kirson NY, Zichlin ML, Ball DE, Matthews BR. Disease severity and minimal clinically important differences in clinical outcome assessments for Alzheimer's disease clinical trials. Alzheimers Dement (N Y). 2019;5:354-363. doi:10.1016/j.trci.2019.06.005

34. Lansdall CJ, McDougall F, Butler LM, et al. Establishing Clinically Meaningful Change on Outcome Assessments Frequently Used in Trials of Mild Cognitive Impairment Due to Alzheimer's Disease. J Prev Alzheimers Dis. 2023;10(1):9-18. doi:10.14283/jpad.2022.102

35. Evans S, McRae-McKee K, Wong MM, Hadjichrysanthou C, De Wolf F, Anderson R. The importance of endpoint selection: How effective does a drug need to be for success in a clinical trial of a possible Alzheimer's disease treatment? Eur J Epidemiol. 2018;33(7):635-644. doi:10.1007/s10654-018-0381-0

36. Posner H, Curiel R, Edgar C, 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(1-2):22-29.

37. Tochel C, Smith M, Baldwin H, et al. What outcomes are important to patients with mild cognitive impairment or Alzheimer's disease, their caregivers, and health-care professionals? A systematic review. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2019;11:231-247. doi:10.1016/j.dadm.2018.12.003

38. Cummings J, Lee G, Nahed P, et al. Alzheimer's disease drug development pipeline: 2022. Alzheimer's & Dementia: Translational Research & Clinical Interventions. 2022;8(1):e12295. doi:10.1002/trc2.12295

39. Kourtis LC, Regele OB, Wright JM, Jones GB. Digital biomarkers for Alzheimer's disease: the mobile/wearable devices opportunity. npj Digital Medicine. 2019;2(1):1-9. doi:10.1038/s41746-019-0084-2

40. Beattie ZT, Gothard S, Barnes LL, et al. The Collaboration and Harmonization of Digital Biomarkers for Alzheimer's Research. Alzheimer's & Dementia. 2022;18:e067984.

41. Papp KV, Samaroo A, Chou HC, et al. Unsupervised mobile cognitive testing for use in preclinical Alzheimer's disease. Alzheimers Dement (Amst).

2021;13(1):e12243. doi:10.1002/dad2.12243

42. Weizenbaum EL, Soberanes D, Hsieh S, et al. Capturing learning curves with the multiday Boston Remote Assessment of Neurocognitive Health (BRANCH): Feasibility, reliability, and validity. Neuropsychology. Published online November 16, 2023. doi:10.1037/neu0000933

43. Popp Z, Low S, Igwe A, et al. Shifting From Active to Passive Monitoring of Alzheimer Disease: The State of the Research. J Am Heart Assoc. 2024;13(2):e031247. doi:10.1161/JAHA.123.031247

44. Dodge HH, Zhu J, Mattek NC, Austin D, Kornfeld J, Kaye JA. Use of High-Frequency In-Home Monitoring Data May Reduce Sample Sizes Needed in Clinical Trials. PLOS ONE. 2015;10(9):e0138095. doi:10.1371/journal.pone.0138095

45. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and Validity of Ambulatory Cognitive Assessments. Assessment. 2018;25(1):14-30. doi:10.1177/1073191116643164

46. Weizenbaum E, Torous J, Fulford D. Cognition in Context: Understanding the Everyday Predictors of Cognitive Performance in a New Era of Measurement. JMIR mHealth and uHealth. 2020;8(7):e14328.

47. Nicosia J, Aschenbrenner AJ, Balota DA, et al. Reliability, validity, and feasibility of a smartphone-based cognitive assessment for preclinical Alzheimer disease. Alzheimer’s & Dementia. 2023;19(S4):e063363. doi:10.1002/alz.063363

48. Anderson M, Perrin A. Tech adoption climbs among older adults. Pew research center. Published online 2017:1-22.

49. Mobile Fact Sheet. Pew Research Center; 2021. https://www.pewresearch.org/internet/fact-sheet/mobile/

50. Holm E, Jacobsen KK, de Lony TB, et al. Frequency of missed or delayed diagnosis in dementia is associated with neighborhood socioeconomic status. Alzheimer’s & Dementia: Translational Research & Clinical Interventions. 2022;8(1):e12271. doi:10.1002/trc2.12271

51. Ferri CP, Jacob KS. Dementia in low-income and middle-income countries: Different realities mandate tailored solutions. PLOS Medicine. 2017;14(3):e1002271. doi:10.1371/journal.pmed.1002271

52. Connolly A, Gaehl E, Martin H, Morris J, Purandare N. Underdiagnosis of dementia in primary care: Variations in the observed prevalence and comparisons to the expected prevalence. Aging & Mental Health. 2011;15(8):978-984. doi:10.1080/13607863.2011.596805

53. Lin PJ, Daly AT, Olchanski N, et al. Dementia Diagnosis Disparities by Race and Ethnicity. Med Care. 2021;59(8):679-686. doi:10.1097/MLR.0000000000001577

54. Gianattasio KZ, Prather C, Glymour MM, Ciarleglio A, Power MC. Racial disparities and temporal trends in dementia misdiagnosis risk in the United States. Alzheimer's & Dementia: Translational Research & Clinical Interventions. 2019;5(1):891-898. doi:10.1016/j.trci.2019.11.008

55. Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. JAMA. 2017;318(13):1215-1216. doi:10.1001/jama.2017.11295

56. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nature Biotechnology. 2015;33:462-463. doi:10.1038/nbt.3223

57. Torous J, Kiang MV, Lorme J, Onnela JP. New Tools for New Research in Psychiatry: A Scalable and Customizable Platform to Empower Data Driven Smartphone Research. JMIR Mental Health. 2016;3(2):e16. doi:10.2196/mental.5165

58. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela JP. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacology. 2018;43(8):1660. doi:10.1038/s41386-018-0030-z

59. Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: Using smartphone sensors to monitor behavior and mental health. Psychiatric Rehabilitation Journal. 2015;38(3):218-226. doi:10.1037/prj0000130

60. Cohen A, Naslund JA, Chang S, et al. Relapse prediction in schizophrenia with smartphone digital phenotyping during COVID-19: a prospective, three-site, two-country, longitudinal study. Schizophr. 2023;9(1):1-8. doi:10.1038/s41537-023-00332-5

61. Kelkar RS, Currey D, Nagendra S, et al. Utility of Smartphone-Based Digital Phenotyping Biomarkers in Assessing Treatment Response to Transcranial Magnetic Stimulation in Depression: Proof-of-Concept Study. JMIR Formative Research. 2023;7(1):e40197. doi:10.2196/40197

62. Lane E, D'Arcey J, Kidd S, et al. Digital Phenotyping in Adults with Schizophrenia: A Narrative Review. Curr Psychiatry Rep. Published online October 20, 2023. doi:10.1007/s11920-023-01467-z

63. Leaning IE, Ikani N, Savage HS, et al. From smartphone data to clinically relevant predictions: A systematic review of digital phenotyping methods in depression. Neuroscience & Biobehavioral Reviews. 2024;158:105541. doi:10.1016/j.neubiorev.2024.105541

64. Ren B, Xia CH, Gehrman P, Barnett I, Satterthwaite T. Measuring Daily Activity Rhythms in Young Adults at Risk of Affective Instability Using Passively

Collected Smartphone Data: Observational Study. JMIR Formative Research. 2022;6(9):e33890. doi:10.2196/33890

65. Saeb S, Zhang M, Karr CJ, et al. Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study. Journal of Medical Internet Research. 2015;17(7):e175. doi:10.2196/jmir.4273

66. Vesel C, Rashidisabet H, Zulueta J, et al. Effects of mood and aging on keystroke dynamics metadata and their diurnal patterns in a large open-science sample: A BiAffect iOS study. J Am Med Inform Assoc. 2020;27(7):1007-1018. doi:10.1093/jamia/ocaa057

67. Zulueta J, Piscitello A, Rasic M, et al. Predicting mood disturbance severity with mobile phone keystroke metadata: A biaffect digital phenotyping study. Journal of Medical Internet Research. 2018;20(7):undefined-undefined. doi:10.2196/jmir.9775

68. Chen MH, Leow A, Ross MK, et al. Associations between smartphone keystroke dynamics and cognition in MS. DIGITAL HEALTH. 2022;8:20552076221143234. doi:10.1177/20552076221143234

69. Chen R, Jankovic F, Marinsek N, et al. Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. ; 2019:2145-2155.

70. Dagum P. Digital biomarkers of cognitive function. npj Digital Medicine. 2018;1(1):10. doi:10.1038/s41746-018-0018-4

71. Gordon ML, Gatys L, Guestrin C, Bigham JP, Trister A, Patel K. App Usage Predicts Cognitive Ability in Older Adults. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. CHI ’19. ACM; 2019:168:1-168:12. doi:10.1145/3290605.3300398

72. Ning E, Cladek AT, Ross MK, et al. Smartphone-derived Virtual Keyboard Dynamics Coupled with Accelerometer Data as a Window into Understanding Brain Health: Smartphone Keyboard and Accelerometer as Window into Brain Health. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems. CHI ’23. Association for Computing Machinery; 2023:1-15. doi:10.1145/3544548.3580906

73. Paolillo EW, Casaletto KB, Clark AL, et al. Passively monitored smartphone battery percentage relates to disease severity in frontotemporal dementia: A proof-of-concept study. Alzheimer’s & Dementia. 2023;19(S18):e079178. doi:10.1002/alz.079178

74. Ross MK, Demos AP, Zulueta J, et al. Naturalistic smartphone keyboard typing reflects processing speed and executive function. Brain and Behavior.

2021;11(11):e2363. doi:10.1002/brb3.2363

75. Bartlett Ellis R, Wright J, Miller LS, et al. Lessons Learned: Beta-Testing the Digital Health Checklist for Researchers Prompts a Call to Action by Behavioral Scientists. J Med Internet Res. 2021;23(12):e25414. doi:10.2196/25414

76. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med. 2020;10(4):1004-1015. doi:10.1093/tbm/ibz074

77. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. npj Digital Medicine. 2019;2(1):1-11. doi:10.1038/s41746-019-0166-1

78. Schmidt S, D'Alfonso S. Clinician perspectives on how digital phenotyping can inform client treatment. Acta Psychologica. 2023;235:103886. doi:10.1016/j.actpsy.2023.103886

79. Vaidyam A, Halamka J, Torous J. Actionable digital phenotyping: a framework for the delivery of just-in-time and longitudinal interventions in clinical healthcare. Mhealth. 2019;5:25. doi:10.21037/mhealth.2019.07.04

80. Torous J, Wisniewski H, Bird B, et al. Creating a Digital Health Smartphone App and Digital Phenotyping Platform for Mental Health and Diverse Healthcare Needs: an Interdisciplinary and Collaborative Approach. J technol behav sci. 2019;4(2):73-85. doi:10.1007/s41347-019-00095-w

81. Vaidyam A, Halamka J, Torous J. Enabling Research and Clinical Use of Patient-Generated Health Data (the mindLAMP Platform): Digital Phenotyping Study. JMIR mHealth and uHealth. 2022;10(1):e30557. doi:10.2196/30557

82. Jak AJ, Bondi MW, Delano-Wood L, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. The American Journal of Geriatric Psychiatry. 2009;17(5):368-375.

83. Melcher J, Hays R, Torous J. Digital phenotyping for mental health of college students: a clinical review. Evid Based Mental Health. 2020;23(4):161-166. doi:10.1136/ebmental-2020-300180

84. Schretlen DJ, Winicki JM, Meyer SM, Testa SM, Pearlson GD, Gordon B. Development, psychometric properties, and validity of the Hopkins Adult Reading Test (HART). The Clinical Neuropsychologist. 2009;23(6):926-943.

85. Folstein M, Folstein S, McHugh P. "Mini-Mental-State": a practical method for grading the cognitive state of patients for the clinician. Journal of Psychological Research. 1975;12:189-198.

86. Reitan RM. Validity of the Trail Making Test as an indicator of organic brain damage.

Perceptual and motor skills. 1958;8(3):271-276.

87. Wechsler D. Manual for the Wechsler memory scale-revised. San Antonio, TX: Psychological Corporation. Published online 1987.

88. Salthouse TA. Mediation of adult age differences in cognition by reductions in working memory and speed of processing. Psychological Science. 1991;2(3):179-183.

89. Brandt J, Benedict RH. Hopkins Verbal Learning Test–Revised: Professional Manual. Psychological Assessment Resources; 2001.

90. Benedict RH, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. Psychological Assessment. 1996;8(2):145.

91. Schretlen D, Testa S, Pearlson G. Calibrated neuropsychological normative system professional manual. Lutz, FL: Psychological Assessment Resources. Published online 2010.

92. Goodglass H, Kaplan E. Boston diagnostic examination for aphasia. Philadelphia: Lea and Febiger. Published online 1983.

93. Farias S, Mungas D, Harvey DJ, Simmons A, Reed BR, DeCarli C. The measurement of everyday cognition: Development and validation of a short form of the Everyday Cognition scales. Alzheimer's & Dementia. 2011;7(6):593-601. doi:10.1016/j.jalz.2011.02.007

94. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of Functional Activities in Older Adults in the Community. J Gerontol. 1982;37(3):323-329. doi:10.1093/geronj/37.3.323

95. Yesavage JA. Geriatric depression scale. Psychopharmacol Bull. 1988;24(4):709-711.

96. Bohannon RW. Reference Values for the Timed Up and Go Test: A Descriptive Meta-Analysis. Journal of Geriatric Physical Therapy. 2006;29(2):64.

97. Podsiadlo D, Richardson S. The Timed "Up & Go": A Test of Basic Functional Mobility for Frail Elderly Persons. Journal of the American Geriatrics Society. 1991;39(2):142-148. doi:10.1111/j.1532-5415.1991.tb01616.x

98. Baker PS, Bodner EV, Allman RM. Measuring Life-Space Mobility in Community-Dwelling Older Adults. Journal of the American Geriatrics Society. 2003;51(11):1610-1614. doi:10.1046/j.1532-5415.2003.51512.x

99. Peel C, Baker PS, Roth DL, Brown CJ, Bodner EV, Allman RM. Assessing Mobility in Older Adults: The UAB Study of Aging Life-Space Assessment. Physical Therapy. 2005;85(10):1008-1019. doi:10.1093/ptj/85.10.1008

100. Crowe M, Andel R, Wadley VG, Okonkwo OC, Sawyer P, Allman RM. Life-Space and Cognitive Decline in a Community-Based Sample of African American and Caucasian Older Adults. The Journals of Gerontology: Series A. 2008;63(11):1241-1245. doi:10.1093/gerona/63.11.1241

101. James BD, Boyle PA, Buchman AS, Barnes LL, Bennett DA. Life Space and Risk of Alzheimer Disease, Mild Cognitive Impairment, and Cognitive Decline in Old Age. The American Journal of Geriatric Psychiatry. 2011;19(11):961-969. doi:10.1097/JGP.0b013e318211c219

102. Berry HL, Rodgers B, Dear KBG. Preliminary development and validation of an Australian community participation questionnaire: Types of participation and associations with distress in a coastal community. Social Science & Medicine. 2007;64(8):1719-1737. doi:10.1016/j.socscimed.2006.12.009

103. Brett L, Georgiou A, Jorgensen M, et al. Ageing well: evaluation of social participation and quality of life tools to enhance community aged care (study protocol). BMC Geriatrics. 2019;19(1):78. doi:10.1186/s12877-019-1094-2

104. van Deursen AJAM, Bolle CL, Hegner SM, Kommers PAM. Modeling habitual and addictive smartphone behavior: The role of smartphone usage types, emotional intelligence, social stress, self-regulation, age, and gender. Computers in Human Behavior. 2015;45:411-420. doi:10.1016/j.chb.2014.12.039

105. Roque NA, Boot WR. A New Tool for Assessing Mobile Device Proficiency in Older Adults: The Mobile Device Proficiency Questionnaire. J Appl Gerontol. 2018;37(2):131-156. doi:10.1177/0733464816642582

106. Barnett I, Onnela JP. Inferring mobility measures from GPS traces with missing data. Biostatistics. 2020;21(2):e98-e112. doi:10.1093/biostatistics/kxy059

107. Ian Barnett. Ian James Barnett GitHub Profile. (https://github.com/ianjamesbarnett).

108. Torgerson WS. Multidimensional scaling: I. Theory and method. Psychometrika. 1952;17(4):401-419.

109. Almeida OP, Almeida SA. Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. Int J Geriatr Psychiatry. 1999;14(10):858-865. doi:10.1002/(sici)1099-1166(199910)14:10<858::aid-gps35>3.0.co;2-8

110. Kiang MV, Chen JT, Krieger N, et al. Sociodemographic Characteristics of Missing Data in Digital Phenotyping. medRxiv. Published online January 4, 2021:2020.12.29.20249002. doi:10.1101/2020.12.29.20249002

111. Onnela JP. Opportunities and challenges in the collection and analysis of digital phenotyping data. Neuropsychopharmacol. 2021;46(1):45-54. doi:10.1038/s41386-020-0771-3

112. Eliacin J, Polsinelli AJ, Epperson F, et al. Barriers and facilitators to participating in Alzheimer's disease biomarker research in black and white older adults. Alzheimers Dement (N Y). 2023;9(2):e12399. doi:10.1002/trc2.12399

113. Indorewalla KK, O'Connor MK, Budson AE, Guess (DiTerlizzi) C, Jackson J. Modifiable Barriers for Recruitment and Retention of Older Adults Participants from Underrepresented Minorities in Alzheimer's Disease Research. Journal of Alzheimer's Disease. 2021;80(3):927-940. doi:10.3233/JAD-201081

114. Minkler M, Chang C. Community-Based Participatory Research. Migration and health: A research methods handbook. Published online 2014:361.

115. Neffa-Creech D, Aggarwal R, Stowell C, et al. Understanding Barriers and Facilitators to Signing Up for a Mobile-Responsive Registry to Recruit Healthy Volunteers and Members of Underrepresented Communities for Alzheimer's Disease Prevention Studies. J Prev Alzheimers Dis. 2023;10(4):865-874. doi:10.14283/jpad.2023.67

116. Bangen KJ, Calcetas AT, Thomas KR, et al. Greater accelerometer-measured physical activity is associated with better cognition and cerebrovascular health in older adults. Journal of the International Neuropsychological Society. 2023;29(9):859-869. doi:10.1017/S1355617723000140

117. Buchman AS, Boyle PA, Wilson RS, Tang Y, Bennett DA. Frailty is Associated With Incident Alzheimer's Disease and Cognitive Decline in the Elderly. Psychosomatic Medicine. 2007;69(5):483. doi:10.1097/psy.0b013e318068de1d

118. 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):1323-1329. doi:10.1212/WNL.0b013e3182535d35

119. Casaletto KB, Renteíıa MA, Pa J, et al. Late-Life Physical and Cognitive Activities Independently Contribute to Brain and Cognitive Resilience. J Alzheimers Dis. 2020;74(1):363-376. doi:10.3233/JAD-191114

120. Gallaway PJ, Miyake H, Buchowski MS, et al. Physical Activity: A Viable Way to Reduce the Risks of Mild Cognitive Impairment, Alzheimer's Disease, and Vascular Dementia in Older Adults. Brain Sciences. 2017;7(2):22. doi:10.3390/brainsci7020022

121. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet. 2020;396(10248):413-446. doi:10.1016/S0140-6736(20)30367-6

122. Petersen J, Austin D, Mattek N, Kaye J. Time Out-of-Home and Cognitive, Physical, and Emotional Wellbeing of Older Adults: A Longitudinal Mixed Effects Model. PLOS ONE. 2015;10(10):e0139643. doi:10.1371/journal.pone.0139643

123. Scarmeas N, Levy G, Tang MX, Manly J, Stern Y. Influence of leisure activity on the incidence of Alzheimer's Disease. Neurology. 2001;57(12):2236-2242.

124. Wang L, Larson EB, Bowen JD, van Belle G. Performance-Based Physical Function and Future Dementia in Older People. Archives of Internal Medicine. 2006;166(10):1115-1120. doi:10.1001/archinte.166.10.1115

125. Rykov Y, Thach TQ, Bojic I, Christopoulos G, Car J. Digital Biomarkers for Depression Screening With Wearable Devices: Cross-sectional Study With Machine Learning Modeling. JMIR Mhealth Uhealth. 2021;9(10):e24872. doi:10.2196/24872

126. Bergua V, Bouisson J, Dartigues JF, et al. Restriction in Instrumental Activities of Daily Living in Older Persons: Association with Preferences for Routines and Psychological Vulnerability. Int J Aging Hum Dev. 2013;77(4):309-329. doi:10.2190/AG.77.4.c

127. Insel KC, Einstein GO, Morrow DG, Koerner KM, Hepworth JT. Multifaceted Prospective Memory Intervention to Improve Medication Adherence. J Am Geriatr Soc. 2016;64(3):561-568. doi:10.1111/jgs.14032

128. Suchy Y, Mora MG, DesRuisseaux LA, Brothers SL. It's complicated: Executive functioning moderates impacts of daily busyness on everyday functioning in community-dwelling older adults. Journal of the International Neuropsychological Society. 2023;29(9):850-858. doi:10.1017/S1355617723000048

129. Tomaszewski Farias S, Schmitter-Edgecombe M, Weakley A, et al. Compensation Strategies in Older Adults: Association With Cognition and Everyday Function. Am J Alzheimers Dis Other Demen. 2018;33(3):184-191. doi:10.1177/1533317517753361

130. Tassoni MB, Drabick DAG, Giovannetti T. The frequency of self-reported memory failures is influenced by everyday context across the lifespan: Implications for neuropsychology research and clinical practice. The Clinical Neuropsychologist. 2023;37(6):1115-1135. doi:10.1080/13854046.2022.2112297

131. Bielak AAM, Hultsch DF, Strauss E, Macdonald SWS, Hunter MA. Intraindividual variability in reaction time predicts cognitive outcomes 5 years later. Neuropsychology. 2010;24(6):731-741. doi:10.1037/a0019802

132. Bielak AAM, Hultsch DF, Strauss E, MacDonald SWS, Hunter MA. Intraindividual variability is related to cognitive change in older adults: evidence for within-person coupling. Psychol Aging. 2010;25(3):575-586. doi:10.1037/a0019503

133. Dodge HH, Mattek NC, Austin D, Hayes TL, Kaye JA. In-home walking speeds and variability trajectories associated with mild cognitive impairment. Neurology. 2012;78(24):1946-1952. doi:10.1212/WNL.0b013e318259e1de

134. Hayes TL, Abendroth F, Adami A, Pavel M, Zitzelberger TA, Kaye JA. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. Alzheimers Dement. 2008;4(6):395-405. doi:10.1016/j.jalz.2008.07.004

135. Kaye J, Mattek N, Dodge HH, et al. Unobtrusive measurement of daily computer use to detect mild cognitive impairment. Alzheimers Dement. 2014;10(1):10-17. doi:10.1016/j.jalz.2013.01.011

136. Jutten RJ, Amariglio RE, Maruff P, et al. Increased intraindividual variability in reaction time performance is associated with emerging cognitive decline in cognitively unimpaired adults. Neuropsychology. Published online November 16, 2023. doi:10.1037/neu0000928

137. Biddle KD, d'Oleire Uquillas F, Jacobs HIL, et al. Social Engagement and Amyloid-β-Related Cognitive Decline in Cognitively Normal Older Adults. The American Journal of Geriatric Psychiatry. 2019;27(11):1247-1256. doi:10.1016/j.jagp.2019.05.005

138. Fratiglioni L, Wang HX, Ericsson K, Maytan M, Winblad B. Influence of social network on occurrence of dementia: a community-based longitudinal study. The Lancet. 2000;355(9212):1315-1319. doi:10.1016/S0140-6736(00)02113-9

139. Lee M, Whitsel E, Avery C, et al. Variation in Population Attributable Fraction of Dementia Associated With Potentially Modifiable Risk Factors by Race and Ethnicity in the US. JAMA Netw Open. 2022;5(7):e2219672. doi:10.1001/jamanetworkopen.2022.19672

140. Penninkilampi R, Casey AN, Singh MF, Brodaty H. The Association between Social Engagement, Loneliness, and Risk of Dementia: A Systematic Review and Meta-Analysis. Journal of Alzheimer's Disease. 2018;66(4):1619-1633. doi:10.3233/JAD-180439

141. Ren Y, Savadlou A, Park S, Siska P, Epp JR, Sargin D. The impact of loneliness and social isolation on the development of cognitive decline and Alzheimer's Disease. Frontiers in Neuroendocrinology. 2023;69:101061. doi:10.1016/j.yfrne.2023.101061

142. Mcalister C, Schmitter-Edgecombe M, Lamb R. Examination of Variables That May Affect the Relationship Between Cognition and Functional Status in Individuals with Mild Cognitive Impairment: A Meta-Analysis. Arch Clin Neuropsychol. 2016;31(2):123-147. doi:10.1093/arclin/acv089

143. Royall DR, Lauterbach EC, Kaufer D, Malloy P, Coburn KL, Black KJ. The Cognitive Correlates of Functional Status: A Review From the Committee on Research of the American Neuropsychiatric Association. JNP. 2007;19(3):249-265. doi:10.1176/jnp.2007.19.3.249

144. Sverdlov O, Curcic J, Hannesdottir K, et al. A Study of Novel Exploratory Tools,

Digital Technologies, and Central Nervous System Biomarkers to Characterize Unipolar Depression. Front Psychiatry. 2021;12. doi:10.3389/fpsyt.2021.640741

145. Torous J, Staples P, Barnett I, Sandoval LR, Keshavan M, Onnela JP. Characterizing the clinical relevance of digital phenotyping data quality with applications to a cohort with schizophrenia. npj Digital Medicine. 2018;1(1):15. doi:10.1038/s41746-018-0022-8

146. Murnane EL, Abdullah S, Matthews M, et al. Mobile Manifestations of Alertness: Connecting Biological Rhythms with Patterns of Smartphone App Use. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services. MobileHCI '16. ACM; 2016:465-477. doi:10.1145/2935334.2935383

147. Quisel T, Kale DC, Foschini L. Intra-day Activity Better Predicts Chronic Conditions. arXiv:161201200 [cs, stat]. Published online December 4, 2016. Accessed November 23, 2019. http://arxiv.org/abs/1612.01200

148. Stachl C, Au Q, Schoedel R, et al. Predicting personality from patterns of behavior collected with smartphones. PNAS. 2020;117(30):17680-17687. doi:10.1073/pnas.1920484117

149. Barnett I, Torous J, Staples P, Keshavan M, Onnela JP. Beyond smartphones and sensors: choosing appropriate statistical methods for the analysis of longitudinal data. J Am Med Inform Assoc. 2018;25(12):1669-1674. doi:10.1093/jamia/ocy121

150. Barnett I, Onnela JP. Inferring mobility measures from GPS traces with missing data. arXiv preprint arXiv:160606328. Published online 2016.

151. Langholm C, Kowatsch T, Bucci S, Cipriani A, Torous J. Exploring the Potential of Apple SensorKit and Digital Phenotyping Data as New Digital Biomarkers for Mental Health Research. Digital Biomarkers. 2023;7(1):104-114. doi:10.1159/000530698

152. Nichols E, Steinmetz JD, Vollset SE, et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. The Lancet Public Health. 2022;7(2):e105-e125. doi:10.1016/S2468-2667(21)00249-8

153. Moore RC, Kaufmann CN, Rooney AS, et al. Feasibility and Acceptability of Ecological Momentary Assessment of Daily Functioning Among Older Adults with HIV. The American Journal of Geriatric Psychiatry. 2017;25(8):829-840. doi:10.1016/j.jagp.2016.11.019

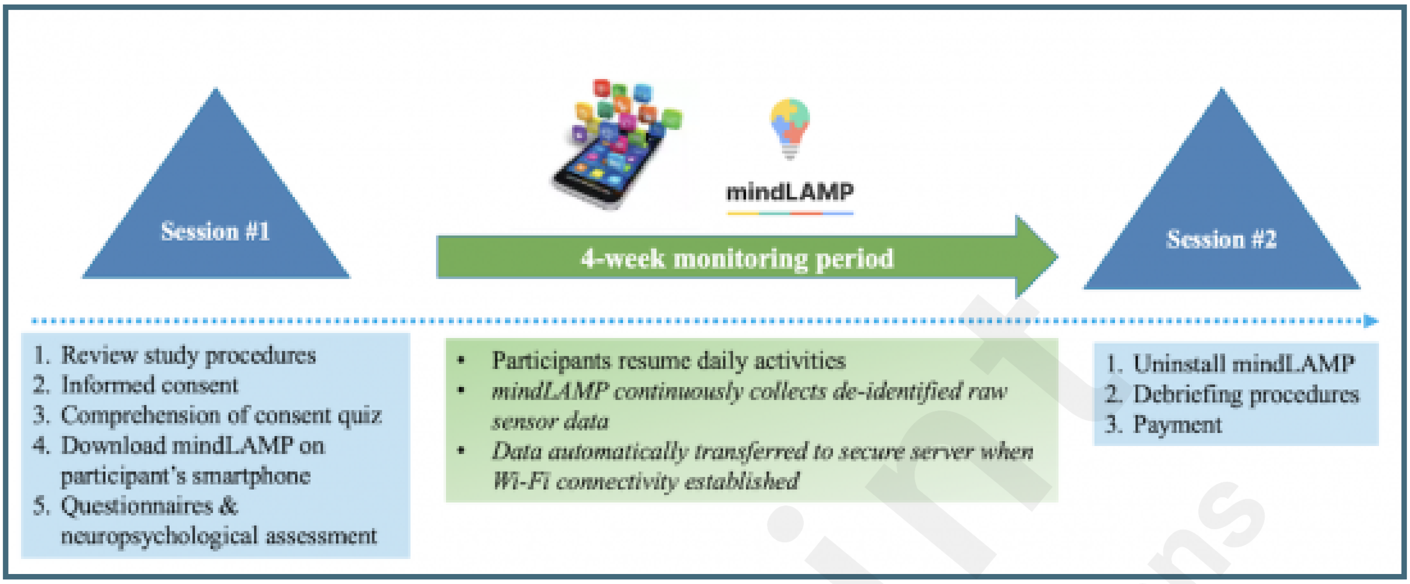
154. Schmidt C, Collette F, Cajochen C, Peigneux P. A time to think: Circadian rhythms in human cognition. Cognitive Neuropsychology. 2007;24(7):755-789. doi:10.1080/02643290701754158

155. Wisniewski H, Henson P, Torous J. Using a Smartphone App to Identify Clinically Relevant Behavior Trends via Symptom Report, Cognition Scores, and Exercise Levels: A Case Series. Frontiers in Psychiatry. 2019;10. Accessed March 18, 2023. https://www.frontiersin.org/articles/10.3389/fpsyt.2019.00652

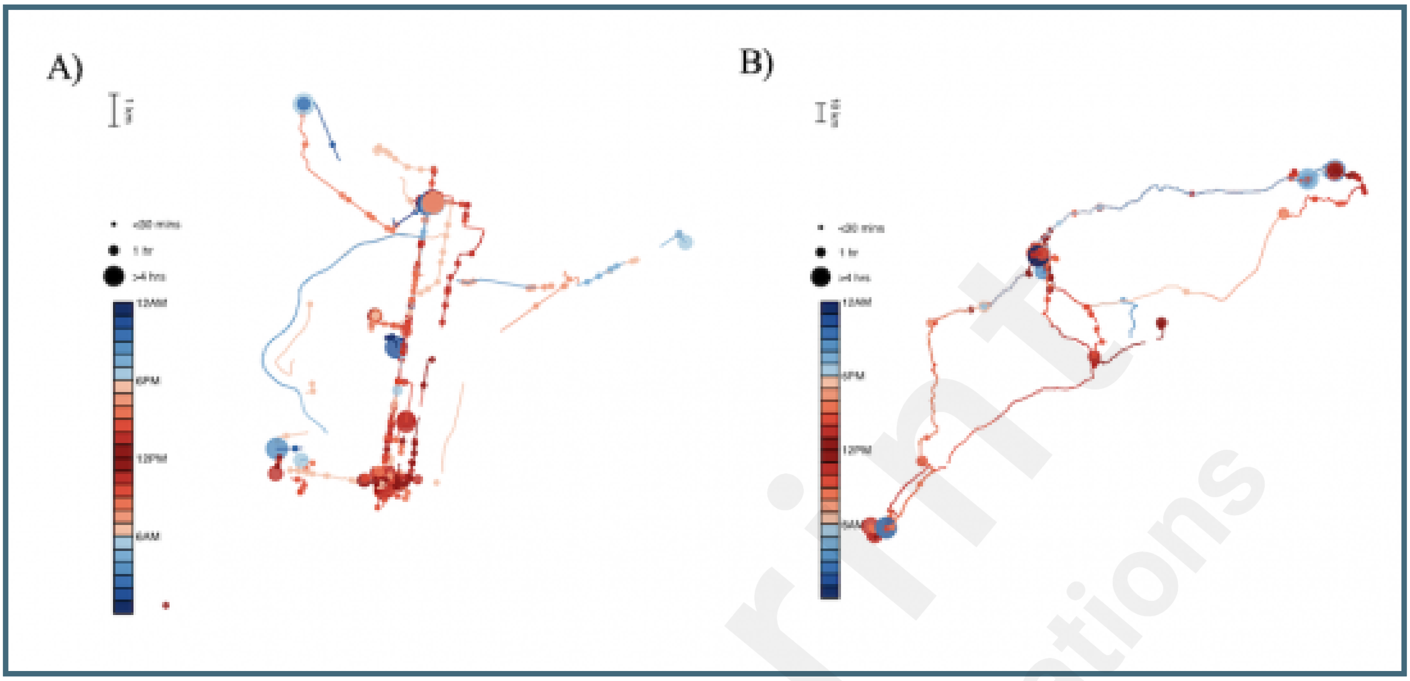
Supplementary Files

Figures

Study Timeline.



Example monthly mobility plots for two participants: A) 67-year-old female living in northern Philadelphia; B) 70-year-old female living in a suburb of Philadelphia.



Multimedia Appendixes

Comprehension of consent quiz administered to participants following informed consent procedures.
URL: http://asset.jmir.pub/assets/dfd71808de625844c0e3f8adb60a288f.png

Tables S1-S10 including supplemental data and analyses.
URL: http://asset.jmir.pub/assets/42b1988e4cf430d55995aea6913220d9.xls

Between-group ANOVA with Welch's test for unequal group variances depicting seasonal differences in GPS activity features, driven by greater activity in summer months.
URL: http://asset.jmir.pub/assets/55767fa28143a339cc74e3d99735 87e1.png

Between-group ANOVA with Welch's test for unequal group variances depicting differences in GPS variability according to occupational status, driven by greater mobility variability in retired participants.
URL: http://asset.jmir.pub/assets/b3cfbd405b981df3b0e8298f39179044.png