Journal of the American Heart Association

ORIGINAL RESEARCH

Comparing Cognitive Tests and Smartphone-Based Assessment in 2 US Community-Based Cohorts

Ileana De Anda-Duran ⓘ, MD, MPH*; Preeti Sunderaraman ⓘ, PhD*; Edward Searls, BA;
Shirine Moukaled ⓘ, MS; Xuanyi Jin, MD; Zachary Popp ⓘ, MPH; Cody Karjadi ⓘ, BS, MS;
Phillip H. Hwang ⓘ, PhD, MPH; Huitong Ding, PhD; Sherral Devine ⓘ, PhD; Ludy C. Shih ⓘ, MD, M.MSc;
Spencer Low ⓘ, BS; Honghuang Lin ⓘ, PhD; Vijaya B. Kolachalama ⓘ, PhD; Lydia Bazzano ⓘ, MD, PhD;
David J. Libon ⓘ, PhD; Rhoda Au ⓘ, PhD, MBA

BACKGROUND: Smartphone-based cognitive assessments have emerged as promising tools, bridging gaps in accessibility and reducing bias in Alzheimer disease and related dementia research. However, their congruence with traditional neuropsychological tests and usefulness in diverse cohorts remain underexplored.

METHODS AND RESULTS: A total of 406 FHS (Framingham Heart Study) and 59 BHS (Bogalusa Heart Study) participants with traditional neuropsychological tests and digital assessments using the Defense Automated Neurocognitive Assessment (DANA) smartphone protocol were included. Regression models investigated associations between DANA task digital measures and a neuropsychological global cognitive Z score (Global Cognitive Score [GCS]), and neuropsychological domain-specific Z scores. FHS participants' mean age was 57 (SD, 9.75) years, and 44% (179) were men. BHS participants' mean age was 49 (4.4) years, and 28% (16) were men. Participants in both cohorts with the lowest neuropsychological performance (lowest quartile, GCS1) demonstrated lower DANA digital scores. In the FHS, GCS1 participants had slower average response times and decreased cognitive efficiency scores in all DANA tasks ($P<0.05$). In BHS, participants in GCS1 had slower average response times and decreased cognitive efficiency scores for DANA Code Substitution and Go/No-Go tasks, although this was not statistically significant. In both cohorts, GCS was significantly associated with DANA tasks, such that higher GCS correlated with faster average response times ($P<0.05$) and increased cognitive efficiency (all $P<0.05$) in the DANA Code Substitution task.

CONCLUSIONS: Our findings demonstrate that smartphone-based cognitive assessments exhibit concurrent validity with a composite measure of traditional neuropsychological tests. This supports the potential of using smartphone-based assessments in cognitive screening across diverse populations and the scalability of digital assessments to community-dwelling individuals.

Key Words: Alzheimer's disease and related dementia ■ cognition ■ community-based cohorts ■ digital cognitive measures

Alzheimer disease and Alzheimer disease-related dementia (ADRD) are life-course conditions that affect >55 million people worldwide and are major contributors to disability and mortality.[1–3] Thus, the early identification of at-risk individuals who are most likely to benefit from preventive/therapeutic measures and clinical trial recruitment is imperative.

Considerable progress has been achieved in diagnostic methods, including blood-based biomarkers, neuroimaging, and cerebrospinal fluid analysis.[4–6]

JAHA is available at: www.ahajournals.org/journal/jaha

RESEARCH PERSPECTIVE

What New Question Does This Study Raise?

- Concurrent validity: this study demonstrates congruence between smartphone-based cognitive tasks and traditional neuropsychological tests in 2 diverse cohorts, suggesting that digital assessments may accurately capture cognitive performance.
- The consistent associations across different cognitive domains support the validity of smartphone-based tools in assessing executive function, attention/processing speed, and memory.

What Question Should Be Addressed Next?

- Future research should investigate the predictive validity of smartphone-based cognitive assessments in identifying cognitive decline or dementia progression, which would provide valuable insights into the potential clinical usefulness of these digital tools and address the regulatory and reimbursement aspects crucial for the integration of digital cognitive testing.

Nonstandard Abbreviations and Acronyms

BHS	Bogalusa Heart Study
DANA	Defense Automated Neurocognitive Assessment
FHS	Framingham Heart Study
GCS	Global Cognitive Score
TMT	Trail-Making Test

Despite remarkable innovations, such methods can be invasive, expensive, and difficult to obtain, particularly in rural areas and low-resource settings, posing enormous challenges when it comes to scaling them at the population level.[7] Furthermore, these methods do not enable the dynamic monitoring of disease progression across the life course that is needed to identify critical time windows for prevention, intervention, or treatment.[8]

Cognitive assessment presents itself as a relatively low-cost, not physically intrusive method for characterizing and diagnosing ADRD.[9,10] Brief cognitive screening tests, such as the Mini-Mental State Examination or the Montreal Cognitive Assessment, have traditionally provided a time-efficient strategy for monitoring cognition, particularly in primary care settings.[11] However, these brief tests have a low sensitivity for identifying individuals in the early stages of the disease.[12,13] Comprehensive neuropsychological protocols are more sensitive and accurate in identifying individuals with ADRD at different stages, yet their administration requires a trained administrator and necessitates participants to visit a specialized center for a detailed assessment. Another major limitation is that some of the commonly administered neuropsychological tests were developed decades ago and are known to exhibit racial, linguistic, and cultural biases.[14–17] Although a comprehensive neurocognitive assessment remains essential for a thorough evaluation, brief cognitive screening can still offer valuable insights for the timely identification of individuals who may require a comprehensive evaluation for a final diagnosis and appropriate treatment and care. Thus, there is an increasing need for unbiased, inexpensive, accessible, and scalable tools that enable frontline health care providers to assess individuals at higher risk for developing dementia.

In recent years, digital technologies in the form of smartphone-based tasks have shown promise for measuring cognitive abilities due to their adequate feasibility, easy accessibility, and increased scalability.[18] The increased global adoption of smartphones can aid in bridging the gap between lengthy traditional methods required to comprehensively assess cognition and the needs of communities in different settings that require accessible time- and cost-efficient methods for cognitive screening.[18,19] These decentralized, scalable approaches that allow for in-home data collection may address challenges posed by geography and segregation, thereby improving the participation of underrepresented groups in ADRD research and overcoming some of the limitations associated with racial, linguistic, and cultural biases when assessing cognitive function.[18] Evidence supporting the feasibility of using smartphone-based platforms for cognition has been shown in community-based longitudinal cohorts, such as the FHS (Framingham Heart Study).[20,21] However, the effectiveness of digital tools as an alternative to traditional assessments and the usefulness of smartphone-based technologies in diverse cohorts remains understudied.

This study investigates the congruence between cognitive performance measured by traditional paper and pencil neuropsychological tests and digital cognitive measures derived from smartphone-based tasks. Participants from 2 distinct epidemiological cohorts, the FHS in the Northeast United States and the BHS (Bogalusa Heart Study) in the Southern United States, with traditional neuropsychological test assessment and digital cognitive measures, participated in this study, wherein the Food and Drug Administration-cleared Defense Automated Neurocognitive Assessment (DANA) protocol was administered via participants' own smartphones.

METHODS

Data Availability

Data, methods, and materials used to conduct the research in the article and additional information can be shared on responsible request.

Study Design and Population

The FHS is a longitudinal, population-based cohort study that began in 1948 in Framingham, Massachusetts. The FHS focuses on the natural history of cardiovascular disease spanning 3 generations of participants' descendants. Neuropsychological evaluations began in 1976 with the collection of a brief cognitive screen and progressed to a comprehensive cognitive evaluation in 1999. The present study includes participants from 4 different subcohorts of the FHS, the Gen 2 (Offspring), Gen 3 (Omni 1, Generation 3), and New Offspring Spouse. In 1971, adult children of initial study participants and their spouses joined the Offspring cohort. A third-generation cohort of people (Gen 3) with at least 1 parent in the Offspring cohort was added in 2001. Omni cohorts of minorities were incorporated in the 1990s. A total of 408 individuals participated in the digital-based cognitive assessment from October 2022 to February 2023; from these, 2 did not have complete digital data, and 62 more had missing data on different neuropsychological paper and paper tests and were excluded from the analysis (Figure S1). The inclusion criteria to participating in the digital-based cognitive assessment included owning a smartphone, proficiency in spoken English, and having a WiFi connection. All FHS participants recruited for the digital study were characterized as cognitively normal.

The BHS began in 1973 and is in the rural town of Bogalusa, Louisiana. The BHS is a community-based longitudinal cohort dedicated to the study of the natural history of cardiovascular disease, with cross-sectional examinations beginning in childhood. During the 2013 to 2016 examination, the first neuropsychological evaluation was implemented as part of the study protocol; the sample of participants consisted of 1298 adults (48.17±5.27 years of age). Additional details about the cohort design have been previously published.[22] From this sample, a total of 67 participants were included in the digital-based cognitive assessment from September 2021 to January 2023; from these, 8 had missing data on different neuropsychological pencil and paper tests and were excluded from the analysis (Figure S2). Similar to FHS, the inclusion criteria to participating in the digital-based cognitive assessment included owning a smartphone, proficiency in spoken English, and having a WiFi connection. All BHS participants were characterized as cognitively normal.

The Boston Medical Center/Boston University Medical Campus institutional review board and the institutional review board of Tulane University Health Sciences Center approved the current study, and all participants provided informed consent.

Traditional Neuropsychological Assessment

A baseline neuropsychological evaluation was administered to participants from each cohort in the FHS. A baseline neuropsychological for Gen 2/Omni 1 was first performed in 1999 to 2005; for Gen 3, Omni 2, and New Offspring Spouse, baseline assessment was done between 2008 and 2013. Since then, subsequent neuropsychological assessments were conducted at an average 2- and 6-year intervals. Details on the neuropsychological tests administered, their summary scores, and normative data have been previously published.[23–25] Briefly, each neuropsychological assessment lasted ≈60 to 90 minutes and was administered by a trained test administrator following a standard research protocol.[26] The specific neuropsychological tests analyzed for the current research included the most recent examination and assessed tests from the following domains: attention and processing speed (Trail-Making Test [TMT] Part A; Wechsler Adult Intelligence Scale [WAIS] Digit Span Forward subtest); executive functioning (WAIS Digit Span Backward, TMT Part B); language (WAIS Similarities subtest, Boston Naming Test [30-item version]); verbal episodic memory (Wechsler Memory Scale Logical Memory–Immediate and Delayed Free Recall); visuospatial perception and organization (Hooper Visual Organization Test); and visual memory (Wechsler Memory Scale Visual Reproduction–Immediate and Delayed Free Recall).

In the BHS, like the FHS protocol, the neuropsychological test evaluation requires trained personnel and takes ≈45 minutes to complete. The BHS examinations are conducted every 4 to 5 years and evaluate executive control, attention/information processing speed, and verbal episodic memory domains. The specific neuropsychological tests analyzed for the current research included the most recent examination, including the Digit Span Forward and Digit Span Backward tests (WAIS-IV), TMT-A and TMT-B, WAIS-IV Logical Memory Immediate and Delayed Free Recall, Delayed Recognition (Wechsler Memory Scale-IV), and WAIS-IV Digit Symbol Coding tests.

A Global Cognitive Score (GCS) was computed using the same approach for FHS and BHS data to ensure the comparability of results. The TMT-A and TMT-B were reverse-scored so that higher scores indicated better performance. Because neuropsychological performance is highly influenced by cultural constructs

Table 1. Description of DANA Task Digital Measures

DANA task	Digital measures	Description
Code Substitution	Average response time	ART for all trials regardless of type (ie, practice or test), ms
	Mean response time correct	ART for all correct test trials, ms
Go/No-Go	Mean response time	ART for all test trials, ms
Simple Reaction Time	Cognitive efficiency	Cognitive efficiency value (measure of both speed and accuracy)

ART indicates average response time; and DANA, Defense Automated Neurocognitive Assessment.

and may result in the overrepresentation of lower and higher scores, we demographically corrected neuropsychological scores to yield more comparable indicators of cognitive function.[14,27] Neuropsychological test raw scores were standardized into Z scores corrected for age, sex, and race, then averaged to compute a GCS. For descriptive purposes, the GCS was divided into quartiles. The 2 middle quartiles were merged into 1 single group to allow for lower and upper tail descriptions resembling better and worse cognitive performance, given that all individuals were cognitively normal.

As a secondary analysis, domain-specific Z scores were created separately in both cohorts to examine the association between DANA task digital outcome measures and executive function abilities (Digit Span Backward and TMT-B), attention and processing speed (Digit Span Forward and TMT-A), and verbal/visual episodic memory (Immediate and Delayed Free Recall).

Sociodemographic Characteristics

Sociodemographic data were collected in accordance with established FHS and BHS protocols, respectively. Variables included age (years) at neuropsychological exam, sex (women, men), and education (less than high school, high school, some college, and college graduate). Information on the educational level was obtained from self-reported questionnaires. Given that race, income, geography, and other social determinants affect the quality of education and are not fully captured by education years alone,[28,29] educational attainment was gauged using a proxy of premorbid cognitive abilities. Specifically, a word reading test (Wide Range Achievement Test-3 for FHS participants and Wide Range Achievement-4 for BHS participants) was used, and the index score was converted into a Z score.

Digital Neuropsychological Measures DANA

The smartphone-based platform comprised a series of tasks divided into 3 assessment blocks, which required 15 to 20 minutes to finish. In Block 1, participants answered questions to confirm their ability to view stimuli on the phone screen and hear instructions clearly. Block 2 involved participants completing a series of cognitive tests and responding to open-ended questions, which were voice recorded. Block 3 aimed to minimize external influences on performance by asking participants about their experiences using smartphone applications. The current study focused on analyzing data from cognitive tests conducted in Block 2, which were collected at both study sites.

The DANA protocol was developed to capture changes in cognition over time and can be used in the clinic and remotely. The protocol can be self-administered on most mobile devices using a Health Insurance Portability and Accountability Act-compliant back-end cloud and is designed for high-frequency, high-sensitivity assessment of an individual over time. The DANA protocol has been compared with the standard Mini-Mental State Examination and was reported to be an electronic, mobile, repeatable, sensitive, and valid method of measuring cognition over time in patients who are depressed and undergoing electroconvulsive therapy treatment.[30] The smartphone-based app is compatible with Android and iPhone and consists of 6 cognitive tests, 3 of which were selected for harmonization across sites included in this study. These 3 tasks included the Code Substitution, the Go/No-Go, and the Simple Reaction Time tests. The Code Substitution test assesses visual scanning, attention, learning, immediate recall, and short-term memory.[31] The Go/No-Go test provides a measure of sustained attention and impulsivity, and the Simple Reaction time assesses the speed and accuracy of targets, omissions, and commissions.[31] All digital outcome measures are described in Table 1.

The vendor-derived measures selected are related to response time and trial iterations (parsed data) and metric data. Metric data are further derived and include averages, medians, and SDs, which were calculated from parsed data. The data collection protocol includes completing the set of tests every 3 months. Each exam is an assigned period where we request the participant to complete the tests (a set of trials) using the DANA smartphone application. Each exam produces 1 set of files for each test with different trials

Table 2. Participant Characteristics by GCS Subgroups in the Framingham Heart Study and Bogalusa Heart Study at Time of Neuropsychological Assessment (N=344)

Characteristic	Total	GCS 1 (lower quartile), n=89	GCS 2 (middle quartiles Q2–Q3), n=173	GCS 3 (upper quartile), n=82	*P* value
Framingham Heart Study (n=344)					
Age, y	56.72 (9.77)	56.27 (9.98)	57.21 (9.62)	56.15 (9.95)	0.64
Sex, men	155 (45.06)	40 (44.94)	84 (48.55)	31 (37.80)	0.30
Race and ethnicity					0.002
White	305 (88.66)	69 (77.53)	158 (91.33)	78 (95.12)	
Black	11 (3.20)	7 (7.87)	3 (1.73)	1 (1.22)	
Other*	28 (8.14)	13 (14.61)	12 (6.94)	3 (3.66)	
Education					<0.001
Less than high school	3 (0.87)	2 (2.25)	1 (0.58)	0 (0.00)	
High school	34 (9.88)	16 (17.98)	15 (8.67)	3 (3.66)	
Some college	54 (15.70)	16 (17.98)	32 (18.50)	6 (7.32)	
College graduate	253 (73.55)	55 (61.80)	125 (72.25)	73 (89.02)	
GCS Z score	−0.003 (−2.47 to 1.20)	−0.74 (−2.46 to −0.33)	0.05 (−0.32 to 0.41)	0.64 (0.41 to 1.19)	<0.001
Educational attainment†	0.004 (−2.90 to 1.70)	−0.73 (−2.90 to 1.38)	0.04 (−2.79 to 1.70)	0.68 (−0.88 to 1.70)	<0.001
Bogalusa Heart Study (n=58)					
Age, y	48.88 (4.4)	49.5 (4.9)	48.2 (4.5)	49.71 (3.7)	0.49
Sex, men	16 (28.07)	5 (38.5)	6 (20.0)	5 (35.71)	0.37
Race					0.71
White	50 (86.2)	13 (92.9)	25 (83.3)	12 (85.71)	
Black	8 (13.8)	1 (7.14)	5 (16.7)	2 (14.29)	
Education					0.01
Less than high school	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
High school	12 (21.05)	6 (46.15)	6 (20.0)	0 (0.0)	
Some college	12 (21.05)	2 (15.38)	8 (26.67)	2 (14.29)	
College graduate	33 (57.9)	5 (38.5)	16 (53.33)	12 (85.71)	
GCS Z score	0.47 (−0.84 to 1.23)	−0.07 (−0.84 to 0.24)	0.51 (0.26 to 0.79)	0.97 (0.83 to 1.23)	<0.001
Educational attainment†	0.60 (−0.98 to 1.33)	−0.13 (−0.98 to 0.94)	0.74 (−0.31 to 1.34)	1.09 (0.85 to 1.34)	<0.001

Data are presented as mean (SD) or median (interquartile range) for continuous measures, and n (%) for categorical measures. GCS is presented as Z score (range). GCS indicates Global Cognitive Score.

*Other includes Asian, Pacific Islander, Hispanic, other, prefer not to answer.

†Educational attainment proxy: Wide Range Achievement Test Z score (range).

corresponding to a single iteration of the specific cognitive test. Both cohorts from the BHS and FHS have a separate infrastructure pipeline but a shared database server where the data are kept separate in different databases. Using Structured Query Language, trial-level data are converted into test- and exam-level data as needed. This process rederives metric data provided by the vendor but allows us to achieve quality control, obtain additional metrics, and more consistency across cohorts. The Structured Query Language queries are available on request. For this study, we included only the first completed exam of each DANA test for analysis.

Statistical Analysis

Differences in participant characteristics by GCS subgroups were compared using 1-way ANOVA for continuous measures and Pearson χ^2 test for frequencies and percentages. Means and SDs of DANA-task digital scores were described by GCS subgroups; no statistical comparisons were made. Three GCS subgroups were created as follows: GCS group 1 (lowest quartile) indicating worse cognitive performance, GCS group 2, and GCS group 3 (highest quartile indicating better cognitive performance). Scatter plots were created to visually explore the concordance between GCS and each DANA task.
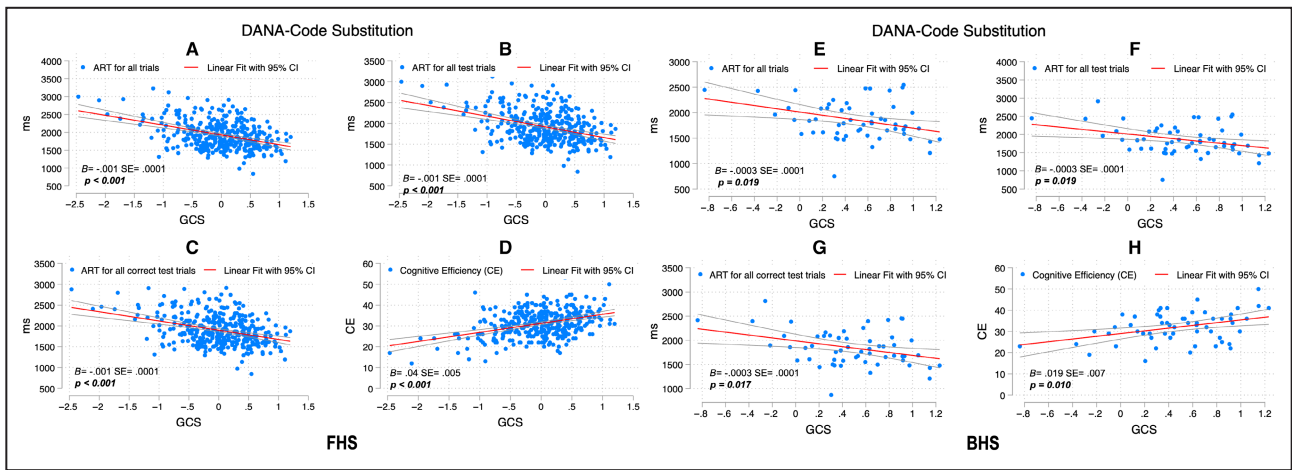
Figure 1. Correlation between Code Substitution digital scores and GCS among FHS and BHS participants.
SEs and *P* values are included for the association between the DANA measure and GCS adjusted for age and education. For FHS, figure parts indicate the correlation between ART for all trials and GCS (**A**), ART for all test trials and GCS (**B**), ART for all correct test trials and GCS (**C**), and CE and GCS (**D**); for BHS, ART for all trials and GCS (**E**), ART for all test trials and GCS (**F**), ART for all correct test trials and GCS (**G**), and CE and GCS (**H**). ART indicates average response time; B, unstandardized coefficients; BHS, Bogalusa Heart Study; CE, cognitive efficiency; DANA, Defense Automated Neurocognitive Assessment; FHS, Framingham Heart Study; and GCS, Global Cognitive Score.

Regression analyses included both unadjusted models and models adjusting for age and self-reported educational level to examine the association of individual DANA-task digital measures with GCS, and with domain-specific cognitive *Z* scores. Results from regression models for GCS were reported in each scatter plot, respectively. Additional models included educational attainment proxy and are included in the supplemental material. Statistical tests were 2-tailed, and *P* values <0.05 were considered significant. All analyses were performed using Stata/IC 15.1 (StataCorp, College Station, TX), and RStudio (version 3.0) software.

RESULTS

Sociodemographic Characteristics

A total of 344 participants were included from the FHS cohort. The mean age of FHS participants was 57 years (SD, 9.77), and 45% (n=155) were men. The BHS cohort included 58 participants with a mean age 49 years (SD, 4.4), and 28% (n=16) were men. A total of 90% (n=307) of FHS participants reported some college and college/graduate education, and 79% (45) of the BHS population had an educational level of at least some college and above (Table 2).
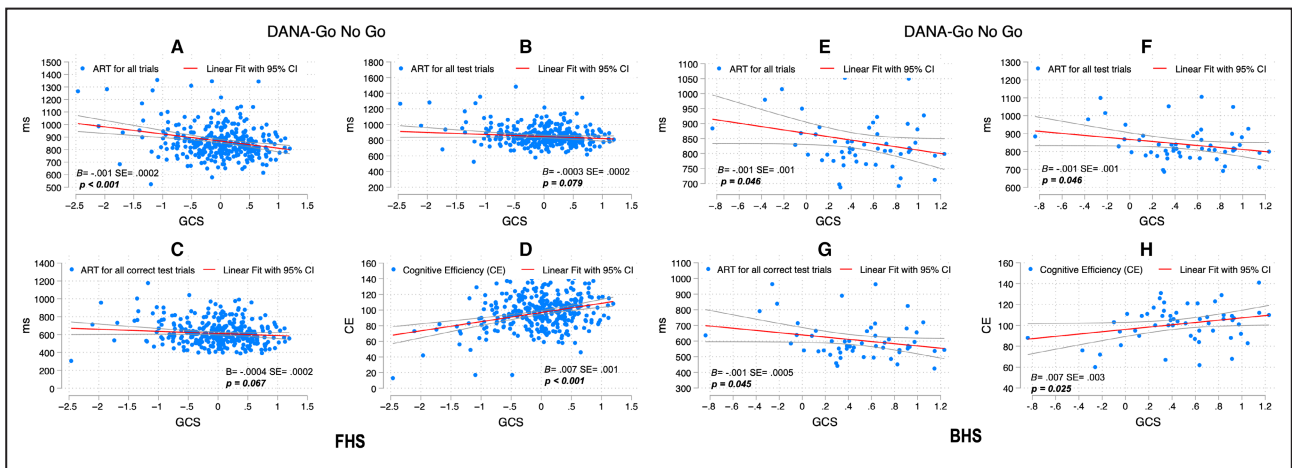

Figure 2. Go/No-Go scores and GCS among FHS and BHS participants.
SEs and *P* values are included for the association between the DANA measure and GCS adjusted for age and education. For FHS, figure parts indicate the correlation between ART for all trials and GCS (**A**), ART for all test trials and GCS (**B**), ART for all correct test trials and GCS (**C**), and CE and GCS (**D**); for BHS, ART for all trials and GCS (**E**), ART for all test trials and GCS (**F**), ART for all correct test trials and GCS (**G**), and CE and GCS (**H**). ART indicates average response time; B, unstandardized coefficients; BHS, Bogalusa Heart Study; CE, cognitive efficiency; DANA, Defense Automated Neurocognitive Assessment; FHS, Framingham Heart Study; and GCS, Global Cognitive Score.
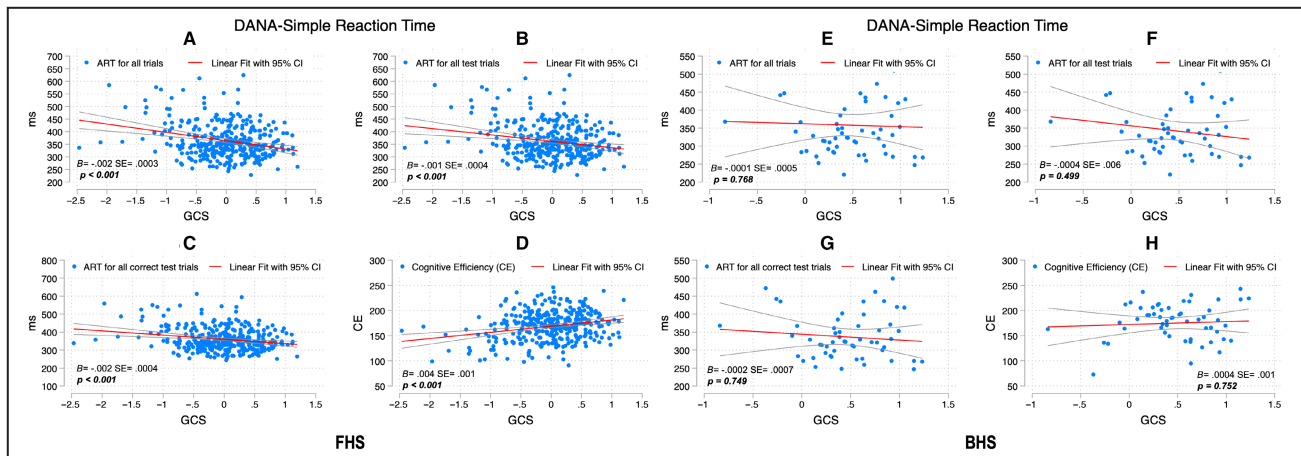
Figure 3. Simple Reaction Time scores and GCS among FHS and BHS participants.
SEs and *P* values are included for the association between the DANA measure and GCS adjusted for age and education. For FHS, figure parts indicate the correlation between ART for all trials and GCS (**A**), ART for all test trials and GCS (**B**), ART for all correct test trials and GCS (**C**), and CE and GCS (**D**); for BHS, ART for all trials and GCS (**E**), ART for all test trials and GCS (**F**), ART for all correct test trials and GCS (**G**), and CE and GCS (**H**). ART indicates average response time; B, unstandardized coefficients; BHS, Bogalusa Heart Study; CE, cognitive efficiency; DANA, Defense Automated Neurocognitive Assessment; FHS, Framingham Heart Study; and GCS, Global Cognitive Score.

There were no significant differences in age or sex across GCS subgroups in both the FHS (P=0.064 and P=0.30, respectively) and BHS (P=0.49 and P=0.37, respectively). Both cohorts showed significant differences in educational level across GCS subgroups (FHS: P<0.001; BHS: P=0.01), and in their estimated premorbid ability (FHS: P<0.001; BHS: P<0.001). In the BHS, only 39%[5] of individuals in GCS 1 reported college education compared with 86%[12] in GCS 3. In the FHS, 64% (65) in GCS 1 compared with 89% (90) of those in GCS 3 (P<0.001) reported a college education (Table 2).

Neuropsychological Global Cognitive Score and DANA Tasks

Among FHS participants, GCS ranged from −2.47 to 1.20 (mean *Z* score: −0.003±0.56); for BHS individuals, GCS ranged from −0.84 to 1.23 (mean *Z* score: 0.47±0.42). The GCS was categorized into 3 groups: lowest (GCS 1), middle (GCS 2), and highest (GCS 3). A total of 26% (n=89) of FHS participants were in GCS 1 corresponding to the lowest quartile of the sample; 50% (n=173) were in GSC 2, corresponding to the second and third quartile; and 24% (n=82) were in GCS 3, corresponding to the highest quartile. In the BHS, 26% (n=15) of participants had the lowest GCS (ie, GCS 1), 50% (n=29) were in GCS 2, and 24% (n=14) were in the highest GCS 3 (Table 2).

Participants with the lowest GCS (GCS 1) from both cohorts also had the lowest scores on digital DANA tasks (Table S1 and S2). Scatterplots supported this trend for both cohorts when exploring correlation patterns between GCS and DANA tasks (Figures 1 through 3).

Code Substitution

Higher GCS was significantly correlated with faster digital DANA scores measuring average response time (ART) for all trials ([FHS: B=−0.001, SE(±) 0.0001; P<0.001]; [BHS: B=−0.0003±0.0001; P=0.019]), and ART for all correct test trials ([FHS: B=−0.001±0.000; P<0.001]; [BHS: B=−0.0003±0.0001; P=0.017]), and ART for all test trials ([FHS: −0.001±0.0001; P<0.001]; [BHS: −0.0003±0.0001; P=0.019]), indicating better cognitive performance (Figure 1). Likewise, higher GCS correlated with higher cognitive efficiency (CE) scores ([FHS: B=0.04±0.005; P<0.001]; [BHS: B=0.019±0.007; P=0.010]), also indicating better performance (Figure 1).

Go/No-Go Task

A similar trend was observed with the exception of the ART for all test trials (FHS: B=−0.0004±0.0002; P=0.05) and the ART for all correct test trials (B=−0.0004±0.0002; P=0.077) in FHS; in BHS, only CE was statistically significant (0.007±0.003; P=0.025) (Figure 2).

Simple Reaction Time Test

Better GCS scores were associated with faster reaction time but only among FHS participants (Figure 3).

Additional models included the word reading *Z* score as a covariate (see Supplemental Material). The association of DANA tasks Code Substitution and Go/No-Go with GCS *Z* score remain significant after adjusting for age and word reading in both cohorts with the exception of the ART for all tests of the Go/No-Go task among FHS participants (−0.0004±0.0002; P=0.057) (Table S3).

Table 3. Association Between DANA Task Score and Neuropsychological Test Domain-Specific Scores in the Framingham Heart Study (N=344)

Task	Executive function		Attention and processing speed		Memory	
	Unadjusted	Adjusted*	Unadjusted	Adjusted*	Unadjusted	Adjusted*
	B±SE (*P* value)	B±SE (*P* value)	B±SE (*P* value)	B±SE (*P* value)	B±SE (*P* value)	B±SE (*P* value)
DANA task, Code Substitution						
ART for all trials	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.0005±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])
ART for all correct test trials	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.0007±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.001±0.0002 (<0.001[†])
Cognitive efficiency	0.04±0.01 (<0.001[†])	0.05±0.01 (<0.001[†])	0.03±0.01 (<0.001[†])	0.03±0.01 (<0.001[†])	0.04±0.01 (<0.001[†])	0.04±0.01 (<0.001[†])
ART for all test trials	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])	−0.001±0.0001 (<0.001[†])
DANA task, Go/No-Go						
ART for all trials	−0.001±0.0003 (0.01[†])	−0.001±0.0003 (0.02[†])	−0.0004±0.0003 (0.25)	−0.0003±0.0003 (0.39)	−0.001±0.0004 (0.003[†])	−0.001±0.0004 (0.007[†])
ART for all correct test trials	−0.00002±0.0003 (0.9)	−0.000003±0.0003 (0.8)	−0.00002±0.0003 (0.94)	−00002±0.0003 (0.95)	−0.0001±0.0003 (0.88)	−0.0001±0.0004 (0.86)
Cognitive efficiency	0.01±0.002 (<0.001[†])	0.01±0.002 (<0.001[†])	0.004±0.002 (0.04)	0.004±0.002 (0.06)	0.01±0.002 (0.004[†])	0.01±0.002 (0.01[†])
ART for all test trials	−0.0001±0.0003 (0.80)	−0.0001±0.0003 (0.67)	−8.23e-06±0.0003 (0.97)	−0.00003±0.0003 (0.91)	1.09e-06±0.0003 (0.99)	−0.00003±0.0003 (0.93)
DANA Task, Simple Reaction Time						
ART for all trials	−0.002±0.001 (<0.001[†])	−0.002±0.001 (0.001[†])	−0.001±0.001 (0.02[†])	−0.001±0.001 (0.03[†])	−0.002±0.001 (0.003[†])	−0.002±0.001 (0.007[†])
ART for all correct test trials	−0.002±0.001 (0.001[†])	−0.002±0.001 (0.003[†])	−0.001±0.001 (0.02[†])	−0.001±0.001 (0.04[†])	−0.001±0.001 (0.08)	−0.001±0.001 (0.22)
Cognitive efficiency	0.006±0.002 (<0.001[†])	0.006±0.002 (<0.001[†])	0.004±0.001 (0.01[†])	0.004±0.001 (0.02[†])	−0.004±0.002 (0.02[†])	−0.004±0.002 (0.04[†])
ART for all test trials	−0.002±0.001 (0.001[†])	−0.002±0.001 (0.005[†])	−0.001±0.001 (0.04[†])	−0.001±0.001 (0.09)	−0.002±0.001 (0.04[†])	−0.001±0.001 (0.12)

Executive function: Digit Span Backward and Trail Making Test-Part B. Attention and processing speed: Digit Span Forward and Trail Making Test-Part A. Memory: Logical Memory-Immediate Recall and Delayed Recall. ART indicates average response time; B, unstandardized coefficients; and DANA, Defense Automated Neurocognitive Assessment.

*Adjusted: covariates included age and education level.

[†]Statistically significant.

Neuropsychological Cognitive Domains and DANA Tasks

Code Substitution

Similar results were observed across both cohorts when examining the association between DANA measures and traditional paper and pencil neuropsychological domain indices. In the FHS cohort, all Code Substitution digital measures were significantly associated with executive function (ART for all trials: −0.001±0.0001; *P*<0.001; ART for all correct test trials: −0.001±0.0001; *P*<0.001; CE: 0.05±0.01; *P*<0.001; ART for all test trials: −0.001±0.0001; *P*<0.001), attention and processing speed (ART for all trials: −0.001±0.0001; *P*<0.001; ART for all correct test trials: −0.0007±0.0001; *P*<0.001; CE: 0.03±0.01; *P*<0.001; ART for all test trials: −0.001±0.0001; *P*<0.001), and memory (ART for all trials: −0.001±0.0001; *P*<0.001; ART for all correct test trials: −0.001±0.0002; *P*<0.001; CE: 0.04±0.01; *P*<0.001; ART for all test trials:

−0.001±0.0001; *P*<0.001), even after adjustment for age and education level (Table 3). In the BHS, similar trends were observed for attention and processing speed (ART for all trials: −0.001±0.0002; *P*=0.001; ART for all correct test trials: −0.001±0.0002; *P*=0.001; CE: 0.03±0.01; *P*=0.02; ART for all test trials: −0.001±0.0002; *P*=0.001) after adjustment for age and education level. For executive function, there was also a significant association (ART for all correct test trials: −0.001±0.0002; *P*=0.03; CE: 0.02±0.01; *P*=0.04; ART for all test trials: −0.0004±0.0002; *P*=0.04), but it did not remain significant after adjusting for age and education (Table 4).

Go/No-Go

In the FHS, the Go/No-Go ART for all trials and the cognitive efficiency digital measures were significantly associated with executive function (−0.001±0.0003; *P*=0.02 and 0.01±0.002; *P*<0.001, respectively] and memory *Z* scores

Table 4. Association Between Dana Task Score and Neuropsychological Test Domain-Specific Scores in the Bogalusa Heart Study (N=58)

Task	Executive function		Attention and processing speed		Memory	
	Unadjusted	Adjusted*	Unadjusted	Adjusted*	Unadjusted	Adjusted*
	B±SE (*P* value)	B±SE (*P* value)	B±SE (*P* value)	B±SE (*P* value)	B±SE (*P* value)	B±SE (*P* value)
DANA task, Code Substitution						
ART for all trials	−0.0004±0.0002 (0.05)	−0.0004±0.0002 (0.06)	−0.001±0.0002 (0.001)	−0.001±0.0002 (0.001)	−0.0001±0.0002 (0.6)	−0.0001±0.0002 (0.8)
ART for all correct test trials	−0.001±0.0002 (0.03)	−0.001±0.0002 (0.05)	−0.001±0.0002 (<0.001)	−0.001±0.0002 (0.001)	−0.0001±0.0002 (0.6)	−0.0001±0.0003 (0.8)
Cognitive efficiency	0.02±0.01 (0.04)	0.02±0.011 (0.07)	0.03±0.01 (0.01)	0.03±0.01 (0.02)	0.02±0.014 (0.3)	0.012±0.014 (0.4)
ART for all test trials	−0.0004±0.0002 (0.04)	−0.0004±0.0002 (0.06)	−0.001±0.0002 (0.001)	−0.001±0.0002 (0.001)	−0.0001±0.0002 (0.6)	−0.0001±0.0002 (0.7)
DANA task, Go/No-Go						
ART for all trials	−0.0003±0.001 (0.7)	−0.001±0.001 (0.5)	−0.002±0.001 (0.02†)	−0.002±0.001 (0.02†)	−0.001±0.001 (0.15)	−0.002±0.001 (0.14)
ART for all correct test trials	−0.0004±0.001 (0.6)	−0.001±0.001 (0.4)	−0.002±0.001 (0.02†)	−0.002±0.001 (0.01†)	−0.001±0.001 (0.21)	−0.001±0.001 (0.20)
Cognitive efficiency	0.003±0.005 (0.5)	0.01±0.005 (0.4)	0.01±0.05 (0.02)	0.01±0.005 (0.01)	0.01±0.006 (0.22)	0.01±0.01 (0.21)
ART for all test trials	−0.0003±0.001 (0.7)	−0.001±0.001 (0.6)	−0.002±0.001 (0.02†)	−0.002±0.001 (0.02†)	−0.002±0.001 (0.14)	−0.002±0.001 (0.14)
DANA task, Simple Reaction Time						
ART for all trials	−0.0004±0.001 (0.5)	−0.001±0.001 (0.5)	−0.001±0.001 (0.1)	−0.001±0.001 (0.1)	−0.0002±0.001 (0.8)	−0.0003±0.001 (0.7)
ART for all correct test trials	−0.001±0.001 (0.3)	−0.001±0.001 (0.4)	−0.001±0.001 (0.4)	−0.001±0.001 (0.5)	0.001±0.001 (0.5)	0.001±0.001 (0.4)
Cognitive efficiency	0.001±0.002 (0.5)	0.001±0.002 (0.5)	0.003±0.002 (0.1)	0.003±0.002 (0.11)	−0.0002±0.002 (0.9)	−0.0002±0.002 (0.9)
ART for all test trials	−0.001±0.001 (0.3)	−0.001±0.001 (0.4)	−0.001±0.001 (0.1)	−0.001±0.001 (0.13)	0.0003±0.001 (0.7)	0.001±0.001 (0.6)

Executive function: Digit Span Backward and Trail Making Test-Part B. Attention and Processing Speed: Digit Span Forward and Trail Making Test-Part A. Memory: Logical Memory–Immediate Recall and Delayed Recall and Memory Recognition. ART indicates average response time; B, unstandardized coefficients; and DANA, Defense Automated Neurocognitive Assessment.

*Adjusted: covariates included age and education level.

†Statistically significant.

[−0.001±0.0004; *P*=0.007 and 0.01±0.002; *P*=0.01, respectively) after adjusting for covariates (Table 3). In the BHS, all Go/No-Go digital measures were associated with attention and processing speed even in the adjusted model (ART for all trials: −0.002±0.001; *P*=0.02; ART for all correct test trials: −0.002±0.001; *P*=0.01; CE: 0.01±0.005; *P*=0.01; ART for all test trials: −0.002±0.001; *P*=0.02). No significant associations were found with executive function and memory *Z* scores (Table 4).

Simple Reaction Time

Lastly, when looking at Simple Reaction Time digital measures and cognitive domains, the association coefficients was similar across cohorts but not statistically significant among BHS individuals. In the FHS, all Simple Reaction Time digital measures were significantly

associated with executive function even after demographic adjustment (ART for all trials: −0.002±0.001; *P*=0.001; ART for all correct trials: −0.002±0.001; *P*=0.003; CE: 0.006±0.002; *P*<0.001; ART for all test trials: −0.002±0.001; *P*=0.005); to attention and processing speed (ART for all trials: −0.001±0.001; *P*=0.03; ART for all correct trials: −0.001±0.001; *P*=0.04; CE: 0.004±0.001; *P*=0.02) except for ART for all test trials (*P*=0.09, adjusted model). For memory, only ART for all trials and CE were associated after controlling for covariates (−0.002±0.001; *P*=0.007; −0.004±0.002 *P*=0.04, respectively) (Table 3).

DISCUSSION

The present study investigated the extent to which smartphone-based measures of cognitive performance

correlate with traditional paper and pencil neuropsychological test-derived GCS in 2 distinct community-based US cohorts. Additionally, we examined how the smartphone-based tests were associated with specific underlying cognitive domains derived from the neuropsychological protocols for each cohort. When examining the global cognitive patterns by quartiles, differences in education were found among the GCS groups, but there were no differences in age and sex. Overall, the findings revealed that performance on smartphone-based cognitive tasks in 2 geographically distinct cohorts was congruent with respect to global cognitive performance measured using a traditional paper and pencil neuropsychological protocol. These data suggest that digital outcomes measures, as described above, may offer an alternative approach to screening for possible cognitive impairment.

The associations between global cognition and each DANA task reveal some alignment in the trends for both cohorts (see Figures 1 through 3). The associations between DANA subtests and individual cognitive domains from the neuropsychological protocol were domain-specific for some DANA subtests within each cohort (see Tables 3 and 4). Specifically, the Code Substitution task, a measure of visual scanning, attention, learning, immediate recall, and short-term memory, correlated significantly with all 3 neuropsychological domains (executive function, attention and processing speed, and memory) from the FHS cohort, and with 2 domains but not memory in the BHS cohort. For the DANA task measuring Simple Reaction Time, a measure of speed and accuracy, all 3 neuropsychological domains were significantly associated in the FHS cohort, whereas none showed a significant association in the BHS cohort. Finally, for the Go/No-Go DANA task, a measure of sustained attention and impulsivity, selective associations were found for the domains of executive function and memory in the FHS cohort, whereas in the BHS cohort, the associations were present for the attention and processing speed domain. The difference in sample size and the demographic makeup between the 2 cohorts might account for some of these differences.

Overall, these findings provide evidence for the validity of the digital tests and suggest that the Code Substitution test may capture cognitive performance to a greater extent across multiple domains and diverse populations. This finding aligns with another study that found DANA subtests correlated with Mini-Mental State Examination in patients undergoing electroconvulsive therapy treatment.[30] Hollinger et al found that among all the DANA subtests, Code Substitution measures were most strongly associated with the Mini-Mental State Examination in their sample.[30] Thus, the findings from the current study lend further support for the potential usefulness of Code Substitution to measure subtle

impairment across diverse clinical groups. However, another study found that Code Substitution was not reliably completed by patients having an Alzheimer disease diagnosis.[32] Whereas the study by Hollinger et al conducted the DANA task with the participant in the inpatient unit, the study by Lathan et al administered the tasks remotely.[32] This could have accounted for differences in participants' ability to complete the Code Substitution task, wherein those with dementia may not have received additional prompts or help and therefore could not complete the task reliably. It could also be that deteriorating cognitive abilities in those with dementia might have posed special challenges in their ability to complete the task. Although additional work is required to understand under what context the tasks can be completed successfully, in the current study, all participants were cognitively normal and completed the tasks successfully in a remote, self-administered, and unsupervised manner.

Although in this study we did not formally examine the statistical differences in educational level between the FHS and BHS cohorts, there were differences in the proportion of individuals who completed a college degree (Table 2) that could be attributable to differences in access and quality of education experiences within these communities.[33] In the United States, educational quality varies across states and over time, with more pronounced differences between northern and southern parts of the country.[34,35] Literacy in adults has been measured using the reading score from the Wide Range Achievement Test, scores from a reading comprehension tests, and scores from writing tests.[29] There is evidence that Wide Range Achievement Test reading scores, when included as covariates, attenuates racial group differences on most neuropsychological tests between Black and White individuals matched on years of education.[29]

In the current study, the overall breakdown of education differed across the cohorts, with relatively more individuals from the BHS having high school and some college-level education. Yet, when accounting for the possible differences in performance on the word reading test, as a proxy for educational attainment, the association between DANA tasks, Code Substitution, Go/No-Go, and Simple Reaction Time digital scores remained significant among the FHS participants. Among the BHS individuals, there was also an association between Code Substitution and Go/No-Go tasks with the GCS Z score. After controlling for both educational level and performance on word reading, results indicate that education did not impact the association between the digital cognitive scores and traditional neuropsychological tests. We used both educational level and a proxy for gauging educational attainment to ensure that educational differences between groups such as the FHS and BHS did not account for varied

results. In this study, while deriving the quartiles for global neuropsychological performance, we also adjusted for sex, age, and race for each individual cohort. After this stage, it is possible to combine data in a harmonized manner while acknowledging that other contributors to cognitive test performance may not be fully controlled for in such pooled analysis. This approach may offer a way of pooling data when various cohorts are considered for research.

The similar pattern of results observed in FHS and BHS cohorts provide evidence for concurrent validity, suggesting that smartphone-based cognitive tests can assess aspects of cognitive abilities known to be associated with ADRDs. The 2 cohorts have different neuropsychological test protocols, thus making room for the argument that this could make generalizability across studies challenging. By deriving a global or composite cognitive score for each cohort, the differences in the test composition for each battery can be minimized, because all tests in the respective protocols are given equal importance. Using a GCS is common in ADRD research and enables the generation of comparable results.[36–38]

The study has limitations related to sample sizes for both cohorts, potentially impacting the findings. Ongoing data collection will seek to replicate these results with larger sample sizes for the FHS and BHS, and will include additional cohorts with similar data. Furthermore, replicating these findings in other diverse epidemiological cohorts will enhance their validity and generalizability of results. This study primarily focused on mapping congruence between smartphone-based cognitive test performance and a global composite score derived from an neuropsychological battery for each cohort.

Despite the limited sample size, the consistent patterns of associations in the 2 cohorts are encouraging, and future efforts to increase the sample will warrant a more in-depth discussion of the findings. Further research will examine the exact nature of associations and extend the study by investigating the predictive validity of these digital measures in identifying cognitive decline or dementia progression. Nevertheless, this study contributes to the growing body of research exploring innovative methods for cognitive assessment and monitoring. To our knowledge, this is one of the first studies to find congruence between smartphone-based DANA tasks and traditional paper and pencil neuropsychological tests in 2 large, community-based diverse cohorts. Other studies examining associations between digital cognitive tasks and traditional paper and pencil tests have generally found moderate to high associations between the 2.[39–42] A major benefit of DANA is that it can be self-administered remotely or in the clinic.[31,43] Recent articles have emphasized the need for urgently incorporating cognitive testing via digital testing in clinical practice using standardized procedures.[44,45] DANA is already Food and Drug Administration cleared and has shown promising psychometric properties. Because more studies use this application, the reliability and validity of in-home and remote assessment using DANA can be further established. Further evidence of different types of reliability and external validity across diverse populations will enable its use in both clinical trials and within clinical settings. A major issue to be tackled is ensuring that testing using digital technology is allowable for billing and reimbursement.[45]

Overall, this study presents valuable insights into the feasibility and potential usefulness of smartphone-based cognitive assessments in capturing cognitive performance in racially and ethnically diverse populations. Smartphone-based cognitive tasks may provide much-needed cognitive screening services to more individuals in an easier manner. The adoption of study designs encompassing digital approaches to collecting data and applying it to clinical, research, and community settings may enable findings with the potential for generalizability across broader populations.

award 20SFRN35490098, Alzheimer's Drug Discovery Foundation, and the Alzheimer's Association Fellowship award AARFD-23-1 150 584.

Disclosures

R.A. is a scientific advisor to Signant Health, and a consultant to Biogen and the Davos Alzheimer's Collaborative. She also serves as Director of the Global Cohort Development program for Davos Alzheimer's Collaborative. V.K. is a consultant to AstraZeneca. The remaining authors have no disclosures to report.

Supplemental Material

Tables S1–S3
Figures S1–S2

REFERENCES

1. 2022 Alzheimer's disease facts and figures. *Alzheimers Dement*. 2022;18:700–789. doi: 10.1002/alz.12638
2. Dementia. World Health Organization. Accessed September 15, 2023. https://www.who.int/news-room/fact-sheets/detail/dementia
3. 2017 National population projections tables: main series. United States Census Bureau. 2017. Accessed August 13, 2023. https://www.census.gov/data/tables/2017/demo/popproj/2017-summary-tables.html
4. Maclin JMA, Wang T, Xiao S. Biomarkers for the diagnosis of Alzheimer's disease, dementia Lewy body, frontotemporal dementia and vascular dementia. *Gen Psychiatr*. 2019;32:e100054. doi: 10.1136/gpsych-2019-100054
5. Teunissen CE, Verberk IMW, Thijssen EH, Vermunt L, Hansson O, Zetterberg H, van der Flier WM, Mielke MM, Del Campo M. Blood-based biomarkers for Alzheimer's disease: towards clinical implementation. *Lancet Neurol*. 2022;21:66–77. doi: 10.1016/S1474-4422(21)00361-6
6. Chouliaras L, O'Brien JT. The use of neuroimaging techniques in the early and differential diagnosis of dementia. [published online August 22, 2023]. *Mol Psychiatry*. doi: 10.1038/s41380-023-02215-8
7. Laske C, Sohrabi HR, Frost SM, López-de-Ipiña K, Garrard P, Buscema M, Dauwels J, Soekadar SR, Mueller S, Linnemann C, et al. Innovative diagnostic tools for early detection of Alzheimer's disease. *Alzheimers Dement*. 2015;11:561–578. doi: 10.1016/j.jalz.2014.06.004
8. Gold M, Amatniek J, Carrillo MC, Cedarbaum JM, Hendrix JA, Miller BB, Robillard JM, Rice JJ, Soares H, Tome MB, et al. Digital technologies as biomarkers, clinical outcomes assessment, and recruitment tools in Alzheimer's disease clinical trials. *Alzheimers Dement (N Y)*. 2018;4:234–242. doi: 10.1016/j.trci.2018.04.003
9. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, Holtzman DM, Jagust W, Jessen F, Karlawish J, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14:535–562.
10. Mattke S, Batie D, Chodosh J, Felten K, Flaherty E, Fowler NR, Kobylarz FA, O'Brien K, Paulsen R, Pohnert A, et al. Expanding the use of brief cognitive assessments to detect suspected early-stage cognitive impairment in primary care. *Alzheimers Dement*. 2023;19:4252–4259. doi: 10.1002/alz.13051
11. Porsteinsson AP, Isaacson RS, Knox S, Sabbagh MN, Rubino I. Diagnosis of early Alzheimer's disease: clinical practice in 2021. *J Prev Alz Dis*. 2021;8:371–386. doi: 10.14283/jpad.2021.23
12. Costa A, Bak T, Caffarra P, Caltagirone C, Ceccaldi M, Collette F, Crutch S, Della Sala S, Démonet JF, Dubois B, et al. The need for harmonisation and innovation of neuropsychological assessment in neurodegenerative dementias in Europe: consensus document of the Joint Program for Neurodegenerative Diseases Working Group. *Alzheimers Res Ther*. 2017;9:27. doi: 10.1186/s13195-017-0254-x
13. Rentz DM, Parra Rodriguez MA, Amariglio R, Stern Y, Sperling R, Ferris S. Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: a selective review. *Alzheimers Res Ther*. 2013;5:58. doi: 10.1186/alzrt222
14. Manly JJ, Miller SW, Heaton RK, Byrd D, Reilly J, Velasquez RJ, Saccuzzo DP, Grant I. The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals. The HIV Neurobehavioral Research Center (HNRC) Group. *J Int Neuropsychol Soc*. 1998;4:291–302. doi: 10.1017/S1355617798002914
15. Manly JJ. Critical issues in cultural neuropsychology: profit from diversity. *Neuropsychol Rev*. 2008;18:179–183. doi: 10.1007/s11065-008-9068-8
16. Loewenstein DA, Argüelles T, Argüelles S, Linn-Fuentes P. Potential cultural bias in the neuropsychological assessment of the older adult. *J Clin Exp Neuropsychol*. 1994;16:623–629. doi: 10.1080/01688639408402673
17. Brickman AM, Cabo R, Manly JJ. Ethical issues in cross-cultural neuropsychology. *Applied Neuropsychol*. 2006;13:91–100. doi: 10.1207/s15324826an1302_4
18. Koo BM, Vizer LM. Mobile technology for cognitive assessment of older adults: a scoping review. *Innov Aging*. 2019;3:igy038. doi: 10.1093/geroni/igy038
19. Au R, Piers RJ, Devine S. How technology is reshaping cognitive assessment: lessons from the Framingham Heart Study. *Neuropsychology*. 2017;31:846–861. doi: 10.1037/neu0000411
20. Öhman F, Berron D, Papp KV, Kern S, Skoog J, Hadarsson Bodin T, Zettergren A, Skoog I, Schöll M. Unsupervised mobile app-based cognitive testing in a population-based study of older adults born 1944. *Front Digit Health*. 2022;4:933265. doi: 10.3389/fdgth.2022.933265
21. Muller M, Grobbee DE, Aleman A, Bots M, van der Schouw YT. Cardiovascular disease and cognitive performance in middle-aged and elderly men. *Atherosclerosis*. 2007;190:143–149. doi: 10.1016/j.atherosclerosis.2006.01.005
22. Berenson GS. Bogalusa Heart Study: a long-term community study of a rural biracial (black/white) population. *Am J Med Sci*. 2001;322:267–274. doi: 10.1097/00000441-200111000-00007
23. Farmer ME, White LR, Kittner SJ, Kaplan E, Moes E, McNamara P, Wolz MM, Wolf PA, Feinleib M. Neuropsychological test performance in Framingham: a descriptive study. *Psychol Rep*. 1987;60:1023–1040. doi: 10.1177/0033294187060003-201.1
24. Au R, Seshadri S, Wolf PA, Elias M, Elias P, Sullivan L, Beiser A, D'Agostino RB. New norms for a new generation: cognitive performance in the Framingham Offspring cohort. *Exp Aging Res*. 2004;30:333–358. doi: 10.1080/03610730490484380
25. Satizabal CL, Beiser AS, Chouraki V, Chêne G, Dufouil C, Seshadri S. Incidence of dementia over three decades in the Framingham Heart Study. *N Engl J Med*. 2016;374:523–532. doi: 10.1056/NEJMoa1504327
26. Gershon RC, Cella D, Fox NA, Havlik RJ, Hendrie HC, Wagster MV. Assessment of neurological and behavioural function: the NIH Toolbox. *Lancet Neurol*. 2010;9:138–139. doi: 10.1016/S1474-4422(09)70335-7
27. Werry AE, Daniel M, Bergström B. Group differences in normal neuropsychological test performance for older non-Hispanic White and Black/African American adults. *Neuropsychology*. 2019;33:1089–1100. doi: 10.1037/neu0000579
28. Ahl RE, Beiser A, Seshadri S, Auerbach S, Wolf PA, Au R. Defining MCI in the Framingham Heart Study Offspring: education versus WRAT-based norms. *Alzheimer Dis Assoc Disord*. 2013;27:330–336. doi: 10.1097/WAD.0b013e31827bde32
29. Manly JJ, Jacobs DM, Touradji P, Small SA, Stern Y. Reading level attenuates differences in neuropsychological test performance between African American and White elders. *J Int Neuropsychol Soc*. 2002;8:341–348. doi: 10.1017/S1355617702813157
30. Hollinger KR, Woods SR, Adams-Clark A, Choi SY, Franke CL, Susukida R, Thompson C, Reti IM, Kaplin AI. Defense automated neurobehavioral assessment accurately measures cognition in patients undergoing electroconvulsive therapy for major depressive disorder. *J ECT*. 2018;34:14–20. doi: 10.1097/YCT.0000000000000448
31. Lathan C, Spira JL, Bleiberg J, Vice J, Tsao JW. Defense automated neurobehavioral assessment (DANA)—psychometric properties of a new field-deployable neurocognitive assessment tool. *Military Med*. 2013;178:365–371. doi: 10.7205/MILMED-D-12-00438
32. Lathan C, Coffma I, Shewbridge R, Lee M, Cirio R, Fonzetti P, Doraiswamy PM, Resnick HE. A pilot to investigate the feasibility of mobile cognitive assessment of elderly patients and caregivers in the home. *J Geriatr Palliat Care*. 2016;4:6. doi: 10.13188/2373-1133.1000017
33. Glymour MM, Manly JJ. Lifecourse social conditions and racial and ethnic patterns of cognitive aging. *Neuropsychol Rev*. 2008;18:223–254. doi: 10.1007/s11065-008-9064-z
34. Sisco S, Gross AL, Shih RA, Sachs BC, Glymour MM, Bangen KJ, Benitez A, Skinner J, Schneider BC, Manly JJ. The role of early-life educational quality and literacy in explaining racial disparities in cognition in late life. *J Gerontol B Psychol Sci Soc Sci*. 2015;70:557–567. doi: 10.1093/geronb/gbt133
35. Berkman LF, Glymour MM. How society shapes aging: the centrality of variability. *Daedalus*. 2006;135:105–114. doi: 10.1162/001152606775321103

36. Malek-Ahmadi M, Chen K, Perez SE, He A, Mufson EJ. Cognitive composite score association with Alzheimer's disease plaque and tangle pathology. *Alz Res Ther*. 2018;10:90. doi: 10.1186/s13195-018-0401-z

37. Schneider LS, Sano M. Current Alzheimer's disease clinical trials: methods and placebo outcomes. *Alzheimers Dement*. 2009;5:388–397. doi: 10.1016/j.jalz.2009.07.038

38. Schneider LS, Mangialasche F, Andreasen N, Feldman H, Giacobini E, Jones R, Mantua V, Mecocci P, Pani L, Winblad B, et al. Clinical trials and late-stage drug development for Alzheimer's disease: an appraisal from 1984 to 2014. *J Intern Med*. 2014;275:251–283. doi: 10.1111/joim.12191

39. Belleville S, LaPlume AA, Purkart R. Web-based cognitive assessment in older adults: where do we stand? *Curr Opin Neurol*. 2023;36:491–497. doi: 10.1097/WCO.0000000000001192

40. Stricker NH, Twohy EL, Albertson SM, Karstens AJ, Kremers WK, Machulda MM, Fields JA, Jack CR Jr, Knopman DS, Mielke MM, et al. Mayo-PACC: a parsimonious preclinical Alzheimer's disease cognitive composite comprised of public-domain measures to facilitate clinical translation. *Alzheimers Dement*. 2023;19:2575–2584. doi: 10.1002/alz.12895

41. Brewster PWH, Rush J, Ozen L, Vendittelli R, Hofer SM. Feasibility and psychometric integrity of mobile phone-based intensive measurement of cognition in older adults. *Exp Aging Res*. 2021;47:303–321. doi: 10.1080/0361073X.2021.1894072

42. Morrison RL, Pei H, Novak G, Kaufer DI, Welsh-Bohmer KA, Ruhmel S, Narayan VA. A computerized, self-administered test of verbal episodic memory in elderly patients with mild cognitive impairment and healthy participants: a randomized, crossover, validation study. *Alzheimers Dement (Amst)*. 2018;10:647–656. doi: 10.1016/j.dadm.2018.08.010

43. DANA Brain Vital. Accessed November 29, 2023. https://danabrainvital.com/

44. Fox-Fuller JT, Rizer S, Andersen SL, Sunderaraman P. Survey findings about the experiences, challenges, and practical advice/solutions regarding teleneuropsychological assessment in adults. *Arch Clin Neuropsychol*. 2022;37:274–291. doi: 10.1093/arclin/acab076

45. Sperling SA, Acheson SK, Fox-Fuller J, Colvin MK, Harder L, Cullum CM, Randolph JJ, Carter KR, Espe-Pfeifer P, Lacritz LH, et al. Teleneuropsychology: from science to policy to practice. [published online September 15, 2023]. *Arch Clin Neuropsychol*. doi: 10.1093/arclin/acad066