Title

An Automated Virtual Reality Program Accurately Diagnoses HIV-Associated Neurocognitive Disorders in Older People With HIV

Journal

Open Forum Infectious Diseases, 10(12)

ISSN

2328-8957

Authors

Moore, Raeanne C

Kuehn, Kevin S

Heaton, Anne

et al.

Publication Date

2023-12-01

DOI

10.1093/ofid/ofad592

Peer reviewed

Open Forum Infectious Diseases

MAJOR ARTICLE

An Automated Virtual Reality Program Accurately Diagnoses HIV-Associated Neurocognitive Disorders in Older People With HIV

Raeanne C. Moore,[1] Kevin S. Kuehn,[1] Anne Heaton,[1] Erin E. Sundermann,[1] Laura M. Campbell,[1,2] Peter Torre,[3] Anya Umlauf,[1] David J. Moore,[1] Nicole Kosoris,[4] David W. Wright,[5] Michelle C. LaPlaca,[6] Drenna Waldrop,[7] and Albert M. Anderson[8]

[1]Department of Psychiatry, School of Medicine, University of California at San Diego, La Jolla, California, USA, [2]University of California San Diego Joint Doctoral Program in Clinical Psychology, San Diego State University, San Diego, California, USA, [3]San Diego State University, San Diego, California, USA, [4]Georgia Tech Research Institute, Atlanta, Georgia, USA, [5]Department of Emergency Medicine, School of Medicine, Emory University, Atlanta, Georgia, USA, [6]Department of Biomedical Engineering, Georgia Tech and Emory University, Atlanta, Georgia, USA, [7]Nell Hodgson Woodruff School of Nursing, Emory University, Atlanta, Georgia, USA, and [8]Division of Infectious Diseases, Department of Medicine, School of Medicine, Emory University, Atlanta, Georgia, USA

Background. HIV-associated neurocognitive disorders (HANDs) remain prevalent despite antiretroviral therapy, particularly among older people with HIV (PWH). However, the diagnosis of HAND is labor intensive and requires expertise to administer neuropsychological tests. Our prior pilot work established the feasibility and accuracy of a computerized self-administered virtual reality program (DETECT; Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury) to measure cognition in younger PWH. The present study expands this to a larger sample of older PWH.

Methods. We enrolled PWH who were ≥60 years old, were undergoing antiretroviral therapy, had undetectable plasma viral loads, and were without significant neuropsychological confounds. HAND status was determined via Frascati criteria. Regression models that controlled for demographic differences (age, sex, education, race/ethnicity) examined the association between DETECT's cognition module and both HAND status and Global Deficit Score (GDS) derived via traditional neuropsychological tests.

Results. Seventy-nine PWH (mean age, 66 years; 28% women) completed a comprehensive neuropsychological battery and DETECT's cognition module. Twenty-five (32%) had HAND based on the comprehensive battery. A significant correlation was found between the DETECT cognition module and the neuropsychological battery ($r = 0.45$, $P < .001$). Furthermore, in two separate regression models, HAND status ($b = -0.79$, $P < .001$) and GDS impairment status ($b = -0.83$, $P < .001$) significantly predicted DETECT performance. Areas under the curve for DETECT were 0.78 for differentiating participants by HAND status (HAND vs no HAND) and 0.85 for detecting GDS impairment.

Conclusions. The DETECT cognition module provides a novel means to identify cognitive impairment in older PWH. As DETECT is fully immersive and self-administered, this virtual reality tool holds promise as a scalable cognitive screening battery.

Keywords. cognition; cognitive aging; cognitive screening; digital health; HIV; virtual reality.

While people with HIV (PWH) are surviving significantly longer in the era of combination antiretroviral therapy (ART) [1, 2], HIV-associated neurocognitive disorders (HAND) remains highly prevalent. Large cohort studies have shown that HAND rates continue to be high among PWH undergoing ART even when confounding comorbidities are excluded [3, 4]. A meta-analysis published in 2020 found a worldwide HAND prevalence of 42.6% (95% CI, 39.7%–45.5%) [5]. In a subset of 51 studies in which >90% of participants were receiving ART, the prevalence of HAND remained essentially constant at 43.1% (95% CI, 38.1%–48.2%). Diminished cognition has been associated with multiple adverse outcomes among persons with HIV. Worse cognition can predict worse ART medication adherence [6], which is known to place PWH at higher risk for clinical complications as well as HIV transmission [7, 8]. Furthermore, cognitive impairment in the setting of HIV is associated with worse quality of life [9] and higher mortality risk [10, 11].

The effects of aging on the brain appear to be particularly significant in PWH. Magnetic resonance imaging studies demonstrate accelerated aging based on brain integrity among PWH despite virologic suppression [12]. Cognitive performance in PWH aged 50 to 65 years is worse than in age-matched people without HIV but similar to people without HIV aged ≥65 years [13]. Older age (>50 years) among PWH is associated with 4.8-fold higher odds of memory impairment as compared with people without HIV [14, 15].

However, current protocols to fully evaluate cognition in PWH require traditional neuropsychological testing. While

Received 20 July 2023; editorial decision 16 November 2023; accepted 27 November 2023; published online 29 November 2023

Correspondence: Raeanne C. Moore, PhD, Department of Psychiatry, University of California San Diego, 220 Dickinson St, Suite B, San Diego, CA 92103 (r6moore@health.ucsd.edu).

Open Forum Infectious Diseases®

https://doi.org/10.1093/ofid/ofad592

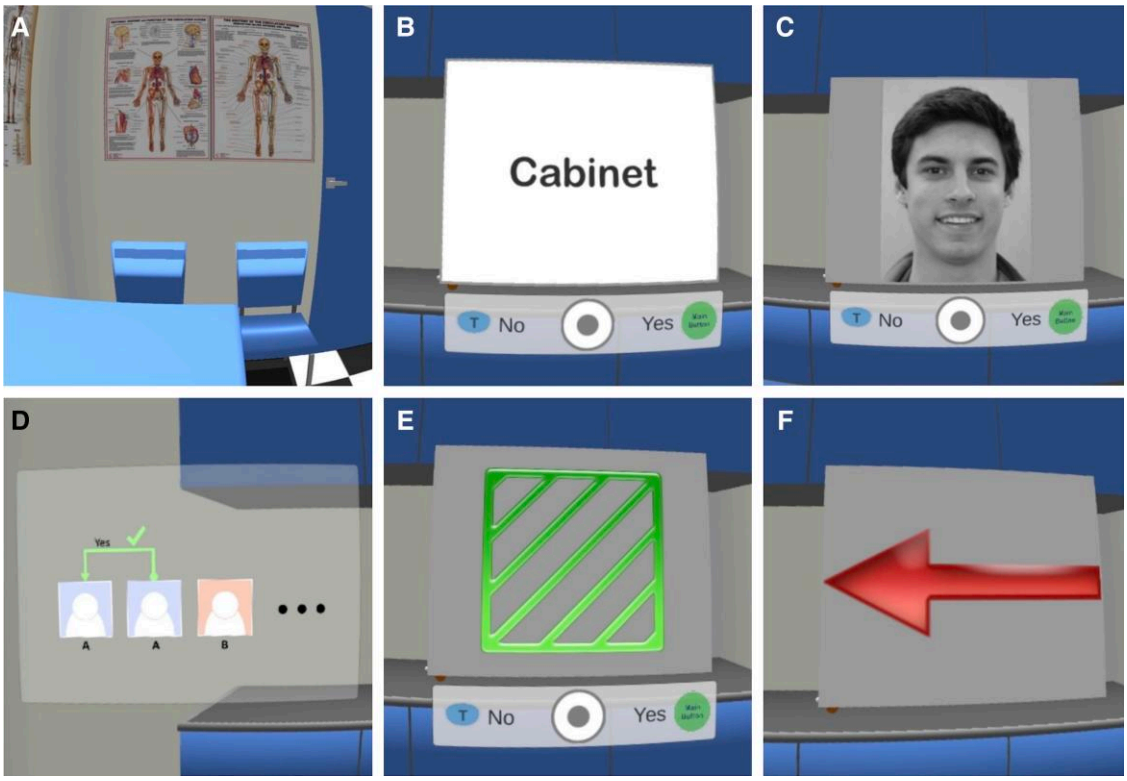

Figure 1. Sample images from DETECT. (*A*) Image of the virtual reality examination room. Sample test stimuli: (*B*) Word Memory, (*C*) N-Back Faces (face), (*D*) N-Back Faces (instructions), (*E*) Shape Comparison (complex), and (*F*) Arrows Test. DETECT, Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury.

traditional testing provides comprehensive information, it is time intensive (typically at least 2 hours) and requires an experienced examiner [16]. Increased access to reliable and valid cognitive testing is needed, particularly among older PWH, who are also at risk for neurodegenerative disorders such as Alzheimer disease. We previously established the feasibility and accuracy of the DETECT program (Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury) as a method to evaluate cognition in a relatively small sample of PWH with an age range of 24 to 59 years [17]. DETECT is a brief automated virtual reality (VR) program designed at Emory University and Georgia Tech that is a novel self-administered approach to evaluate cognition. The program has 7 tests (Figure 1) that cover multiple domains, including immediate and delayed recognition memory, working memory, processing speed, and executive function. The program resides on a smartphone, which is inserted into a VR headset that simulates an examination room environment. The program runs automatically with verbal directions, which obviates the need for an examiner.

In this study, our goal was to expand our prior work to older PWH who are at higher risk for HAND (HAND vs no HAND). The first aim was to examine the convergent validity of the DETECT cognitive module with a comprehensive traditional neuropsychological battery. We predicted that DETECT scores would correlate closely with traditional neuropsychological testing in the sample, as it has in prior studies [17–19]. The second aim was to examine the sensitivity and specificity of DETECT for identifying those with and without HAND. Last, the third aim of this study was to examine the relationships of DETECT cognition scores with demographic characteristics (age, sex, education, race/ethnicity) and functional independence. We hypothesized that poorer performance on DETECT would be associated with older age, fewer years of education, and greater dependence on performing instrumental activities of daily living (IADL).

METHODS

Participants

This study—called "DETECT: A Novel Device to Assess How HIV Affects Neurocognitive Decline and Postural Instability in Older Adults at Risk for Alzheimer's Disease"—is a multisite longitudinal study conducted at the HIV Neurobehavioral Research Program (HNRP) at the University of California San Diego (UCSD) and at the Emory University Center for AIDS Research. Participants were recruited between September 2019 and December 2022. The baseline visit data

were analyzed in the current study. Participants enrolled at UCSD were recruited from prior and ongoing studies at the HNRP and through local outreach in the San Diego community. Participants enrolled at Emory were recruited from clinics affiliated with the Emory Center for AIDS Research and through local outreach in the Atlanta metropolitan area.

The parent study enrolled persons with and without HIV. For this study, inclusion criteria included age ≥60 years, diagnosis of HIV with combination ART and HIV RNA <200 copies/mL for at least 6 months (virologically suppressed per the guidelines of the US Department of Health and Human Services), English fluency, and ability to provide written informed consent. Exclusion criteria were as follows: probable dementia (score ≤10 on the HIV Dementia Scale [HDS]; due to concerns about performing the VR assessment), persons in hospice who could not be followed longitudinally, plans to move out of the area within the following 3 years, and neurologic confounds unrelated to HIV that could confound the HAND diagnosis: stroke, head injury with loss of consciousness >30 minutes and cognitive sequelae (eg, memory deficits and sleep disturbances; deficits in attention that can be attributed to several factors, such as traumatic brain injury severity, complications, and chronicity of the brain injury), history of central nervous system opportunistic infection, and serious mental illness (eg, schizophrenia or bipolar disorder). Individuals were also excluded if they had significant visual or hearing impairments that would impede study assessments.

Participant Consent Statement

This study was approved by each institution's institutional review board (UCSD and Emory), and all participants provided written informed consent.

Procedures

Participants completed a comprehensive neuropsychological evaluation, a neuromedical interview, the DETECT VR assessment, a blood draw, and an audiology examination (audiology data presented elsewhere [20]). The HNRP gold standard neuropsychological battery of tests (Table 1) assessed 7 neurocognitive domains: verbal fluency, speed of information processing, attention/working memory, executive functioning, learning, memory, and motor skills [21]. At the HNRP, modifications to administer some of the tests remotely via video conferencing were made to accommodate testing during the COVID-19 pandemic [22]. At Emory, all testing was completed in person throughout the COVID-19 pandemic. At both study sites, the order was always the same such that DETECT was administered prior to the traditional neuropsychological battery on the day of testing. At UCSD, the subset of assessments administered remotely was completed on a different day, not immediately following the administration of DETECT but always after it. This decision was made for 2 reasons: (1) so that exposure to the DETECT tests was completed prior to exposure to the other neuropsychological tests that day, which could have priming or interference effects, and (2) to minimize the impact of fatigue on the DETECT performance. The neuromedical interview included the collection of self-reported medical history and HIV disease characteristics (eg, estimated duration of HIV disease, nadir CD4 count, antiretroviral history). Plasma HIV RNA was confirmed to be <200 copies/mL at baseline.

Table 1. Neuropsychological Tests by Neurocognitive Domain

Verbal fluency	**Executive functioning**
Controlled Oral Word Association Test: F-A-S	Wisconsin Card Sorting Test: computerized 64-card version
Category fluency: animals/actions	Trail Making Test: part B
Speed of information processing	Stroop Color-Word Trial
WAIS-III: digit symbol	**Learning**
WAIS-III: symbol search	HVLT-R: immediate recall
Trail Making Test: part A	BVMT-R: immediate recall
Stroop Color Trial	**Memory**
Attention/working memory	HVLT-R: delayed recall
WAIS-III: letter-number sequencing	BVMT-R: delayed recall
Paced Auditory Serial Addition Task: 50-item single-trial version	**Motor** Grooved Pegboard

Abbreviations: BVMT-R, Brief Visuospatial Memory Test–Revised; HVLT-R, Hopkins Verbal Learning Test–Revised; WAIS-III, Wechsler Adult Intelligence Scale–Third Edition.

Neurocognitive Impairment

Raw scores on the neuropsychological battery were first converted to scaled scores on a standardized scale (mean = 10, SD = 3). Scaled scores were then converted to *T* scores (mean = 50, SD = 10), adjusted to correct for age, sex, education, and race/ethnicity [21, 23, 24]. *T* scores were averaged across tests within a domain to obtain domain *T* scores. *T* scores were then converted into continuously distributed deficit scores, which ranged from 0 (corresponding to a *T* score >39; no impairment) to 5 (*T* score <20; severe impairment) and were averaged across all tests to derive domain-based scores and Global Deficit Scores (GDS). The GDS weights the neuropsychological data in a similar manner to clinical judgments, by reflecting the frequency and severity of neurocognitive impairments across the entire test battery [25]. Impaired neurocognitive performance was defined by a GDS cut point ≥0.50, and participants are classified as GDS impaired vs not GDS impaired. A continuous GDS score was also used in some analyses (see Analytic Plan section). Scores account for practice effects for participants with prior administrations of the battery [26].

Participants were classified as having HAND if they met Frascati criteria for asymptomatic neurocognitive impairment (ANI) or mild neurocognitive disorder [16]. No participants met criteria for HIV-associated dementia. ANI is defined by impaired performance (demographically adjusted normative scores that fall at least 1 SD below the mean) in at least 2 cognitive domains. A diagnosis of mild neurocognitive disorder requires the same impaired cognitive performance as ANI but

with the presence of functional impairment in everyday living [14]. Functional dependence vs independence was assessed via a modified version of the IADL scale [27]. To meet criteria, these diagnoses of HAND require that impairment be attributed to HIV, at least in part, and severely confounding comorbidities not be present (eg, learning disability, stroke, or non–HIV-related central nervous system infections) [14]. As outlined in the exclusion criteria for the current study, persons with such non–HIV-related conditions were not enrolled into this study. While conditions such as anemia and chronic kidney disease are associated with worse cognition [28, 29], these conditions and other medical confounds are classified as mild-moderately confounding conditions and do not preclude diagnoses of HAND [14].

DETECT Evaluation

The current version of the DETECT custom software was created jointly by the Georgia Institute of Technology and Emory University. For the administration of DETECT, we used the commercially available SAMSUNG Gear VR SM-R325 with an Oculus controller. The device is compatible with multiple Samsung Galaxy smartphone models, including the Galaxy S8, which was utilized for this study. The DETECT program is administered with a VR headset, which simulates in-room clinic conditions that individuals experience during traditional neuropsychological evaluation and clinical visits. The program is immersive when used with noise-canceling headphones; for this study, which was conducted in a private testing room, headphones were not used. The system is designed such that participants with prescription glasses can easily wear them inside the headset.

The self-administered DETECT cognitive module takes approximately 15 minutes to complete and includes 7 tests: Word Memory (immediate and delayed recall subtests), N-Back Faces (1- and 2-back subtests), Shape Comparison (simple and complex shape subtests), and Arrows Test.

Word Memory: Learning and Delayed Recall Subtests

Domains assessed: immediate and delayed recognition memory. Participants were shown 12 target words, with each word randomly assigned at each administration from >700 possible words, representing 65 possible word categories. Next, participants were shown 24 words 1 at a time for 3 seconds each. Twelve words were from the original word list, and the other 12 were randomized distracter words that matched the same word category as 1 of the target words. Participants were instructed to respond to each word by pressing the *yes* button if the word was from the original list or *no* if it was not from the original list. Participants were given 10 seconds to respond to each word. After the immediate recall trial, participants proceeded to the other neuropsychological tests. Once the other tests were completed (approximately 10 minutes later), participants completed the delayed recall trial in which 24 words were presented in a random order, 1 at a time for 2 seconds each. Twelve words were from the original target words presented at the beginning of the evaluation, and 12 were randomized distracter words that matched the same category as 1 of the target words. For the immediate and delayed recall, responses were coded as correct (true positive or negative), incorrect (false positive or negative), or no response (no response within 10 seconds).

N-Back Faces: 1- and 2-Back Subtests

Domain assessed: working memory. Participants were shown black-and-white photographs of actual faces, 1 at a time for 2 seconds each. When presented with each face (not counting the first), participants were instructed to press the *yes* button if the face presented is the same as the face presented just before (ie, 1 back) or *no* if not the same as the face presented just before. Fourteen face trials were recorded. This test of working memory requires participants to continuously change their frames of reference to which each new face would be compared.

In the 2-back test, participants were presented with faces in the same manner as for the 1-back test (1 at a time for 2 seconds each), but this time they were instructed to compare each new face to the 1 presented 2 faces ago (ie, 2 back). Thirteen face trials were recorded for the 2-back test. For the 1- and 2-back tests, participants were given 2 seconds to respond to each image.

Shape Comparison: Simple and Complex Subtests

Domain assessed: processing speed. Participants were first shown a target shape to which they were to compare subsequent shapes. For each shape presented after the target shape, participants were instructed to press the *yes* button if it exactly matched the target shape or *no* if it did not exactly match. Participants were given 3 seconds to respond to each image. The simple shape subtest consisted of 20 trials, in which all shapes were a gray circle, triangle, square, or diamond.

In the complex shape subtest (test of working memory), a target image was again presented, to which participants were to compare against the subsequent images. However, now each image had a particular shape, color, and line orientation. For example, each image could be 1 of 4 possible shapes (circle, triangle, square, or diamond), in 1 of 5 colors (blue, red, purple, green, or black), with parallel lines filling the shape in 1 of 3 orientations (lines in the horizontal, vertical, or diagonal direction). Participants were instructed to press the *yes* button if each image exactly matched the target image in its color, shape, and line orientation or *no* button if it did not exactly match. Participants were given 3 seconds to respond to each image; there were 24 trials in the complex shape subtest.

Arrows Test

Domain assessed: executive function. In each trial, participants were presented with a red or blue arrow, pointing to the right

or left. If the arrow was blue, participants were instructed to press the *yes* button if it was pointing to the right and *no* if to the left. Conversely, if the arrow was red, participants were instructed to press the *yes* button if it was pointing to the left and *no* if to the right. Participants were given 3 seconds to respond to each arrow, presented 1 at a time for a total of 20 trials.

DETECT Test Scoring

The DETECT software recorded the response and reaction time of each trial for all tests. Participants were allowed 2, 3, or 10 seconds to respond to each trial (depending on the test); if they did not press either button within that time limit, the trial timed out and coded as *no response*. For each test and subtest, 2 primary summary scores were calculated: total and response percentage correct.

$$\textit{Total percentage correct} = \frac{\textit{number of trials with correct response}}{\textit{total number of trials}}.$$

$$\textit{Response percentage correct} = \frac{\textit{number of trials with correct response}}{\textit{number of trials with a response}}$$

VR Usability

Participants completed a feedback questionnaire after the VR evaluation (Supplementary Table 1), which included items asking them to assess difficulties with operating the controller (yes/no), using the headset (yes/no), viewing the VR screen (yes/no), as well as understanding the instructions for the 4 cognitive tests (yes/no for each item). To calculate a score of usability, we summed the first 3 items and recoded difficulties in understanding test instructions to 0 (no difficulties on any test), 1 (difficulty on 1 test), and 2 (difficulties on ≥2 tests). We then summed the first 3 items with this recoded variable to create a composite measure of usability (range, 0–5). Additionally, participants were asked if they "felt motion sickness when completing the VR assessments (eg, nausea, dizziness)."

Analytic Plan

We first controlled for the influence of age, sex, race, ethnicity, and education level by partialing out the variance from these variables in 2 separate regression models predicting DETECT performance and the global mean scaled scores from the traditional neuropsychological battery and subsequently extracting the residuals. We then generated Pearson *R* values to test for correlations between total and response percentage correct from DETECT with the global mean scaled scores (after accounting for demographic differences). Next, we used linear regression to test for relationships between DETECT performance and HAND status, GDS, GDS impairment status, and daily functioning (Lawton and Brody IADL scale). We also examined the area under the curve (AUC) and sensitivity and specificity of DETECT in differentiating these clinical characteristics. Finally, we examined demographic differences in DETECT performance (without demographic variables partialed out). We first used linear regression for the analyses of demographic relationships; however, as the residuals were not normally distributed, we relied on nonparametric models. For these, we used either Kendall–Theil Sen Siegel nonparametric linear regressions for continuous outcomes or generalized additive models for binary/nominal variables. There were no missing data in any of the variables used for the present analyses.

RESULTS

Descriptive Statistics

Demographic and clinical characteristics are presented in Table 2. There were 79 PWH with a mean age of 66 years (SD, 5). Twenty-eight percent of the participants were women, and the sample was racially and ethnically diverse. A significant percentage (65%) of the participants had a history of AIDS, while 25 (32%) met criteria for HAND. The median estimated duration of HIV disease was 29 years.

Aim 1: Convergent Validity

We used Pearson *R* to test for correlations among the total and response percentage correct from DETECT and the global mean scaled score from the neuropsychological battery. These results are presented in Table 3 and Figure 2. Total and response percentage correct values were moderately to highly correlated ($r = 0.72$; 95% CI, .59–.81; $P < .001$). The zero-order correlation between total percentage correct from DETECT and the global mean scaled scores derived from the neuropsychological battery was 0.56 (95% CI, .38–.69; $P < .001$), while the correlation between response percentage correct and the scaled scores was 0.57 (95% CI, .40–.70; $P < .001$). After adjusting DETECT and global mean scaled scores for age, sex, ethnicity, race, and education, the correlations between the scores and the total percentage correct ($r = 0.45$; 95% CI, .26–.61; $P < .001$) and response percentage correct ($r = 0.52$; 95% CI, .34–.66; $P < .001$) remained significant. A correlation matrix for DETECT total percentage correct with domain scores from the neuropsychological test battery is presented in Figure 3.

Aim 2: Discriminant Validity

We then tested for differences based on clinical and neuropsychological characteristics in the total percentage of correct responses from DETECT, after partialing out the variance explained by age, sex, ethnicity, race, and education. Results from these analyses are presented in Tables 4 and 5. In linear regression models, HAND

Table 2. Demographic and Clinical Characteristics of the Sample (79 People With HIV)

	No. (%), Mean ± SD, or Median [IQR]
Demographics	
Age, y (range)	66.5 ± 4.8 (60–80)
Male	57 (72.2)
Race/ethnicity	
Non-Hispanic White	33 (41.8)
African American/Black	35 (44.4)
Hispanic/Latinx	8 (10.1)
Other	3 (3.8)
Education, y	14.2 ± 3.0
Medical and psychiatric comorbidities	
Charlson Comorbidity Index[a]	7.8 ± 4.3
Lifetime Major Depressive Disorder	37 (46.8)
Current Substance Use Dependence	0 (0)
Cognitive and everyday functioning	
Scaled scores	
Global	7.9 ± 1.9
Verbal fluency	8.6 ± 2.6
Executive function	7.6 ± 2.3
Speed of information processing	8.8 ± 2.3
Learning	7.4 ± 2.5
Delayed recall	7.7 ± 2.6
Working memory	8.4 ± 2.5
IADL: dependent	16 (20.3)
DETECT, % correct	
Total	64.2 ± 12.5
Response	81.8 ± 8.8
HAND	25 (31.7)
HIV characteristics	
AIDS history	51 (64.6)
CD4	
Current	671.6 [406.4–936.9]
Nadir	119.5 [0–367.5]
Estimated duration of HIV disease, y	29 [16.5–41.5]
Undergoing ART	79 (100)
Viral load <200	72 (100)[b]

Abbreviations: ART, antiretroviral therapy; DETECT, Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury; HAND, HIV-associated neurocognitive disorder; IADL, instrumental activities of daily living.

[a]n = 58.

[b]n = 72; seven (9%) had missing viral loads (laboratory tests not completed), but all self-reported being undetectable for years.

status predicted DETECT total percentage scores ($b = -0.79$; 95% CI, −1.25 to −.34; $t[df = 76] = -3.51$, $P < .001$; AUC = 0.78, sensitivity = 0.65, positive predictive value = 0.44, negative predictive value = 0.85). This implies that individuals with HAND scored, on average, 7 percentage points lower than those without HAND (Figure 4). Similarly, individuals identified as impaired by way of the GDS cut score also performed worse on DETECT ($b = -0.83$, 95% CI, −1.27 to .39; $t[df = 76] = -3.75$, $P < .001$; AUC = 0.85, sensitivity = .63, positive predictive value = .46, negative predictive value = .81). This suggests that individuals with GDS impairment scored 9 percentage points lower on DETECT than individuals who were not GDS impaired.

Table 3. Correlations Between DETECT Performance and Traditional Neuropsychological Battery

	% Correct on DETECT	
Variable	Total	Response
Total	…	…
Response	0.72***	…
NP battery–global mean scaled score	0.45***	0.52***

All scores are corrected for age, sex, ethnicity, race, and education levels. Correlations reflect Pearson *R*.

Abbreviations: DETECT, Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury; NP, neuropsychological.

***$P < .001$.

Aim 3: Demographic Differences in DETECT Performance

Next, we tested for differences in unadjusted DETECT performance based on demographic characteristics (Table 4). Performance on DETECT did not vary by age, sex, or ethnicity; however, African American/Black participants had lower scores on DETECT ($b = -0.26$; 95% CI, −.49 to −.02; $t[df = 76] = -1.42$, $P = .03$) than non-Hispanic White participants. Additionally, participants with higher levels of education performed better on DETECT ($b = 0.01$; 95% CI, .00–.01; $V = 2082$, $P = .01$). For each additional year of education, there was a corresponding 1% increase in DETECT performance. The global scaled scores from the neuropsychological battery were related to IADL independence ($b = 1.19$; 95% CI, .40–1.90; $P < .001$), whereas DETECT performance was unrelated to IADL independence in this sample ($b = 0.32$; 95% CI, .23−.89; $P = .25$).

VR Usability

Of 79 participants, 6 (7.6%) reported feelings of motion sickness during the VR cognitive tests, with all 6 indicating that these feelings would not "discourage [them] from using a headset like this again." Nineteen participants (24.1%) reported feeling motion sickness at some point during the VR assessment, which included a nonpostural balance test, in which the background and foreground were moving in the final "seasick" trials. Of the 79 participants, 28 (35%) indicated zero difficulties using DETECT, while 87% of the sample reported ≤2 difficulties on the DETECT composite usability score. DETECT performance was not related to the composite measure of VR usability ($r = -0.07$, $P = .52$).

DISCUSSION

Given the high prevalence of HAND despite effective ART and the time/administrative limitations of traditional neuropsychological testing, along with older PWH being at risk for Alzheimer disease, new methods are needed to evaluate cognition in this growing population. Many abbreviated tests of cognition are available, but their overall performance for the

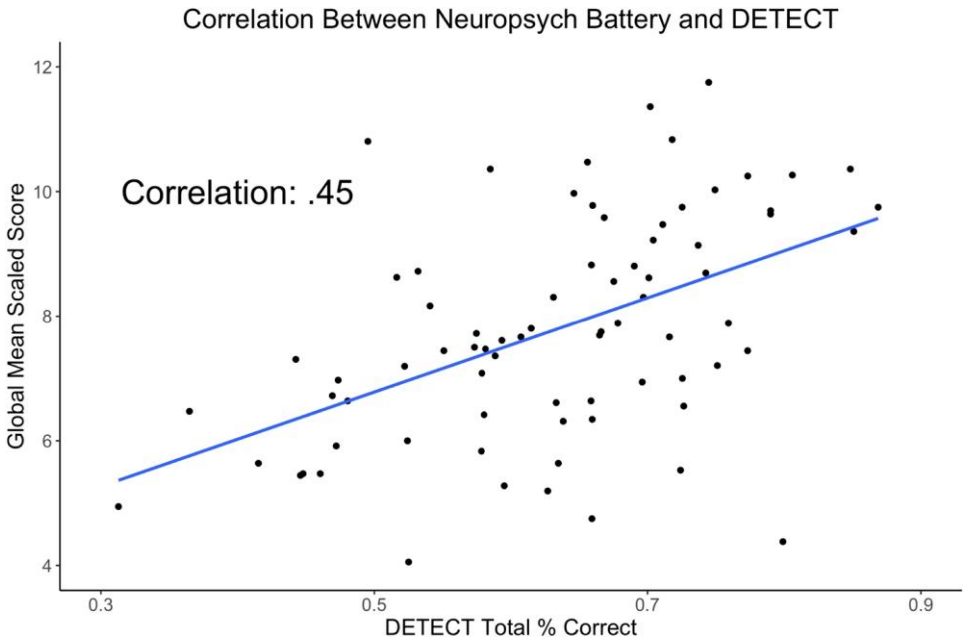

Figure 2. Correlation between comprehensive neuropsychological test battery and DETECT total percentage correct. DETECT, Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury.

diagnosis of HAND has been suboptimal. For example, the HDS appears to be relatively sensitive in detecting dementia among PWH who are hospitalized [30, 31]. Yet, most individuals with HAND in the ART era are outpatients with mild cognitive deficits. In this context, the HDS scores that were originally recommended are not sufficiently sensitive for the detection of ANI and mild neurocognitive disorder [32, 33]. The International HIV Dementia Scale was introduced for settings in which language barriers may affect the administration of the HDS. Much like the HDS, however, the International HIV Dementia Scale has been found to lack sensitivity and specificity for milder forms of HAND, and investigations have been limited to PWH aged <60 years [34].

Several other rapid tests designed for the general population have been evaluated in PWH, but again their diagnostic accuracy has been suboptimal. For example, while some studies have shown that the Montreal Cognitive Assessment has promise in the evaluation of PWH, others have shown that this rapid test has relatively low overall sensitivity and specificity for the diagnosis of HAND [33, 35, 36]. This includes a study demonstrating that the test provided no advantage over the HIV Dementia Scale [37]. Also, the evaluations of these other rapid tests have focused on PWH aged <60 years, and the tests themselves still require an administrator. More research is needed on older PWH, who carry a higher risk for HAND as well as neurodegenerative diseases such as Alzheimer disease [13, 14].

In this study, we evaluated the DETECT cognitive module in PWH aged ≥60 years. The study showed moderate convergent validity between the brief DETECT cognitive tests and the laboratory-based comprehensive neuropsychological battery. Furthermore, DETECT scores were able to differentiate participants with HAND (AUC = 0.78) as well as impairment based on GDS (AUC = 0.85). Screening older PWH for cognitive impairment should be a foundational component of geriatric assessments within HIV clinical care, yet the evaluation of cognition is often overlooked due to short clinical visits with competing priorities. Formal neuropsychological assessments are typically lengthy and costly; primary care physicians often refer patients only when there are overt signs of cognitive impairment. Abbreviated and validated cognitive screening tools are needed, and DETECT shows promise for real-world implementation as a low-cost self-administered solution. Due to the slightly stronger correlation between the response percentage correct score and traditional assessments than the total percentage correct score, we advise that the response percentage correct be used. However, this may have been due to test instructions that could have been more clear and to the relatively short time-out periods for some of the tests, which may not have been appropriate for our population. Future studies that replicate this finding are needed.

Contrary to our hypotheses, we did not find a relationship between DETECT performance and age, although this may have been due to the restricted age range of our participants (most participants were <70 years old despite a range of 60–80 years). While there appeared to be associations between DETECT performance and race and education, these

Figure 3. Correlation matrix for DETECT performance with neuropsychological test battery domains. DETECT, Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury; Pro Speed, processing speed.

Table 4. Differences in DETECT Performance Based on Demographic and Clinical Characteristics

			95% CI		
Effect	Estimate	SE	LL	UL	*P* value
Clinical characteristic					
HAND diagnosis	−0.79	0.23	−1.25	−.34	<.001
GDS: impaired, ≥0.5	−0.83	0.22	−1.27	−.39	<.001
GDS: continuous	−0.94	0.26	−1.45	−.43	<.001
IADL: independent	0.32	0.28	.23	.89	.25
Demographic characteristic					
Age[a]	0.00	0.01	.00	.00	.52
Sex: male[b]	0.03	0.03	−.03	.09	.40
Ethnicity: not Hispanic/ Latinx[b]	0.01	0.05	−.08	.10	.86
Race: Black[b]	−0.26	0.11	−.49	−.02	.03
Education[a]	0.00	0.02	.00	.01	.01

Abbreviations: DETECT, Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury; GDS, Global Deficit Score; HAND, HIV-associated neurocognitive disorder; IADL, instrumental activities of daily living; LL, lower limit; UL, upper limit.

[a]Kendall–Theil Sen Siegel nonparametric linear regression.

[b]Generalized additive model.

Table 5. AUC and Sensitivity/Specificity Analyses of Differences in DETECT Performance Based on Clinical Characteristics

Variable	AUC	Sensitivity	PPV	NPV
HAND	0.78	0.65	0.47	0.84
GDS	0.85	0.63	0.46	0.81
IADL	0.55	0.62	0.23	0.89

Abbreviations: AUC, area under the curve; DETECT, Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury; GDS, Global Deficit Score; HAND, HIV-associated neurocognitive disorder; IADL, instrumental activities of daily living; NPV, negative predictive value; PPV, positive predictive value.

associations have been well described with neuropsychological tests in general. There also was a lack of association between DETECT performance and IADL dependence, which the comprehensive neuropsychological battery was able to detect. Ongoing active efforts are underway by the developers of the DETECT technology to advance the platform, which may help improve the overall sensitivity of DETECT as a predictive

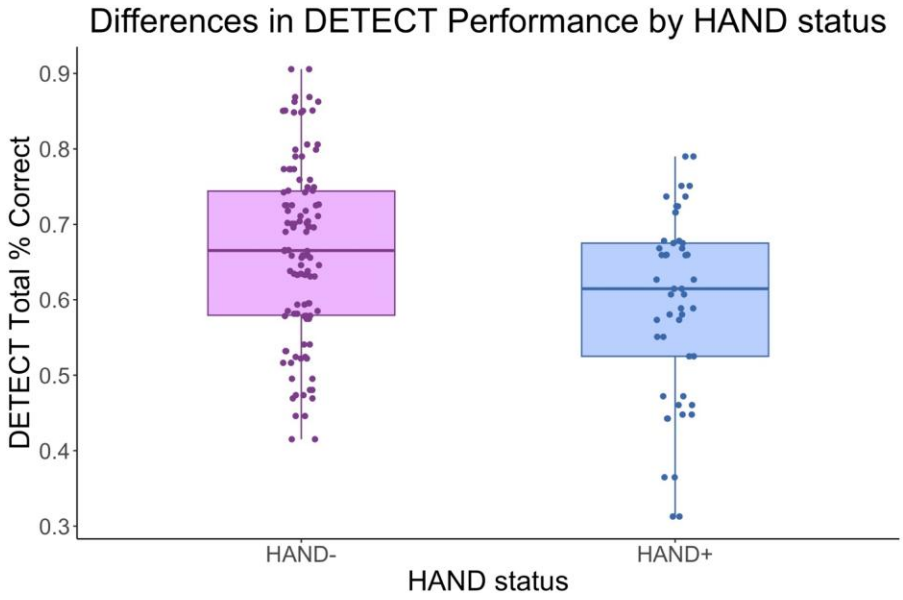

Figure 4. DETECT performance by HAND status. Data are presented as median (horizontal line), IQR (box), and range (vertical lines). DETECT, Display Enhanced Testing for Cognitive Impairment and Traumatic Brain Injury; HAND, HIV-associated neurocognitive disorder.

screening tool for cognitive impairment. Examples of these efforts include working on the next generation of VR technology, including the user interface, patient and provider facing, and the analysis algorithm that leverages components from the different tests to better inform the output. Additional efforts include collecting normative data across a wide demographic population and continuing to analyze and publish data from previous studies.

It is promising that the participants in this study reported minimal difficulties using the VR technology. Only 6 indicated motion sickness on the cognitive tests, while about 25% of the sample cited motion sickness at some point during the assessment, which included motion-related tests that are not included in the current study. The fact that usability was not related to DETECT performance suggests that the few participants who did experience some difficulties in using the VR technology were not the ones who performed worse on cognitive tests.

We acknowledge the limitations of the current study. Limitations regarding use of DETECT in this population included the following: some participants had difficulty knowing what buttons to press on the controller; the time limit to respond to some test items was quite short (eg, only 2 seconds to respond to N-Back Faces test stimuli); and practice trials were not included prior to the initiation of the actual test trials, which may have aided with ensuring understanding of instructions. The lack of practice trials may also have resulted in poorer test performance. A potential limitation of self-administered assessments such as DETECT, as compared with traditionally administered assessments, is that participants cannot ask the examiner questions to clarify test instructions. Regarding the study design, one limitation is that the current analysis is cross-sectional; therefore, we are not able to evaluate changes over time. Another limitation is the sample, which was a relatively clean sample of PWH and excluded participants who were not virologically suppressed, were non-English speaking, had probable dementia as the baseline visit, and had neurologic and mental health confounds. While the sampling was appropriate for the study design, it limits generalizability, and further study is needed to broaden these findings. Moreover, while the DETECT cognition module includes assessments of immediate and delayed recognition memory, working memory, and executive function, not all domains are represented. Specifically, there are no language fluency or motor tests, and the addition of a motor test in particular could strengthen the utility of DETECT as a clinical screening tool, as it is a HAND-relevant domain that is missing from the battery. Another limitation of the DETECT tests is an overrepresentation of Caucasian faces in the N-Back Faces test, which could limit cross-cultural validity. Despite these limitations, our findings suggest that the DETECT cognition module is a moderately sensitive screening tool for cognitive impairment associated with HIV and could be useful for identifying PWH who need formal follow-up cognitive testing. Furthermore, it is worthy for additional study to evaluate the performance of the test in different settings and to identify elements that might be improved or removed from the battery. Considerations for further research include evaluating neural-based biomarkers in relation to DETECT cognitive performance to better understand the biological mechanisms that

underpin cognition in PWH as well as to evaluate DETECT longitudinally.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Notes

Financial support. This work was supported by the National Institute of Mental Health (NIMH) and National Institute on Aging (NIA) (R01 AG062387 to R. C. M. and A. M. A., multiple principal investigators; R21 MH118092, R01 MH128872, and R01 MH128158 to A. M. A., principal investigator; P30 MH062512 to D. J. M., principal investigator); National Institute of Allergy and Infectious Diseases (NIAID) (T32AI007384 to K. S. K., coinvestigator); and Emory Center for AIDS Research (P30AI050409).

Potential conflicts of interest. R. C. M. is a cofounder of KeyWise AI and NeuroUX. K. S. K. received compensation as an employee and consultant from KeyWise AI. A. M. A. is currently performing a study at Emory University for which Eli Lilly is providing the study drug. All other authors report no potential conflicts. D. W. W., M. C. L. and N. K. hold 2 patents (U.S. 8,568,311 and 10,506,966) for the DETECT technology described in this study, the intellectual property for which is owned by Georgia Institute of Technology and Emory University. Should the technology be licensed to an outside entity they would be entitled to royalties derived from potential commercialization.

References

1. Palella FJ Jr, Delaney KM, Moorman AC, et al; HIV Outpatient Study Investigators. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. N Engl J Med **1998**; 338:853–60.
2. Samji H, Cescon A, Hogg RS, et al. Closing the gap: increases in life expectancy among treated HIV-positive individuals in the United States and Canada. PLoS One **2013**; 8:e81355.
3. Heaton RK, Clifford DB, Franklin DR Jr, et al. HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER study. Neurology **2010**; 75:2087–96.
4. Sacktor N, Skolasky RL, Seaberg E, et al. Prevalence of HIV-associated neurocognitive disorders in the Multicenter AIDS Cohort Study. Neurology **2016**; 86: 334–40.
5. Wang Y, Liu M, Lu Q, et al. Global prevalence and burden of HIV-associated neurocognitive disorder: a meta-analysis. Neurology **2020**; 95:e2610–21.
6. Ettenhofer ML, Foley J, Castellon SA, Hinkin CH. Reciprocal prediction of medication adherence and neurocognition in HIV/AIDS. Neurology **2010**; 74:1217–22.
7. Cohen MS, Chen YQ, McCauley M, et al. Antiretroviral therapy for the prevention of HIV-1 transmission. N Engl J Med **2016**; 375:830–9.
8. Strategies for Management of Antiretroviral Therapy Study Group; El-Sadr WM, Lundgren J, Neaton JD, et al. CD4+ count-guided interruption of antiretroviral treatment. N Engl J Med **2006**; 355:2283–96.
9. Tozzi V, Balestra P, Galgani S, et al. Neurocognitive performance and quality of life in patients with HIV infection. AIDS Res Hum Retroviruses **2003**; 19:643–52.
10. Saylor D, Nakigozi G, Nakasujja N, et al. HIV associated neurocognitive disorder leads to death. In: 2019 Conference on Retroviruses and Opportunistic Infections. Seattle, Washington; **2019** March 4–7. Abstract 425.
11. Tozzi V, Balestra P, Serraino D, et al. Neurocognitive impairment and survival in a cohort of HIV-infected patients treated with HAART. AIDS Res Hum Retroviruses **2005**; 21:706–13.
12. Petersen KJ, Metcalf N, Cooley S, et al. Accelerated brain aging and cerebral blood flow reduction in persons with HIV. Clin Infect Dis **2021**; 73:1813–21.
13. Sheppard DP, Iudicello JE, Morgan EE, et al. Accelerated and accentuated neurocognitive aging in HIV infection. J Neurovirol **2017**; 23:492–500.
14. Tan IL, Smith BR, Hammond E, et al. Older individuals with HIV infection have greater memory deficits than younger individuals. J Neurovirol **2013**; 19:531–6.
15. Aung HL, Aghvinian M, Gouse H, et al. Is there any evidence of premature, accentuated and accelerated aging effects on neurocognition in people living with HIV? A systematic review. AIDS Behav **2021**; 25:917–60.
16. Antinori A, Arendt G, Becker JT, et al. Updated research nosology for HIV-associated neurocognitive disorders. Neurology **2007**; 69:1789–99.
17. Anderson AM, Lennox JL, Nguyen ML, Waldrop-Valverde D, Tyor WR, Loring DW. Preliminary study of a novel cognitive assessment device for the evaluation of HIV-associated neurocognitive impairment. J Neurovirol **2016**; 22:816–22.
18. Wright DW, Nevarez H, Kilgo P, et al. A novel technology to screen for cognitive impairment in the elderly. Am J Alzheimers Dis Other Demen **2011**; 26:484–91.
19. Wright D, Goldstein F, Kilgo P, et al. Use of a novel technology for presenting screening measures to detect mild cognitive impairment in elderly patients. Int J Clin Pract **2010**; 64:1190–7.
20. Torre P III, Sundermann EE, Brandino A, et al. Auditory and cognitive function in older adults living with and without HIV. AIDS **2023**; 37:1971–8.
21. Heaton R, Miller S, Taylor M, Grant I. Revised comprehensive norms for an expanded Halstead-Reitan Battery: demographically adjusted neuropsychological norms for African American and Caucasian adults. Lutz: Psychological Assessment Resources, **2004**.
22. Kohli M, Fisher A, Sun-Suslow N, et al. Concurrent validity and reliability of at-home teleneuropsychological evaluations among people with and without HIV. J Int Neuropsychol Soc **2023**; 29:193–204.
23. Norman M, Moore D, Taylor M, et al. Demographically corrected norms for African Americans and Caucasians on the Hopkins Verbal Learning Test–Revised, Brief Visuospatial Memory Test–Revised, Stroop Color and Word Test, and Wisconsin Card Sorting Test 64-Card Version. J Clin Exp Neuropsychol **2011**; 33:793–804.
24. Heaton RK, Taylor M, Manly J, et al. Demographic effects and use of demographically corrected norms with the WAIS-III and WMS-III. In: Tulsky DS Saklofske D, Heaton RK, eds. Clinical interpretation of the WAIS-III and WMS-III. San Diego: Academic, **2002**:181–210.
25. Blackstone K, Moore DJ, Franklin DR, et al. Defining neurocognitive impairment in HIV: deficit scores versus clinical ratings. Clin Neuropsychol **2012**; 26:894–908.
26. Cysique LA, Franklin DR, Abramson I, et al. Normative data and validation of a regression based summary score for assessing meaningful neuropsychological change. J Clin Exp Neuropsychol **2011**; 33:505–22.
27. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist **1969**; 9(3 pt 1):179–86.
28. Becker JT, Kingsley L, Mullen J, et al. Vascular risk factors, HIV serostatus, and cognitive dysfunction in gay and bisexual men. Neurology **2009**; 73:1292–9.
29. Kallianpur AR, Wang Q, Jia P, et al. Anemia and red blood cell indices predict HIV-associated neurocognitive impairment in the highly active antiretroviral therapy era. J Infect Dis **2016**; 213:1065–73.
30. Berghuis JP, Uldall KK, Lalonde B. Validity of two scales in identifying HIV-associated dementia. J Acquir Immune Defic Syndr **1999**; 21:134–40.
31. Valcour V, Paul R, Chiao S, Wendelken LA, Miller B. Screening for cognitive impairment in human immunodeficiency virus. Clin Infect Dis **2011**; 53:836–42.
32. Bottiggi KA, Chang JJ, Schmitt FA, et al. The HIV Dementia Scale: predictive power in mild dementia and HAART. J Neurol Sci **2007**; 260(1–2):11–5.
33. Aita SL, Kaewpoowat Q, Yasri S, et al. Psychometric utility of the International HIV Dementia Scale and Montreal Cognitive Assessment in HIV-associated asymptomatic neurocognitive impairment. J Neurovirol **2021**; 27:568–78.
34. Haddow LJ, Floyd S, Copas A, Gilson RJ. A systematic review of the screening accuracy of the HIV Dementia Scale and International HIV Dementia Scale. PLoS One **2013**; 8:e61826.
35. Overton ET, Azad TD, Parker N, et al. The Alzheimer's Disease–8 and Montreal Cognitive Assessment as screening tools for neurocognitive impairment in HIV-infected persons. J Neurovirol **2013**; 19:109–16.
36. Rosca EC, Albarqouni L, Simu M. Montreal Cognitive Assessment (MoCA) for HIV-associated neurocognitive disorders. Neuropsychol Rev **2019**; 29:313–27.
37. Janssen MA, Bosch M, Koopmans PP, Kessels RP. Validity of the Montreal Cognitive Assessment and the HIV Dementia Scale in the assessment of cognitive impairment in HIV-1 infected patients. J Neurovirol **2015**; 21:383–90.