

Validity of Cognitive Screens for HIV-Associated Neurocognitive Disorder: A Systematic Review and an Informed Screen Selection Guide

Jody Kamminga · Lucette A. Cysique · Grace Lu · Jennifer Batchelor · Bruce J. Brew

Published online: 27 September 2013
© The Author(s) 2013. This article is published with open access at Springerlink.com

Abstract Various screening tools have been proposed to identify HIV-Associated Neurocognitive Disorder (HAND). However, there has been no systematic review of their strengths and weaknesses in detecting HAND when compared to gold standard neuropsychological testing. Thirty-five studies assessing HAND screens that were conducted in the era of combination antiretroviral therapy were retrieved using standard search procedures. Of those, 19 (54 %) compared their screen to standard neuropsychological testing. Studies were characterised by a wide variation in criterion validity primarily due to non-

standard definition of neurocognitive impairment, and to the demographic and clinical heterogeneity of samples. Assessment of construct validity was lacking, and longitudinal useability was not established. To address these limitations, the current review proposed a summary of the most sensitive and specific studies (>70 %), as well as providing explicit caution regarding their weaknesses, and recommendations for their use in HIV primary care settings.

Keywords HIV · AIDS · HAND · Cognitive screen · Computerised screen · HIV dementia scale · Criterion validity · Construct validity · Systematic review · Neuropsychology

Electronic supplementary material The online version of this article (doi:10.1007/s11904-013-0176-6) contains supplementary material, which is available to authorized users.

J. Kamminga (✉) · L. A. Cysique · G. Lu · B. J. Brew
St Vincent's Clinical School, St. Vincent's Hospital Applied Medical Research Centre, The University of New South Wales, Sydney, Australia
e-mail: j.kamminga@neura.edu.au

L. A. Cysique
e-mail: lcysique@unsw.edu.au

G. Lu
e-mail: g.lu@unsw.edu.au

B. J. Brew
e-mail: bbrew@stvincents.edu.au

J. Kamminga · L. A. Cysique
Neuroscience Research Australia, Barker Street, Randwick, Sydney, NSW 2031, Australia

J. Batchelor
Department of Psychology, Macquarie University, Sydney, Australia
e-mail: jennifer.batchelor@mq.edu.au

L. A. Cysique · B. J. Brew
Brain Sciences Institute, The University of New South Wales, Sydney, Australia

L. A. Cysique · G. Lu · B. J. Brew
Department of Neurology, St. Vincent's Hospital, Xavier Building Level 4 390 Victoria Street, Darlinghurst 2010, NSW, Australia

Introduction

The introduction of combination antiretroviral therapies (cART) has changed Human Immunodeficiency Virus (HIV)/Acquired Immunodeficiency Syndrome (AIDS) from a life threatening disease to one that is chronic [1]. Since the introduction of cART, HIV-associated neurocognitive disorder (HAND) has remained one of the main central nervous system complications in HIV infection, with its prevalence remaining stable at 30–50 % from the pre- to post- cART era [2–4]. This overall stability has in part been attributed to the increased lifespan associated with receiving cART [5] as well as the chronic effects of HIV on the brain [6]. Despite this overall stability, however, the more severe form of HAND, dementia, has been attenuated by cART, and mild forms of HIV-related neurocognitive impairment are now more common [7••].

This epidemiological shift prompted neuroHIV experts to reformulate the original HAND diagnostic criteria [8, 9] to better discriminate between demented and non-demented forms of the disease. As a result, the current American Academy of Neurology (AAN) HAND diagnostic classification criteria categorise three degrees of severity: mild (Asymptomatic Neurocognitive Disorder; ANI), moderate (Mild Neurocognitive Disorder;

MND), and severe (HIV-Associated Dementia; HAD) [10]. As the critical feature that distinguishes MND from ANI is the presence of significantly declined instrumental activities of daily living (IADL), accurate assessment of IADL status is imperative to diagnosis [10].

The majority of HIV positive (HIV+) individuals with HAND in the cART era do not present with HAD [11•]. It is therefore only through comprehensive neuropsychological assessment that ANI/MND are revealed [12•]. If not treated early, these non-demented forms of HAND have been found to be a risk factor for HAD [13]. Having mild HAND also yields a greater chance of developing future cognitive difficulties, particularly as HIV+ individuals age [14]. *Early* detection through screening is therefore imperative to minimise progression of ANI/MND to HAD, at which point there is likely less chance of complete recovery, even after cART is initiated. *Accurate* detection is also crucial to therapeutic and clinical care of those with HAND, particularly for ANI, as it enables adequate follow up in HIV+ individuals who would otherwise not be targeted for neurological care.

Although comprehensive neuropsychological assessment is the recommended gold standard for detecting HAND, cognitive screening measures such as the HIV Dementia Scale (HDS) [15] have been suggested when time and resources are limited [10]. As not all HIV+ individuals develop HAND, an initial brief screen followed by a comprehensive neuropsychological assessment in *only* those with an impaired screen is a cost-effective strategy that has previously been proposed [16•]. This strategy is only useful however, if the screen is able to detect at least mild HAND with adequate sensitivity and specificity.

Valcour and colleagues [17•] recently summarised existing HAND screens, however their review was not exhaustive and did not differentiate between those validated in pre-cART versus cART samples or against a gold standard neuropsychological assessment. More recently, a systematic review [18•] was conducted, although it was limited to the screening accuracy of the HDS and International HDS, and calculation excluded those with ANI, the most common form of HAND. To our knowledge, no study to date has systematically reviewed the criterion and construct validity of all existing cART era HAND cognitive screening studies under comprehensive methodological constraints. A review of this nature would establish a precedent in laying out the detailed strengths and limitations of the validation procedures that have been conducted, and hence the *practical* useability of existing screens in the cART era. Such a review would also serve as the basis for defining guidelines for improved validation strategies.

Based on this need, the aims of the current review were to:

1. Provide a systematic review of *all* HAND screening instruments in the cART era that have been assessed against standard neuropsychological testing as gold standard.
2. Critically review the findings, with particular attention to the validation procedure, base impairment rate, and criterion and construct validity.
3. Guide researchers and clinicians towards an informed screen selection, and to delineate recommendations for an improved screen validation.

Methods

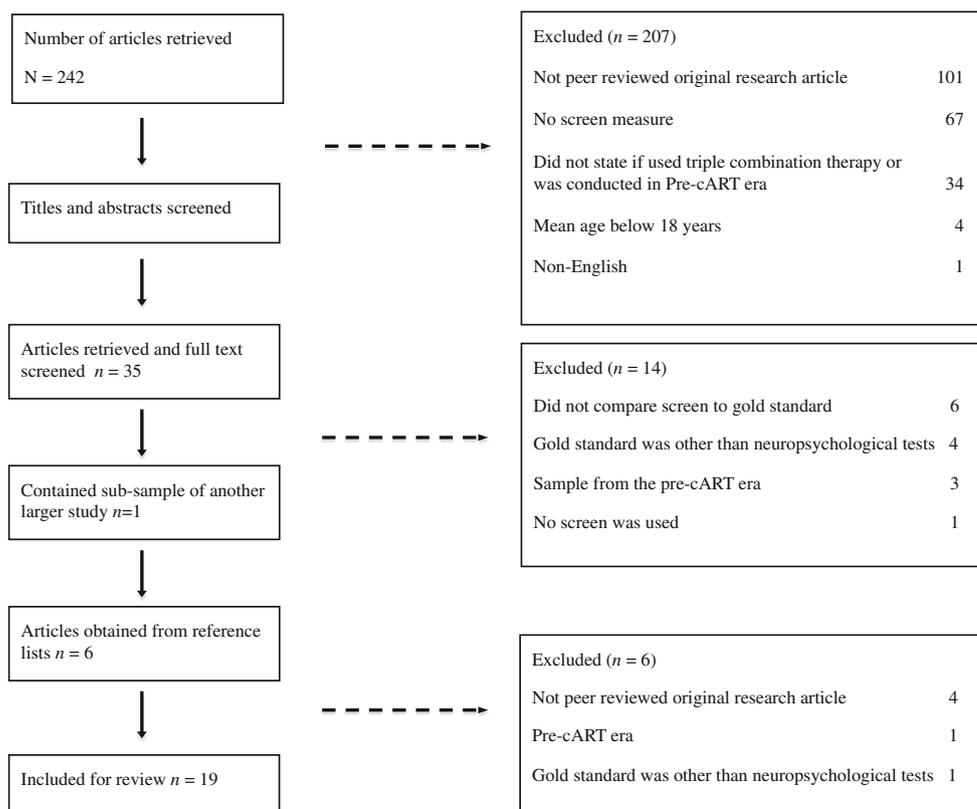
Seven databases were searched between June 2012 and May 2013, including Pubmed, psycINFO, Cochrane Library Databases, Scopus, and ScienceDirect. Each database was searched (keyword/title/abstract) using the following keywords: HIV and cognitive/neurocognitive/neuropsych*/computer* and screen*/test/battery and HAART/cART. The reference lists of included articles were also examined for eligible studies. After each search was completed, all returned abstracts were extracted, and duplicates removed.

As indicated in Fig. 1, the search returned 242 abstracts. Studies were excluded if:

- Written in a language other than English,
- Not a published peer-reviewed original research article (e.g. review, commentary or conference abstracts),
- It was not stated whether HIV+ individuals had received cART or were conducted in the pre-cART era (prior to 1996),
- A cognitive screening measure that assessed HIV-associated cognitive impairment was not included;
- The mean age of participants was below 18 years.

As outlined in Fig. 1, full text of 35 articles was retrieved. Of those, a further 14 were removed after consultation of the full text due to the following reasons: they did not compare the screen to a gold standard comprehensive neuropsychological assessment, the gold standard used was other than neuropsychological tests, the sample derived from the pre-cART era, and no screen was used. An additional study [19] was excluded as it used a sub-sample of another study [20]. Based on a manual search of the reference lists of retrieved articles, six abstracts of interest were retrieved and subsequently excluded (see Fig. 1). Each of the final 19 studies was systematically reviewed in a non-blind fashion, and data was extracted.

As the study methods and designs, and type of screening measure used varied substantially between the 19 studies, a systematic review was conducted rather than a meta-analysis. For each study in the review, the cognitive domains and specific items of the screening measures and NP gold standard were summarised. Sample characteristics were then presented for the HIV+ and HIV - groups using summary statistics including medians and ranges where appropriate. Next, impairment classification criteria and rate were summarised for

Fig. 1 Flow diagram of study selection

the neuropsychological gold standard and screening measures using medians and ranges. Finally criterion and construct validity were examined and medians and ranges were reported.

Results

The 19 empirical papers were published between 2003 and 2013. All but one were cross-sectional in design; one study [21] computed criterion validity indexes at baseline and evaluated stability at 12 weeks and 24 weeks follow-up, although longitudinal results were not reviewed here. See Online Resource 1 for a comprehensive overview of all 19 studies.

HAND Screening Measures

Of the 19 studies, 17 compared the utility of one screen to a gold standard (comprehensive neuropsychological testing), one evaluated three screens [22], and two studies compared two screens [23, 24]. Three studies used computerised screens. One study [25] used the original CogState battery that assesses psychomotor speed, attention and working memory, and visual learning and memory. Another study [21] used the Computerised Assessment of Mild Cognitive Impairment (CAMCI) [26], an eight task battery assessing attention,

psychomotor speed, working memory, learning and memory, and executive function (inhibition). The final computerised study [27] used the CalCAP Mini Battery [28], which targets one domain, psychomotor speed/reaction time, through two tasks. The most common paper and pen screen was the HDS ($n = 10$), followed by the IHDS ($n = 5$) [29]. Four studies used combinations of standardised neuropsychological tests; one Mexican study [24] used a large neuropsychological battery to assess the domains of orientation, attention, memory, language, visuospatial abilities, and executive functions through 16 tests, the remaining studies used smaller combinations (e.g. two, three or four) of neuropsychological tests [23, 30, 31].

Neuropsychological Gold Standard

All but one study [21] specified the individual tests used in the gold standard test battery. One study [29] included two samples that used a different battery of neuropsychological tests as the gold standard and were considered separately, generating a total of 20 reviewed studies. Comprehensiveness of the gold standard neuropsychological test battery varied widely across studies in terms of the domains assessed (see Table 1). Eight cognitive domains were assessed across the 20 studies. These included premorbid ability (not included in impairment rate), motor/fine motor skills, attention and working memory, psychomotor

Table 1 Neuropsychological tests used in each study by cognitive domain

| Study | Premorbid ability | Motor/ fine motor | Attention / working memory | Psychomotor speed/ reaction time | Learning and memory | Executive function | Language | Visuoconstruction |
|-------------------------|-------------------|-------------------|----------------------------|----------------------------------|---------------------|---------------------|----------|--------------------|
| [38] Chalermchai | | GP | | TMTA | BVMT | CT2 | SemFl | SemFl block design |
| | | FT | | CT1 | AVLT | First names fluency | | |
| [20] Sakamoto | | GP | LNS PASAT | TMTA | SMLT | WCST | SemFl | |
| | | | | Coding SS | FMLT | LetFl TMTB | | |
| [31] Moore | | GP | PASAT Digit Span | TMTA | HVLT | CWT | SemFl | |
| | | | | Coding SS | BVMT | WCST | | |
| | | | | Stroop Colour | | LetFl ActFl TMTB | | |
| [21] Becker | | | | X | ^a X | X | X | X |
| [34] Joska | | FT | | Coding | BVMT | CT2 | | |
| | | GP | | CT1 | HVLT | CWT | | |
| | | | | TMTA | | MAT | | |
| | | | | | | MCT | | |
| [24] Levine | | GP | LNS PASAT | Coding SS | BVMT | CWT | | |
| | | | | | HVLT | TMTB | | |
| | | | | | | LetFl | | |
| | | | | | | WCST | | |
| [33] Simioni | NART | | Digit Span SpWM | RT | | TMTB | | |
| | | | | TMTA | | | | |
| | | | | VIP | | | | |
| [22] Skinner | | GP | | ^b CalCAP | | TMTB | | |
| | | | | SDMT | | | | |
| | | | | TMTA | | | | |
| [40] Morgan | | GP | LNS PASAT | Coding SS | BVMT | TMTB | | |
| | | | | | HVLT | LetFl | | |
| | | | | TMTA | | WCST | | |
| [35] Singh | | | Digit Span | TMTA | | TMTB | | |
| [32] Bottiggi | NART | FO | Ruff 2 and 7 | SDMT | FMLT | CWT | | RCFT copy |
| | | GP | Seq RT | TMTA | RAVLT | TMTB | | |
| | | TG | | | | LetFl | | |
| [37] Wonja | | GP | | SDMT | RAVLT | CWT | | |
| | | | | TMTA | | | | |
| | | | | VCRT | | TMTB | | |
| [25] Cysique | | GP | Digit Span | SDMT | CVLT | Similarities | BNT | RCFT copy |
| | | | | TMTA | RCFT 3 min | TMTB | SemFl | |
| | | | | | | LetFl | | |
| [30] Ellis | Vocabulary | TG | PASAT | SS | HVLT | | | |
| | | GP | | | | | | |
| [45] Richardson | | GP | Digit Span | ^b CalCAP | RAVLT | CWT | | RCFT copy |
| | | | | TMTA | | | | |
| | | | | | | TMTB | | |
| [29] Sacktor (American) | | GP | | ^b CalCAP | RAVLT | Odd man | | RCFT copy |
| | | | | SDMT | | Out | | |

Table 1 (continued)

| Study | Premorbid ability | Motor/ fine motor | Attention / working memory | Psychomotor speed/ reaction time | Learning and memory | Executive function | Language | Visuoconstruction |
|------------------------|-------------------|-------------------|------------------------------|----------------------------------|---------------------|-------------------------------------|----------|-------------------|
| [29] Sacktor (Ugandan) | | GP | Digit Span | CT1 | VLT | LetFl CT2 | | |
| [23] Carey | | TG GP | LNS PASAT | SDMT Coding SS TMTA | BVMT HVL | CWT HCT TMTB LetFl WCST | SemFl | |
| [27] Gonzalez | | GP | Digit span PASAT VisSp | CT1 SDMT TMTA | RAVLT | CT2 CWT HCT TMTB LetFl | | Block design |
| [39] Smith | | GP | LNS | ^b CalCAP CT1 | CVLT Faces | CT2 CWT | | |

NART = National Adult Reading Test; GP = Grooved Pegboard; FT = Finger Tapping; FO = Finger Oscillation; TG = Timed Gait; LNS = Letter Number Sequencing; PASAT = Paced Auditory Serial Addition Test; SpWM = Spatial Working Memory, CANTAB; SeqRT = Sequential Reaction Time, CalCAP; VisSp = Visual Span; TMTA = Trial Making Test A; CT1 = Colour Trails Test 1; SS = Symbol Search; SDMT = Symbol Digit Modalities Test; RT = Reaction Time, CANTAB; VIP = Visual Information Processing, CANTAB; VCRT = Visual Choice Reaction Time; HVL = Hopkins Verbal Learning Test; BVMT = Brief Visuospatial Memory Test; SMLT = Story memory Learning Test; FMLT = Figure Memory Learning Test; RAVLT = Rey Auditory Verbal Learning Test; CVLT = California Verbal Learning Test; RCFT 3 min = Rey Complex Figure Test, 3 Minute Delay; VLT = World Health Organisation /University of California Verbal Learning Test; MAT = Mental Alternation Test; MCT = Mental Control Test; CT2 = Colour Trails 2; CWT = Stroop Colour Word Interference Test; WCST = Wisconsin Card Sort Test; LetFl = Letter Fluency; ActFl = Action Fluency; TMTB = Trail Making Test B; HCT = Halstead Category Test; SemFl = Semantic Fluency; BNT = Boston Naming Test; RCFT Copy = Rey Complex Figure Test Copy
“X” = Test not specified

^a verbal and visuospatial learning and memory

^b Individual tests not specified, assumed that the whole battery was administered

speed/reaction time, learning and memory, executive function, language, and visuoconstruction. No single study assessed all domains; instead, the median was five (range=3-7). The most frequently assessed domain was information processing speed/reaction time, with all studies using at least one measure, followed by executive function ($n=18$), learning and memory ($n=17$) and motor/fine motor skill ($n=17$). Ten studies used a measure of both verbal and visuospatial memory.

All studies that included a measure of motor/fine motor skill used the Grooved Pegboard. The Paced Auditory Serial Addition Test and Digit Span were the most frequently used measure of attention and working memory (both $n=6$), as was the Trail Making Test A ($n=13$) for information processing speed/reaction time. Trail Making Test B ($n=12$) was the most common measure of executive function. Within the verbal domain of learning and memory, the most frequently used task of verbal learning and memory was the Hopkins Verbal Learning Test ($n=6$), and the Brief Visuospatial Memory Test for visuospatial learning and memory ($n=6$). Semantic fluency was the preferred measure of language ($n=6$), and the Rey Complex Figure copy the most commonly employed test of

visuoconstruction ($n=3$). Two studies [32, 33] also included general screening measures in the neuropsychological gold standard test battery.

Sample Characteristics

Details of exclusion criteria, demographic characteristics for HIV+ and HIV-negative (HIV-) groups, and HIV infection characteristics can be consulted in Online Resource 1. In brief, studies varied considerably in their exclusion criteria, sample size, population, country location, demographic composition and HIV-specific features, such as years since diagnosis, CD4 cell count, viral load, and proportion receiving cART or with AIDS. The main sample source was tertiary referral. One study recruited a sample from a primary health care centre [34], one used a sample of US military beneficiaries [31], and four studies did not report their sample source. Thirteen studies were conducted in a Western country, four in non-Western countries, and one recruited both an American and Ugandan sample [29].

Impairment Classification Criteria and Rate

There was considerable heterogeneity in the methods used to calculate the overall impairment rate for the neuropsychological gold standard and screening measures. Of the 20 separate studies, seven explicitly reported the gold standard overall impairment rate, whilst 13 provided sufficient information for it to be calculated. In terms of impairment classification method, criteria for nine analyses were solely based on neuropsychological test scores, with one study reporting poorly defined impairment criteria (i.e. moderate cognitive impairment: “beyond the norms on at least two tests” without specific reference to the normative sample or cut-off criterion) [35]. Eleven analyses used clinical ratings that incorporated neuropsychological test scores, neurological information, brain imaging, and functional ability to classify individuals. Of those, ten used a version of the AAN criteria, and one used a system based on the Memorial Sloan Kettering (MSK) dementia scale [36], which contains gradations that range from minor cognitive disturbance to profound and incapacitating disorders, and integrates neurological deficits related to myelopathy. There was considerable disparity across studies regarding whether asymptomatic or subclinical cognitive impairment was considered in the overall impairment rate. While five studies [20, 24, 33, 37, 38] included this group as cognitively ‘impaired’, three studies [21, 22, 35] considered those individuals as ‘unimpaired’, and another [29] explicitly excluded this group from analysis. Conversely, a further study [39] excluded those with moderate to severe dementia from their sample. Considering the heterogeneity of the term ‘impairment’ and varying methods for classification, impairment rates based on the neuropsychological gold standard ranged widely across studies, from 19 % to 81 % (median=51 %).

In most instances studies generated multiple screen impairment ratings associated with the use of more than one screen or multiple methods for generating impairment (by varying cut-offs or including only a subset of the overall sample), generating 55 analyses see (Table 2). In general, screen-based impairment rate was poorly reported across studies. It was explicitly stated in one fifth, and six provided sufficient detail so that it could be calculated. Two thirds of the studies, however, failed to report the impairment rate or provide information to permit calculation. Across the 18 available analyses, median impairment rate was 33 % (range =4–71 %). In direct comparison to the neuropsychological gold standard, impairment for all screening measures was classified using test scores only; no IADL measure was used.

Criterion Validity

In most instances studies generated multiple criterion validity indexes based on varying cut-off scores, demographic adjustments, and on different subsamples of the overall sample,

generating a total of 55 analyses (presented in Table 2). Across all 55 screen analyses, sensitivity (the proportion of individuals identified as impaired on both the gold standard and the screen) ranged widely between 0 % and 93 % (median=66 %). Specificity (the proportion of individuals identified as unimpaired on both the gold standard and the screen) ranged between 25 % and 100 % (median=80 %). Overall classification accuracy was reported in 20 % ($n=12$) of the analyses, and ranged between 42 % and 87 % (median=65 %). Ten analyses (18 %) produced sensitivity and specificity over 70 %, and were ranked in order from highest to lowest sensitivity in Table 3. Twenty-five analyses (45 %) used standard raw cut-off criterion (≤ 10) for the HDS/IHDS, and 14 (25 %) analyses were based on subsets of larger samples.

In regards to varying levels of HAND severity, one study [20] that used the HDS found sensitivity decreased with the inclusion of milder HAND (77 % moderate-severe impairment, 65 % mild-moderate impairment, 63 % mild impairment). Another study [40] generated HDS criterion validity for HAD, MND, and ANI, separately and found sensitivity was 36 %, 23 %, 0 %, respectively, as compared to 17 % for the overall sample. Yet another [24] found 100 % of those diagnosed as ANI by the neuropsychological gold standard were misclassified when using the MMSE, and 50 % when using the HDS and NEUROPSI. Similarly, the IHDS was found not to distinguish between those who were subclinical (MSK=0.5) and those who were cognitively ‘unimpaired’ (MSK=0) across America and Ugandan samples [29]. By removing those who were diagnosed by the neuropsychological gold standard as ANI from the ‘impaired’ group, categorising them as ‘unimpaired’ and comparing this group to those with moderate to severe impairment (MND/HAD), sensitivity markedly improved for the IHDS (45 % to 81 %; [34]), but less so for the HDS (67 % to 70 %), and NEUROPSI (75 % to 80 %), and least of all for the Mini Mental Status Examination (MMSE; 8 % to 10 %). Lastly, for the HDS to be sensitive to milder forms of HAND, the cut-off was required to be raised to a score of 12 [33] or 14 or less [37].

By using demographic adjustments ($T < 40$), the HDS was shown to improve markedly in sensitivity as compared to the standard raw cut-off (≤ 10) for each level of HAND severity (HAD: 36 %-93 %; MND: 23 %-77 %; ANI: 0 %-50 %; [40]), and was associated with a marked yet acceptable decrease in specificity (from 94 % to 73 %). Indeed, seven out of the top ten screens with the highest sensitivity used demographically adjusted cut-off scores, and all screens that used short combinations of NP tests (20 %) used demographically adjusted cut-off criterion.

Three studies investigated concomitant criterion validity of several screens. One study [24] compared the utility of three screens using the same sample and found that none exhibited optimal values for *both* sensitivity and specificity, and the MMSE performed particularly poorly when

Table 2 Screen versus gold standard NP impairment rates and standard criterion validity indexes across studies (%)

| Study | Screen | NP IR | Screen IR | Cut-off | Sub-sample | Sensitivity | Specificity | Accuracy |
|------------------|----------------------------|-------------------|-------------------|--|--|-------------|-------------|----------|
| [40] Morgan | HDS | 43 ^a | - | $T < 40$ | HAD only | 93 | 73 | - |
| [31] Moore | 4 NP tests | 19 ^b | - | 4 tests $T < 40$, or 2 tests $T < 40 + 1$ test $T < 35$, or 2 tests $T < 35$, or 1 test $T < 40 +$ 1 test $T < 30$, or 1 test $T < 25$ | - | 87 | 87 | - |
| [31] Moore | 3 NP tests | 19 ^b | - | 3 tests $T < 40$ or 1 test $T < 40 + 1$ test $T < 35$, or 1 test $T < 30$ | - | 87 | 76 | - |
| [25] Cysique | CogState | 62 ^c | 62 ^c | - | - | 81 | 70 | - |
| [23] Carey | NP tests (HVLt-R& ndGP) | 29 ^c | 34 ^c | $T < 40$ on 1 test or $T < 35$ on 2 tests | - | 78 | 85 | 83 |
| [40] Morgan | HDS | 43 ^c | - | $T < 40$ | MND only | 77 | 73 | - |
| [23] Carey | NP tests (HVLt-R& Cod) | 29 ^c | - | $T < 40$ on 1 test or $T < 35$ on 2 tests | - | 75 | 92 | 87 |
| [31] Moore | 2 NP tests | 19 ^c | - | 2 tests $T < 40$ or 1 test $T \leq 35$ | - | 73 | 83 | - |
| [29] Sacktor | IHDS(American) | 38 ^c | - | ≤ 10.5 | - | 71 | 79 | - |
| [21] Becker | CAMCI | 31 ^c | 30 ^c | - | - ^g | 72 | 98 | - |
| [35] Singh | IHDS | 80 ^{c,f} | - | ≤ 10.5 | - | 94 | 25 | - |
| [32] Bottiggi | HDS | 52 ^c | - | ≤ 10 | - | 93 | 38 | - |
| [33] Simioni | HDS | 74 ^c | - | ≤ 14 | No self reported cognitive complaints | 88 | 67 | - |
| [35] Singh | IHDS | 80 ^{c,f} | - | ≤ 10 | - | 88 | 50 | - |
| [29] Sacktor | IHDS (Ugandan) | 31 ^c | - | ≤ 9.5 | - | 88 | 48 | - |
| [37] Wonja | HDS | 68 ^b | 48 ^{b,e} | ≤ 13 | - | 87 | 46 | - |
| [38] Chalermchai | IHDS | 51 ^c | - | ≤ 10 | - | 86 | - | - |
| [33] Simioni | HDS | 74 ^c | - | ≤ 14 | Self reported cognitive complaints | 83 | 63 | - |
| [34] Joska | IHDS | 81 ^b | - | ≤ 11 | - | 81 | 54 | - |
| [29] Sacktor | IHDS (American) | 38 ^b | - | ≤ 10 | - | 80 | 57 | - |
| [22] Skinner | IHDS | 40 ^b | - | ≤ 10 | - | 77 | 65 | 70 |
| [20] Sakamoto | HDS | 51 ^c | 12 ^b | ≤ 10 | moderate-severe impairment | 77 | - | - |
| [24] Levine | NP tests | 52 ^b | 71 ^c | T score (varied) | - | 75 | 61 | - |
| [20] Sakamoto | HDS | 51 ^c | 56 ^c | $T < 40$ | - | 69 | 56 | 63 |
| [27] Gonzalez | CalCAP | 57 ^c | 49 ^c | Average deficit | - | 68 | 77 | 72 |
| [24] Levine | HDS | 52 ^b | 62 ^c | ≤ 10 | - | 67 | 50 | - |
| [20] Sakamoto | HDS | 51 ^c | - | ≤ 14 | - | 66 | 61 | - |
| [20] Sakamoto | HDS | 51 ^c | - | ≤ 10 | virologically suppressed only | 66 | 55 | 61 |
| [20] Sakamoto | HDS | 51 ^c | 19 ^b | ≤ 10 | mild-moderate impairment | 65 | - | - |
| [37] Wonja | HDS | 68 ^b | 48 ^{b,e} | ≤ 12 | - | 63 | 84 | - |
| [20] Sakamoto | HDS | 51 ^c | 20 ^b | ≤ 10 | mild impairment | 63 | - | - |
| [22] Skinner | HDS | 40 ^b | - | ≤ 11 | - | 62 | 80 | - |
| [30] Ellis | 3 NP tests | 56 ^{c,d} | - | Averaged z-scores ≤ 0.5 | - | 58 | 82 | 68 |
| [32] Bottiggi | HDS | 52 ^c | - | ≤ 10 | Moderate & severe impairment | 57 | 84 | - |

Table 2 (continued)

| Study | Screen | NP IR | Screen IR | Cut-off | Sub-sample | Sensitivity | Specificity | Accuracy |
|------------------|------------|-------------------|-----------------|------------------------------------|------------------------|-------------|-------------|----------|
| [45] Richaradson | HDS | 50 ^c | 40 ^c | ≤10 | - | 55 | 75 | 63 |
| [33] Simioni | HDS | 74 ^c | - | ≤10 | - | 54 | 96 | - |
| [38] Chalermchai | IHDS | 51 ^c | - | ≤10 | - | 53 | 89 | - |
| [40] Morgan | HDS | 43 ^b | - | T<40 | ANI only | 50 | 73 | - |
| [22] Skinner | HDS | 40 ^b | - | ≤10 | - | 46 | 80 | 67 |
| [22] Skinner | MMSE | 40 ^b | - | ≤27 | - | 46 | 55 | - |
| [34] Joska | IHDS | 81 ^b | - | ≤10 | - | 45 | 80 | 56 |
| [30] Ellis | 3 NP tests | 56 ^{c,d} | 32 ^c | ≤1 SD 1 test | - | 44 | 84 | - |
| [39] Smith | HDS | 49 ^b | - | ≤10 | - | 39 | 85 | - |
| [40] Morgan | HDS | 43 ^b | - | ≤10 | HAD only | 36 | 94 | - |
| [32] Bottiggi | HDS | 52 ^c | - | ≤10 | Severe impairment only | 36 | 94 | - |
| [38] Chalermchai | IHDS | 51 ^c | - | ≤10 | - | 34 | 87 | - |
| [30] Ellis | 3 NP tests | 56 ^{c,d} | 15 ^c | ≤1 SD on 1 test / ≤2 SD on 2 tests | - | 24 | 98 | - |
| [20] Sakamoto | HDS | 51 ^c | 17 ^c | ≤10 | - | 24 | 92 | 57 |
| [40] Morgan | HDS | 43 ^b | - | ≤10 | MND only | 23 | 94 | - |
| [40] Morgan | HDS | 43 ^b | - | ≤10 | - | 17 | 94 | - |
| [23] Carey | HDS | 29 ^b | 4 ^b | ≤11 | - | 9 | 98 | - |
| [24] Levine | MMSE | 52 ^b | 14 ^c | ≤25 | - | 8 | 88 | - |
| [30] Ellis | 3 NP tests | 56 ^{c,d} | - | Clinical rating ≥5 | - | 2 | 100 | 42 |
| [40] Morgan | HDS | 43 ^b | - | ≤10 | ANI only | 0 | 94 | - |
| N=55 | Median | 51 | 33 | | | 66 | 80 | 65 |
| | Range | 19-81 | 4-71 | | | 0-93 | 25-100 | 42-87 |

See reference list for numbered studies. Rounded to nearest whole value. Ordered with balance of highest sensitivity and specificity. Studies with ≥70 % sensitivity and specificity are bold. Studies ≤70 % sensitivity and specificity are in grey. Values are based on standard cutoff scores unless otherwise stated. Values represent percentages unless otherwise stated, and are to the nearest decimal reported. NP = neuropsychology. IR = Impairment rate. HDS = HIV Dementia Scale. NP = neuropsychological. HVLT-R = Hopkins Verbal Learning Test, Revised. ndGP = non-dominant hand Grooved Pegboard. Cod = Coding. IHDS = International HIV Dementia Scale. MMSE = Mini Mental Status Exam. Accuracy = classification accuracy. ‘-’ = not reported. Sensitivity refers to the proportion of individuals identified as impaired on both the gold standard and the screen. Specificity refers to the proportion of individuals identified as unimpaired on both the gold standard and the screen. PPP relates to specificity and is defined as the probability that an individual has HAND given their positive screen result. NPP relates to sensitivity and is the probability that an individual does not have HAND given their negative screen result. ^a proportion; ^b manually calculated; ^c values reported; ^d based on test scores only; ^e based on a cutoff score of 12 or less; ^f “any neurocognitive impairment”; ^g sample excluded ANI category

compared to the HDS and NEUROPSI. Similarly, another study [22] compared the utility of the HDS and IHDS within the same sample and found that while the HDS showed the highest sensitivity (76.9 % versus 46.2 %), the IHDS had superior specificity (65 % versus 55 %). One study [23] showed that within their sample, the pairing of the Hopkins Verbal Learning Test Recall and Grooved Pegboard non-dominant scores drastically outperformed the HDS in terms of sensitivity (78 % versus 9 %), though demonstrated inferior specificity (85 % versus 98 %). Of note, an additional study found the IHDS to increase by 30 % (53.3 % vs 86 %) with the inclusion of Trail Making Test part A [38].

Lastly, one study [20] that used the HDS with standard cut-off criterion found that by including only those in the sample

who were viruologically suppressed, sensitivity and overall classification accuracy were raised (24 % to 66 %, and 57 % to 61 %, respectively), although this was associated with a decreased in specificity (92 % to 50 %).

Construct Validity

Only three studies reported correlations between their respective screening measure and neuropsychological tests and/or cognitive domains (see Table 4). Very small to moderate associations were observed generally across these studies and no screen was clearly primarily related to neuropsychological tests associated with core HAND domains (e.g. information processing speed or attention and working memory). In many instances although significant correlations were

Table 3 Screen versus gold standard NP impairment rates and standard criterion validity indexes for studies where sensitivity and specificity is 70 % or higher (%)

| Study | Screen | NP IR | Screen IR | Cut-off | Sample | Sensitivity | Specificity | Accuracy |
|----------------------------|-------------------------|-----------------|-----------------|--|----------------------------|-------------|-------------|----------|
| [40] Morgan | HDS | 43 ^a | - | $T < 40$ | HAD only | 93 | 73 | - |
| [31] Moore | 4 NP tests | 19 ^b | - | 4 tests $T < 40$, or 2 tests $T < 40 + 1$ test $T < 35$, or 2 tests $T < 35$, or 1 test $T < 40 + 1$ test $T < 30$, or 1 test $T < 25$ | Entire sample | 87 | 87 | - |
| [31] Moore | 3 NP tests | 19 ^b | - | 3 tests $T < 40$ or 1 test $T < 40 + 1$ test $T < 35$, or 1 test $T < 30$ | Entire sample | 87 | 76 | - |
| [25] Cysique | CogState | 62 ^a | 62 ^a | - | Entire sample | 81 | 70 | - |
| [23] Carey | NP tests (HVL-R & ndGP) | 29 ^a | 34 ^a | $T < 40$ on 1 test or $T < 35$ on 2 tests | Entire sample | 78 | 85 | 83 |
| [40] Morgan | HDS | 43 ^a | - | $T < 40$ | MND only | 77 | 73 | - |
| [23] Carey | NP tests (HVL-R & Cod) | 29 ^a | - | $T < 40$ on 1 test or $T < 35$ on 2 tests | Entire sample | 75 | 92 | 87 |
| [31] Moore | 2 NP tests | 19 ^b | - | 2 tests $T < 40$ or 1 test $T \leq 35$ | Entire sample | 73 | 83 | - |
| [29] Sacktor (American) | IHDS | 38 ^a | - | ≤ 10.5 | Entire sample | 71 | 79 | - |
| [21] Becker | CAMCI | 31 ^a | 30 ^b | - | Entire sample ^c | 72 | 98 | - |

Consideration of the sample in question is necessary for proper screen selection. See reference list for numbered studies. Rounded to nearest whole value. Ordered with balance of highest sensitivity and specificity. NP = neuropsychological. IR = Impairment rate. Accuracy = overall correct classification accuracy. '-' = not reported. Sensitivity refers to the proportion of individuals identified as impaired on both the gold standard and the screen. Specificity refers to the proportion of individuals identified as unimpaired on both the gold standard and the screen. HAD = HIV Associated Dementia. MND = Mild Neurocognitive Disorder. HVL-R = Hopkins Verbal Learning Test, revised. ndGP = nondominant hand Grooved Pegboard. Cod = Coding. ^a manually calculated; ^b values reported; ^c sample excluded ANI category

observed with core HAND domains, they were often also related to other domains (e.g. visuoconstruction) to a similar magnitude.

Discussion

The search for an optimal screening tool in HIV infection is ongoing and remains a major challenge [41]. In keeping with current recommendations regarding reliable HAND diagnosis [10], we propose that a key factor for optimal HAND screen validation is the comparison to a gold standard comprehensive neuropsychological battery. Of the initial 35 studies identified that screened for HAND, 19 (54 %) were compared to a neuropsychological gold standard.

Among those, there was substantial variation in six key factors that hindered optimal screen validation and straightforward interpretation of the literature. These were:

1. Inclusion/non-inclusion of some HIV+ individuals without an optimal rationale
2. Lack of assessment of screening tool construct validity
3. Variation in overall gold standard impairment rate and use of non-standard definitions of impairment
4. Lack of reporting screen overall impairment rate and variation in screen impairment definition with no optimal statistical rationale
5. Inclusion/non-inclusion of a control (HIV-) group; and
6. Lack of reporting of important HIV and/or demographic characteristics.

To aid in selecting an optimal screen, all analyses were ranked by optimal sensitivity and specificity, and those with both indexes greater than 70 % were presented separately (see Table 3). A cut-off of 70 % was selected to yield a demanding threshold for type I and type II errors simultaneously. Depending on the clinical context, however, lower ranked screens such as those with high sensitivity and low specificity may be more relevant and were also included in Table 2 as a guide for selecting an appropriate screen. Caution is advised,

Table 4 Reported correlations between screening measure and neuropsychological test and cognitive domain for the four studies

| Study (screening measure) | Highest | <i>r</i> | Lowest | <i>r</i> |
|---------------------------|---|------------------|-------------------------------|-------------------|
| | Domain (NP test) | | Domain (NP test) | |
| [34] Joska (IHDS) | Information processing speed (CT1) / memory (BVMT recall) | .35 ^a | Executive function (CT2) | .21 ^a |
| [25] Cysique (CogState) | Fine motor skill (GPd) | .62 ^b | Visuoconstruction (RCFT copy) | -.22 ^c |
| [39] Smith (CalCAP mini) | Executive function | .43 ^d | Fine motor skill | .22 ^e |

The follow studies did not report any characteristics for HIV+ group: Chalermchai et al., 2013, Sakamoto et al., 2013, Moore et al., 2012, Becker et al. 2011, Singh et al., 2008, Bottiggi et al., 2007, Carey et al., 2004, Ellis et al., 2005, Levine et al., 2011, Morgan et al., 2005, Richardson et al., 2005, Sacktor et al., 2005 (American or Ugandan), Simioni et al., 2010, Skinner et al., 2009, Smith et al., 2003, Wonja et al., 2007. Only correlations that reached statistical significance ($p < .05$) were reported. NP = Neuropsychological. r = Pearson's correlation. IHDS = International HIV Dementia Scale. CT1 = Colour Trails Test part 1. BVMT = Brief Visuospatial Memory Test. CT2 = Colour Trails Test part 2. GPn = Grooved Pegboard non-dominant hand. GPd = Grooved Pegboard dominant hand. RCFT = Rey Complex Figure Test

^a Total score

^b Identification (RT)

^c Matching (RT)

^d Summary deficit score of Halstead Category Test, Trail Making Test part B, and Stroop Color Word Interference Test

^e Choice RT 4 (RT)

however, when using screens that have sensitivity close to chance level (50 % or less). For this reason, the MMSE should *not* be used to assess HAND. Of note, only one study assessed the longitudinal validity of a screening tool. Consequently, the following recommendations apply to cross-sectional studies that examined the validity of screens, and it will be necessary to better establish the longitudinal validity of any of the current screen included in this review.

How to Best Interpret Criterion Validity

It is recommended that attention be paid to the characteristics of the sample that the screen was validated in. In particular, it is imperative that the level of clinical comorbidities (which may or may not impact cognition), referral source (tertiary or not), and the degree of HIV disease severity are comparable to the context that the screen is to be used in, as greater levels of comorbidities and advanced HIV (more common in tertiary settings) are likely to lead to a higher base rate of impairment and inflated criterion validity when compared to the wider HIV+community.

It is also important to select a screen that has been validated against a standard definition of impairment, and preferably the current AAN HAND nomenclature [10], as these criteria have a robust statistical grounding upon which to define impairment and increases comparability across studies [11••]. The sole use of neuropsychological test scores without assessment of functional status precludes the use of the AAN criteria and the discrimination of ANI from MND [10]. Caution is also warranted if the validation study excluded ANI, as this is the

most common form of HIV-related cognitive impairment in the cART era [7••]. Although the validity of ANI as a sub-clinical form of HAND was initially equivocal, cumulative evidence now suggests it to be a valid neuropathological entity [42] that is predictive of further cognitive decline [11••]. Finally, as the ability to accurately calculate sensitivity and specificity for a test directly depends on the correct use of base rates [43], artificial manipulation of the base-rate impairment invalidates the standard definitions of HAND and results in an unreliable estimation of impairment in the sample.

Strengths and Weaknesses of Screens with Highest Criterion Validity

The screen that exhibited the highest criterion validity was the demographically adjusted HDS [40]. However, this occurred only when the subset of HAD was considered, as the inclusion of MND and ANI resulted in inadequate (chance-level) sensitivity and is therefore inappropriate for use in the cART era where ANI is the most prevalent form of HAND [7••]. Overall, the screens that examined short neuropsychological test combinations [23, 31, 40] consistently demonstrated the highest performing criterion validity. While such a finding would be expected, this method may not be feasible in a busy clinical practice where time is limited and where there may not be anyone qualified to administer, and most importantly, to correctly score and interpret such tests. As an additional barrier, for legal and ethical reasons some tests can only be interpreted by a qualified neuropsychologist.

The CogState computerised battery [25] was as good as paired neuropsychological tests. However, this result may have been inflated by the advanced HIV disease stage of all participants and associated high base rate of impairment. Moreover, the specific battery lacked a verbal list-learning task of memory which has been shown to be particularly sensitive to the deficits characterising mild HAND [44]. That withstanding, a key benefit of using a computerised task is that compared to neuropsychological tests, it requires minimal skill regarding administration, interpretation, and scoring. A caveat to this however, is that the final interpretation will likely need to be monitored by a neuropsychologist as the battery is sensitive to not only HAND, but cognitive impairment in general. Lastly, although the CAMCI demonstrated one of the highest criterion validity indexes [21], the sample excluded those with ANI, which both artificially inflated the sensitivity and rendered the screen non-usable to detect ANI. In addition, replication of the study is impossible as the individual neuropsychological tests employed were not reported.

Finally, *all* HAND screen validation studies failed to include a measure of IADL status in their screening measure. This remains a key limitation as it precludes optimal operationalisation of the current HAND criteria when using the screen and its ability to differentiate between ANI and MND [10].

Construct Validity

Construct validity was reported in a minority of studies (3/19 or 16 %). Neglecting to report such data results in an inability to determine which tests are more/less valuable as screening measures and limits progression in the field. There are two possible explanations for the failure to report construct validity. First, most studies used samples of convenience, rendering it probable that the test battery was not selected to compare with a screen. Second, the majority of screens were initially developed in the pre cART era. As such, the construct validity may have been assumed which is problematic as in the pre-cART era HAND was more prevalent. Construct validity is therefore a vital index of validity as HAND has become milder.

In those studies that reported construct validity, it was found that nearly all associations were within the small to moderate range, and none consistently and uniquely associated with core HAND domains. One explanation of this finding is that these screens measure different facets of cognition compared to traditional neuropsychological tests. Such an assertion is consistent with previous findings [28] that used factor analysis to demonstrate that the subtests from the CalCAP screen loaded on separate factors to neuropsychological tests, and such an interpretation is reinforced by the finding that the screens that used neuropsychological test

pairings were among the most sensitive and specific. This argues for renewed effort and resources to be directed toward tackling this clinical research question. One possible means to improve criterion validity of computerised screens is to include reaction time measures in the gold standard neuropsychological battery (e.g. [45]).

Key Clinical and Demographic Information

Several studies did not adequately report all key clinical data and demographic variables for their sample. This hinders a straightforward interpretation of the clinical relevance of the findings as these factors have likely impacted the level of gold standard and screen impairment rate and the related level of criterion validity. Caution is therefore recommended when using the optimal screens highlighted in this review. Prior to the use of any screen, it should be ensured that the population of interest is adequately represented by the validation sample.

Recommendations for Optimal Validation of a HAND Screening Procedure

From the analysis of existing studies several key aspects should be considered when developing and validating a HAND screening procedure. Prospective studies are required which adopt the following recommendations:

1. Selection of adequate screen outcome measures that target core HAND domains that are known to be predominantly affected in the cART era.
2. Assess IADL with a standard instrument to allow screen operationalisation of the AAN criteria. It should also be noted that IADL is best interpreted in the context of concomitant assessment of mood status [46].
3. Comparison of the screen to a *comprehensive* neuropsychological gold standard (i.e. assessment of core HAND domains in addition to fine motor skill, language, visuospatial skills, and premorbid ability) in combination with IADL status, to improve the reliability of base-rate of impairment.
4. Use of the most representative sample of the HIV population within one country or country region (not exclusively tertiary healthcare samples, for example).
5. The inclusion of a control (HIV-) group with similar characteristics to optimally assess HAND specificity.
6. Explicit rationale for screen impairment criteria.
7. Reporting of all standard criterion validity indexes.
8. Reporting of construct validity.
9. Assessment of the screen feasibility in practice.
10. Assess the longitudinal validity of the screening tool including correction for practice effects.

Recommendations for Optimal use of HAND Screening, Patient Results and Feedback in HIV Primary Care

The overarching suggestion is that rather than seeking to implement a screening tool in isolation, a *procedure* should instead be applied to yield an optimal *indication* of the need of each patient for further neurological examination. This decision is based on whether there is:

1. HIV-related neurocognitive impairment on the cognitive screen tool, and if so, to what degree;
2. Whether any other neuropsychiatric confounds (based on standardised questionnaires or clinical interviews regarding mood/substance use and medical history) may contribute to the presence or severity of neurocognitive impairment, and
3. Whether the detected neurocognitive impairment interferes significantly with everyday functioning (i.e., IADL status).

If any of the above three points are present in combination or in isolation, the primary care physician should refer the patient to a neurologist with an accompanying referral letter detailing the aforementioned information. Importantly, as the AAN HAND criteria [10] outlines, the presence of neurocognitive impairment alone is not sufficient to reach a diagnosis of HAND. This is particularly the case when cognitive function has been assessed via a brief screen. It is therefore only following neurological examination in combination with further clinical investigation including brain imaging, extensive blood test panel and if possible extensive cerebrospinal fluid panel, and careful recording of neurological history, that a HAND diagnosis can be reached [10].

The success of any HAND screening procedure at the primary care level also requires the practice to have necessary resources available to optimally train the non-neuropsychologist (e.g. nurses) in administration of the screening assessment, as well as for conducting long-term quality control on test administration and neurocognitive and clinical data handling. It would also be optimal for a key primary care physician within the clinic with additional training in HIV neurology to oversee the implementation of this process. To adequately deal with complex cases (e.g. with multiple comorbid factors such as mood and substance disorders), and to manage long-term quality assurance, it is also advised that the lead primary care physician seeks collaboration with a senior neuropsychologist.

One final but vital aspect to consider is the dissemination of results to the patient. As emphasised above, *results from a screening procedure serve only to indicate the need for further neurological examination*. It is therefore strongly advised that feedback from the HAND screen is disseminated in specifically these terms *only*. As the screening procedure is not diagnostic, mention of the presence or absence of “neurocognitive impairment” is *not* appropriate. Additionally, if co-morbid life-threatening conditions are present (e.g. major depressive

disorder with suicidal ideations or severe cardiac health issues) treatment of these issues via psychiatric or specialist intervention should take precedence over that of HAND. Once these issues have been addressed and are stable, the patient should then be re-assessed on the screen and directed for neurological care as necessary.

Acknowledgements This work was supported by the Peter Duncan Applied Neuroscience Unit at the St. Vincent’s Hospital Applied Medical Research Centre, Sydney Australia (Head of Group. Prof. Bruce Brew) and the NHMRC CDF APP1045400 (CIA-Cysique).

Compliance with Ethics Guidelines

Conflict of Interest Lucette A. Cysique received grants from MSD and NHMRC; consulted for CogState and Canadian Trial Network in 2012; and received honoraria from Abbvie and ViiVhealthcare in 2011.

Grace Lu awarded Dementia Collaborative Research Centre Honours Scholarship of \$5000.

Bruce J. Brew received payment as board member for GlaxoSmithKline, Biogen Idec, ViiV Healthcare, Merck Serono, and *International Journal for Tryptophan Research*; has received grants from National Health and Medical Research Council of Australia and from MSRA; received honoraria from ViiV Healthcare, Boehringer Ingelheim, Abbott, Abbvie, and Biogen Idec; receives royalties from *HIV Neurology* (Oxford University Press, 2001) and *Palliative Neurology* (Cambridge University Press, 2006); received travel/accommodation expenses covered or reimbursed by Abbott; and has research support funding from BI Lilly, GlaxoSmithKline, ViiV Healthcare, Merck Serono, and St. Vincent’s Clinic Research Foundation.

Jody Kamminga and Jennifer Batchelor declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Papers of particular interest, published recently, have been highlighted as:

- Of importance
- Of major importance

1. UNAIDS: UNAIDS report on the global AIDS epidemic, 2012. Available at http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2012/gr2012/20121120_UNAIDS_Global_Report_2012_en.pdf.
2. Cysique LA, Maruff P, Brew BJ. Prevalence and pattern of neuropsychological impairment in human immunodeficiency virus-infected/acquired immunodeficiency syndrome (HIV/AIDS) patients across pre- and post-highly active antiretroviral therapy eras: a combined study of two cohorts. *J Neurovirol.* 2004;10:350–7.
3. Heaton RK, Grant I, Butters N, et al. The HNRCC 500—neuropsychology of HIV infection at different disease stages. *HIV Neurobehavioral Research Center. J Int Neuropsychol Soc.* 1995;1:231–51.

4. Robertson KR, Smurzynski M, Parsons TD, et al. The prevalence and incidence of neurocognitive impairment in the HAART era. *AIDS*. 2007;21:1915–21.
5. Tozzi V, Balestra P, Serraino D, et al. Neurocognitive impairment and survival in a cohort of HIV-infected patients treated with HAART. *AIDS Res and Hum Retroviruses*. 2005;21:706–13.
6. • Mothobi NZ, Brew BJ. Neurocognitive dysfunction in the highly active antiretroviral therapy era. *Curr Opin Infect Dis*. 2012;25(Mothobi NZ, Brew BJ):4–9. *Summarises the most likely cause of HIV brain impairment in cART era*.
7. •• Heaton RK, Clifford DB, Franklin Jr DR, et al. HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study. *Neurology*. 2010;75:2087–96. *Characterisation of HAND in the cART era, with mild HAND being the prominent form of HAND today*.
8. The Dana Consortium on Therapy for HIV Dementia and Related Cognitive Disorders. Clinical confirmation of the American Academy of Neurology algorithm for HIV-1-associated cognitive/motor disorder. *Neurology*. 1996, 47:1247-53.
9. [No authors listed]. Nomenclature and research case definitions for neurologic manifestations of human immunodeficiency virus-type 1 (HIV-1) infection. *Neurology*. 1991, 41:778-85.
10. Antinori A, Arendt G, Becker JT, et al. Updated research nosology for HIV-associated neurocognitive disorders. *Neurology*. 2007;69:1789–99.
11. •• Cysique LA, Bain MP, Lane TA, Brew BJ. Management issues in HIV-associated neurocognitive disorders. *Neurobehav HIV Med*. 2012;4:63–73. *Current review on HIV/AIDS management from a clinical neuropsychological perspective*.
12. • Schouten J, Cinque P, Gisslen M, et al. HIV-1 infection and cognitive impairment in the cART era: a review. *AIDS*. 2011;25:561–75. *Up to date review of neurological management of HIV/AIDS*.
13. Stern Y, McDermott MP, Albert S, et al. Factors associated with incident human immunodeficiency virus-dementia. *Arch Neurol*. 2001;58:473–9.
14. Heaton RK, Franklin D, Woods S, et al.: Asymptomatic mild HIV-associated neurocognitive disorder increases risk for future symptomatic decline: A CHARTER longitudinal study. Presented at the conference on retroviruses and opportunistic infections. Seattle, United States of America March, 2012.
15. Power C, Selnes OA, Grim JA, McArthur JC. HIV Dementia scale: a rapid screening test. *J Acquir Immune Defic Syndr Hum Retrovirol*. 1995;8:273–8.
16. •• Cysique LA, Letendre SL, Ake C, et al. Incidence and nature of cognitive decline over 1 year among HIV-infected former plasma donors in China. *AIDS*. 2010;24:983–90. *Describes the initial algorithmic use of a cognitive screening strategy*.
17. •• Valcour V, Paul R, Chiao S, et al. Screening for cognitive impairment in human immunodeficiency virus. *Clin Infect Dis*. 2011;53: 836–42. *Existing literature review of current HAND screens. Note, this does not include all available studies, and did not differentiate between pre and post cART era or between those who used neuropsychological assesment as the gold standard*.
18. •• 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. *Comprehensive review of existing HIV Dementia Scale and International HIV Dementia Scale*.
19. Overton ET, Kauwe JS, Paul R, et al. Performances on the CogState and standard neuropsychological batteries among HIV patients without dementia. *AIDS Behav*. 2011;15:1902–9.
20. Sakamoto M, Marcotte TD, Umlauf A, et al. Concurrent classification accuracy of the HIV dementia scale for HIV-associated neurocognitive disorders in the CHARTER Cohort. *J Acquir Immune Defic Syndr Hum Retrovirol*. 2013;62:36–42.
21. Becker JT, Dew MA, Aizenstein HJ, et al. Concurrent validity of a computer-based cognitive screening tool for use in adults with HIV disease. *AIDS Patient Care STDS*. 2011;25:351–7.
22. Skinner S, Adewale AJ, DeBlock L, et al. Neurocognitive screening tools in HIV/AIDS: comparative performance among patients exposed to antiretroviral therapy. *HIV Med*. 2009;10:246–52.
23. Carey CL, Woods SP, Rippeth JD, et al. Initial validation of a screening battery for the detection of HIV-associated cognitive impairment. *Clin Neuropsychol*. 2004;18:234–48.
24. Levine AJ, Palomo M, Hinkin CH, et al. A comparison of screening batteries in the detection of neurocognitive impairment in HIV-infected Spanish speakers. *Neurobehav HIV Med*. 2011;2011:79–86.
25. Cysique LA, Maruff P, Darby D, Brew BJ. The assessment of cognitive function in advanced HIV-1 infection and AIDS dementia complex using a new computerised cognitive test battery. *Arch Clin Neuropsychol*. 2006;21:185–94.
26. Saxton J, Morrow L, Eschman A, et al. Computer assessment of mild cognitive impairment. *Postgrad Med*. 2009;121:177–85.
27. Gonzalez R, Heaton RK, Moore DJ, et al. Computerized reaction time battery versus a traditional neuropsychological battery: detecting HIV-related impairments. *J Int Neuropsychol Soc*. 2003;9:64–71.
28. Miller EN. California computerised assessment battery (CalCAP) manual. Los Angeles: Norland Software; 1990.
29. Sacktor NC, Wong M, Nakasujja N, et al. The International HIV Dementia Scale: a new rapid screening test for HIV dementia. *AIDS*. 2005;19:1367–74.
30. Ellis RJ, Evans SR, Clifford DB, et al. Clinical validation of the NeuroScreen. *J Neurovirol*. 2005;11:503–11.
31. Moore DJ, Roediger MJ, Eberly LE, et al. Identification of an abbreviated test battery for detection of HIV-associated neurocognitive impairment in an early-managed HIV-Infected Cohort. *PLoS One*. 2012;7:e47310.
32. Bottiggi KA, Chang JJ, Schmitt FA, et al. The HIV Dementia Scale: predictive power in mild dementia and HAART. *J Neurol Sci*. 2007;260:11–5.
33. Simioni S, Cavassini M, Annoni JM, et al. Cognitive dysfunction in HIV patients despite long-standing suppression of viremia. *AIDS*. 2010;24:1243–50.
34. Joska JA, Westgarth-Taylor J, Hoare J, et al. Validity of the international HIV dementia scale in South Africa. *AIDS Patient Care STDS*. 2011;25:95–101.
35. Singh D, Sunpath H, John S, Eastham J, Gouden R. The utility of a rapid screening tool for depression and HIV dementia amongst patients with low CD4 counts - a preliminary report. *African J Psychiatry*. 2008;11:282–6.
36. Price RW, Brew BJ. The AIDS dementia complex. *J Infect Dis*. 1988;158:1079–83.
37. Wonja V, Skolasky RL, McArthur JC, et al. Spanish validation of the HIV dementia scale in women. *AIDS Patient Care STDS*. 2007;21: 930–41.
38. Chalermchai T, Sithinamsuwan P, Clifford D, et al. Trail Making Test A improve performance characteristics of the International HIV Dementia Scale to identify symptomatic HAND. *J Neurovirol*. 2013;19:137–43.
39. Smith CA, van Gorp WG, Ryan ER, et al. Screening subtle HIV-related cognitive dysfunction: the clinical utility of the HIV dementia scale. *J Acquir Immune Defic Syndr Hum Retrovirol*. 2003;33:116–8.
40. Morgan EE, Woods SP, Scott JC, et al. Predictive validity of demographically adjusted normative standards for the HIV Dementia Scale. *J Clin Exp Neuropsychol*. 2008;30:83–90.
41. Mind Exchange Working Group. Assessment, diagnosis, and treatment of HIV-Associated Neurocognitive Disorder: a consensus report of the mind exchange program. *Clin Infect Dis*. 2013;56:1004–17.
42. Cherner M, Cysique L, Heaton RK, et al. Neuropathologic confirmation of definitional criteria for human immunodeficiency virus-associated neurocognitive disorders. *J Neurovirol*. 2007;13:23–8.
43. Elwood R. Psychological tests and clinical discriminations: beginning to adress the base rate problem. *Clinical Psychol Rev*. 1993;13: 409–19.

44. Cysique LA, Maruff P, Brew BJ. Variable benefit in neuropsychological function in HIV-infected HAART-treated patients. *Neurology*. 2006;66:1447–50.
45. Richardson MA, Morgan EE, Vielhauer MJ, et al. Utility of the HIV dementia scale in assessing risk for significant HIV-related cognitive-motor deficits in a high-risk urban adult sample. *AIDS Care*. 2005;17:1013–21.
46. Cysique L, Deutsch R, Atkinson J, et al. Incident major depression does not affect neuropsychological functioning in HIV-infected men. *J Int Neuropsychol Soc*. 2007;13:1–11.