Evidence-based perspectives on the implementation of screening for neurocognitive impairment in HIV

Lewis J Haddow¹
Tristan J Barber²,³
Louise Breuer⁴
Jonathan D Cartledge⁴
Katherine Pierce⁵

¹Research Department of Infection and Population Health, University College London, ²St Stephen’s AIDS Trust, Chelsea and Westminster Hospital NHS Foundation Trust, ³Faculty of Medicine, Imperial College London, ⁴Department of Genitourinary Medicine, ⁵Camden Neurological and Stroke Service, Central and North West London NHS Foundation Trust, London, UK

Abstract: This review considers the justifications, accuracy, limitations, and potential applications of screening instruments for neurocognitive impairment (NCI) in HIV-positive patients. Wilson and Jungner’s classic screening principles provide thinking tools to guide discussion: the condition should be an important health problem; there should be an accepted treatment; there should be a suitable test; and the natural history of the condition should be adequately understood. NCI appears to be common among HIV-positive patients, but evidence of its progression in those established on suppressive antiretroviral therapy is inconclusive. Also, there are limited data on the optimum management in patients who are found to have NCI but are already receiving effective antiretroviral therapy. The accuracy of screening tests, such as the HIV Dementia Scale and the Montreal Cognitive Assessment, is such that their positive and negative likelihood ratios are close to 1; therefore, many false positive and negative results will arise when using these tests in most clinical scenarios. We argue that the place for screening for neurocognitive impairment is within a management pathway that aims to identify a range of psychological and neurological problems. Widespread implementation of screening for HIV-associated neurocognitive disorders is premature, on the basis of available evidence.

Keywords: HIV, cognition disorders, neuropsychological tests, sensitivity and specificity, diagnosis

Neuropsychological health of people living with HIV

HIV is known to establish productive infection of microglia and nonproductive infection of astrocytes,¹ causing HIV-associated dementia (HAD) or HIV encephalopathy (HIVE) in some untreated patients.² Effects include cognitive, motor, and behavioral deficits, often affecting attention and executive functions. Prior to 1995, when it was progressive and fatal in the absence of treatment, HAD/HIVE was reported in observational cohorts at a rate of 6.5 cases per 1,000 patient-years.³ With the increasing availability of antiretroviral therapy (ART), the incidence, prevalence, severity, and character have changed.⁴ ⁶ Neurocognitive impairment (NCI) has been reported in 19%–69% of people living with HIV (PLWH) in recent cross-sectional studies from Western Europe and the US, with variations in prevalence depending on the population sampled and the control group.⁷ ¹⁹

There is no diagnostic test for HIV-associated neurocognitive disorder (HAND) from neuroimaging or laboratory investigations.⁷ ²¹ Diagnostic criteria have evolved over the past 3 decades to reflect the changing pathology and epidemiology of HIV
neuropathogenesis, with the Frascati criteria being the most recent.\textsuperscript{2,17,22} Diagnosis relies on neuropsychological testing, where normative neuropsychological data for the populations of interest are often lacking and diagnostic thresholds are arbitrary.\textsuperscript{23–25} Other pathologies are excluded through neuroimaging and lumbar puncture, but many HIV-positive (HIV+) patients have other psychological morbidities that confound diagnosis.

Many other mental health problems are common in PLWH, with depression perhaps the most prominent.\textsuperscript{26} Meta-analysis of ten studies comparing depression rates in HIV+ and HIV− negative individuals found approximately double the risk of depression compared with the general population, although the variation between studies was high.\textsuperscript{27} Generalized and specific anxiety disorders frequently occur in PLWH,\textsuperscript{28–30} and suicidality is also an important concern.\textsuperscript{31} Linked to these disorders, alcohol use is prevalent in some HIV+ groups, particularly men and drug users.\textsuperscript{32} Ilicit drug use has long been associated with the AIDS epidemic, with injecting drug users being a core risk group. Heavy use of recreational drugs is recognized as being prevalent in HIV+ men who have sex with men, often in the context of higher risk sexual behavior.\textsuperscript{33} Drugs of abuse, particularly methamphetamine, opiates, and cocaine, have overlapping toxicities with the neuropathogenesis of HIV.\textsuperscript{34,35} These patterns of mental health may be caused by the stress, stigma, and social difficulties of living with HIV, or they may precede or coassociate with the infection. The overlap between mental disorders and NCI may both increase the risk of HAND and confound its diagnosis.

In future, there may be a higher rate of cognitive decline in PLWH resulting from incidental neurodegenerative disease. As the survival of treated HIV+ populations approaches that of the general population,\textsuperscript{36} the proportion living to middle and older age increases.\textsuperscript{37,38} Those in age groups at highest risk of Alzheimer’s disease are still relatively uncommon. HIV’s association with chronic inflammation, other cardiovascular diseases, endothelial dysfunction,\textsuperscript{39} and ischemic stroke\textsuperscript{40,41} suggests a higher risk of vascular dementia in PLWH.

Outcomes and impacts of HIV-associated NCI in treated patients

The first of Wilson and Jungner’s criteria\textsuperscript{42} is the condition sought should be an important health problem. The high prevalence of NCI in cross-sectional studies suggests that it is a common, persistent problem even in treated HIV.\textsuperscript{43} Notwithstanding the uncertainty surrounding its true prevalence, NCI is an important concern for patients and clinicians. Patients with good virological control are still at risk of mild-to-moderate impairment, and complications include increased mortality,\textsuperscript{44} impaired quality of life,\textsuperscript{45} impaired everyday function,\textsuperscript{46} and poor adherence to ART.\textsuperscript{43}

It is now apparent that ART can prevent or reverse the disease process, and the specter of HAD has diminished since the 1980s.\textsuperscript{6} This is supported by population-level cohort data\textsuperscript{14} and individual-level results in which the effects of ART were measured relative to a pretreatment baseline. The effects are reassuring in patients with significant NCI before treatment, with 30%–70% experiencing improved cognitive function in the first few years of ART.\textsuperscript{46}

Without wishing to minimize the potential impact of NCI on PLWH, there are also reassuring data suggesting that cognitive function generally remains stable over 1–5 years. Cole et al\textsuperscript{47} analyzed data from 345 stable HIV+ patients in US Multicenter AIDS Cohort Study with a brief cognitive and psychomotor assessment and found preserved function over 5 years compared to controls. Recent data from the Multicenter AIDS Cohort Study, using a more comprehensive neuropsychological battery, found little change in the prevalence of HAND from 2007 to 2012 and minimal progression over 4 years.\textsuperscript{48} In a 1-year study in the People’s Republic of China\textsuperscript{49} and a 3-year study in the US,\textsuperscript{50} both cohorts including a high proportion of nonvirologically suppressed patients, the proportion of patients who declined in neurocognitive function was similar to the proportion who improved. Perhaps controversially, there was a considerable amount of fluctuation in neurocognitive performance.\textsuperscript{51} In two 2-year cohorts from Italy and France with 95% ART coverage, the prevalence of NCI was very similar both at baseline and at study completion.\textsuperscript{52,53} An earlier US cohort study, enrolled between 2001 and 2005, found declining cognitive function in 35% of the 215 patients followed up (of 276 originally enrolled), but there was a 56% prevalence of “substance abuse or dependence”, and treatment and adherence may have been suboptimal by current standards.\textsuperscript{54}

Other classic screening criteria are that the natural history of a condition, including its development from silent to clinical disease, should be well-understood, and there should be a recognizable latent or early symptomatic stage. Beyond the beneficial effects of ART, physicians are faced with uncertainty about how patients’ cognitive difficulties will evolve over time, and more longitudinal studies are needed. There are reports that patients with asymptomatic NCI (ANI)\textsuperscript{57} experience further decline in cognitive function. In a multicenter US HIV cohort study (n=387), there was a
substantially greater rate of change from ANI to more severe grades of HAND than in those without NCI. Similar results were found in a Canadian cohort where virological suppression rates were much higher than in the US study, but there was still a 32% risk of progression from ANI to other grades of HAND, compared to 17% risk of progression from normal cognitive function to symptomatic HAND. Limitations to both these studies lie in the use of “time to event” analysis, which in this context assumes that once a patient changes from ANI to more severe grades of HAND, they cannot revert to having ANI or normal cognitive function. As a result, the analysis will capture greater variability in function as well as faster decline. Also in the Canadian study, progression was determined from participants’ self-reported abilities, which could have been significantly influenced by mood and other factors.

Treatment options for HAND

Another important prerequisite for screening is that there should be an accepted treatment and there should be an agreed policy on whom to treat. As already noted, ART is highly effective in reversing HIV, and as a result, HIV treatment guidelines have included symptomatic HAND as an indication for ART. But the Strategic Timing of AntiRetroviral Treatment study suggests that initiation of ART is not the issue: we should be treating everyone with HIV, and the results of neuropsychological assessments have no bearing on whether or not a patient should be on ART. There is a case to be made for HIV testing in individuals with NCI who are not already diagnosed, but that argument has been strongly made elsewhere.

If the benefits of ART in treating or preventing HAND or HIV are uncontroversial, we must consider whether there is a case for screening patients already on ART so that other therapeutic interventions can be offered. Considerations include whether any particular ART drug or regimen affects neurocognitive outcomes in patients with demonstrable NCI, or in PLWH in general; whether additional benefit can be achieved from adjunctive treatments; and whether more general approaches, such as management of comorbidities and psychological and pharmacological therapies can improve outcomes. Regarding the last point, there is a benefit in identifying individuals with functional impairment because of the opportunities to help those individuals in other ways, such as social support and compensatory neurorehabilitation. But for those assessed as having ANI, there is no clear path for treatment despite evidence of an increased risk of progression.

The central nervous system (CNS) penetration effectiveness score grades the ability of antiretroviral drugs to suppress viral replication in the CNS by achieving adequate tissue concentrations. Similarly, the monocyte efficacy score grades drugs on their ability to achieve high intracellular concentrations in the macrophage-monocyte cell line, which includes microglia. To date, there have been inconsistent findings regarding the effects of drugs with different CNS penetrations. The inevitable biases that emerge in these observational studies make their findings difficult to interpret. In an attempt to overcome these biases, a randomized treatment modification trial compared high- and low-penetration regimens but was terminated early because of under-recruitment. Concerns have been raised about the effectiveness of boosted protease inhibitor (PI) monotherapy because of PIs’ relatively poor penetration into the CNS, but two trials reported no differences in neurocognitive function between patients randomized to PI monotherapy and combination antiretroviral therapy over at least 96 weeks of follow-up. While British and European HIV treatment guidelines recommend any standard combination antiretroviral therapy regimen in most patients with NCI, American guidelines name dolutegravir and darunavir as preferred options. There are also potential adverse neuropsychological effects of ART, particularly efavirens and efavirenz-containing regimens should be avoided in those with neurocognitive difficulties.

Some patients with NCI have an underlying phenomenon of detectable virus in cerebrospinal fluid (CSF) but undetectable virus in plasma, often known as CSF viral escape or discordance. This has been described in case series by Canestri et al. Given the association between the phenomenon and neurocognitive or neurological syndromes and the theoretical risk of mutation and “spill-over” of resistant strains into the wider circulation, it is generally recommended that attempts are made in such patients to genotype the virus in both the CSF and plasma and adjust the ART regimen accordingly. One could view screening for NCI as a method of ultimately funneling down to identify such individuals.

Trials of adjunctive pharmacological treatment, including minocycline, selegiline, valproate, rivastigmine, and lithium, have not shown benefits in cognitive function in PLWH. It may be that the trials were underpowered or too short in duration to show benefit, or that too heterogeneous a group of patients was targeted. Or, given current gaps in knowledge on neuropathogenesis in patients receiving virologically suppressive ART, it may be that the wrong interventions were chosen.

HIV-specific screening tools for neurocognitive disorders

Screening also requires a suitable test or examination, which should be acceptable to the population.
screening tool was the HIV Dementia Scale (HDS), published in 1995 and intended as “a brief but sensitive instrument to identify dementia (in HIV+ patients)”.87 The scale comprises four subtests (Table 1), targeted mainly at subcortical cognitive processes. A modified HDS, which omits the antisaccadic errors subtest, has also been proposed.88 Development of the scale involved assessment of 101 HIV+ and 29 seronegative participants in 1991–1992.

There have been four systematic reviews and three meta-analyses (Table 2) of the accuracy of the HDS.89–92 It is clear from the pooled estimates that the sensitivity of the scale is low and the specificity is only slightly better, while differences in methodology and study selection have led to heterogeneous estimates between meta-analyses and individual studies. The HDS is designed to be highly standardized, so heterogeneity is more likely to result from the diverse range of testing batteries, case definitions, and diagnostic thresholds used in the “gold standard”. One can resolve this heterogeneity by reducing sensitivity and specificity into thresholds used in the “gold standard”. One can resolve this heterogeneity by reducing sensitivity and specificity into a single diagnostic odds ratio (DOR), at the expense of a considerable amount of useful information. The results of two of the meta-analyses give the DORs of 5.88 and 7.52.89,91

The three reviews have highlighted problems with study quality in diagnostic accuracy studies. The first major concern was the possibility of validation bias, which occurs when assessors in a diagnostic accuracy study are not blinded to the results of the other test, ie, they are not blinded to the results of the HDS when carrying out the neuropsychological battery and they are not blinded to the full battery when conducting the HDS. Validation bias can lead to falsely high estimates of accuracy. The second major concern was participant selection, which in almost all studies involved some exclusion of patients with competing psychiatric, neurological, or systemic conditions. By excluding such individuals, diagnostic accuracy studies create an artificially “clean” study sample, leading to an overestimate of diagnostic accuracy and results that may not be generalizable to real clinical scenarios. The use of such entry criteria, in conjunction with nonrandom, nonconsecutive sampling strategies, is likely to have led to overestimation of accuracy.93

It was highlighted in the review by Haddow et al that almost all published studies of the HDS were conducted in North America,93 although the scale has been specifically tested in Spanish speakers.94,95 In a Zambian study, it was observed that all 48 HIV+ participants and 15 presumed seronegative controls tested positively on the HDS.96 In response to difficulties with culturally specific elements of the HDS and the administration of the antisaccadic errors subtest, the

### Table 1 Subtests of the HIV Dementia Scale and International HIV Dementia Scale

<table>
<thead>
<tr>
<th>Subtest</th>
<th>HIV Dementia Scale</th>
<th>International HIV Dementia Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall of four items at 5 minutes</td>
<td>Recall memory</td>
<td>Recall memory</td>
</tr>
<tr>
<td>Timed written alphabet</td>
<td>Psychomotor speed</td>
<td>Finger tapping</td>
</tr>
<tr>
<td>Cube copy</td>
<td>Visuospatial/constructional</td>
<td>Fist–palm–side test (Luria test)</td>
</tr>
<tr>
<td>Antisaccadic errors</td>
<td>Attention</td>
<td>Psychomotor speed and executive function</td>
</tr>
<tr>
<td>Maximum score</td>
<td>16</td>
<td>Maximum score</td>
</tr>
<tr>
<td>Score required for positive test</td>
<td>≤10</td>
<td>Score required for positive test</td>
</tr>
</tbody>
</table>

### Table 2 Summary of three meta-analyses of the accuracy of the HDS and IHDS

<table>
<thead>
<tr>
<th>Reference</th>
<th>Standard diagnosis</th>
<th>Studies pooled</th>
<th>Pooled sensitivity (%)</th>
<th>Pooled specificity (%)</th>
<th>Positive likelihood ratio*</th>
<th>Negative likelihood ratio*</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDS studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hu et al 91</td>
<td>All HAND</td>
<td>7 (648)</td>
<td>61</td>
<td>79</td>
<td>2.90</td>
<td>0.49</td>
</tr>
<tr>
<td>Zipursky et al 90</td>
<td>All HAND</td>
<td>8 (1,338)</td>
<td>48</td>
<td>ND b</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>Haddow et al 99</td>
<td>MND or HAD</td>
<td>10 (2,291)</td>
<td>42</td>
<td>91</td>
<td>4.77</td>
<td>0.64</td>
</tr>
<tr>
<td>Haddow et al 99</td>
<td>HAD</td>
<td>13 (1,277)</td>
<td>68</td>
<td>78</td>
<td>3.08</td>
<td>0.41</td>
</tr>
<tr>
<td>IHDS studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hu et al 91</td>
<td>All HAND</td>
<td>5 (296)</td>
<td>64</td>
<td>59</td>
<td>1.56</td>
<td>0.61</td>
</tr>
<tr>
<td>Zipursky et al 90</td>
<td>All HAND</td>
<td>4 (457)</td>
<td>62</td>
<td>ND b</td>
<td>b</td>
<td></td>
</tr>
<tr>
<td>Haddow et al 99</td>
<td>MND or HAD</td>
<td>11 (416)</td>
<td>64</td>
<td>66</td>
<td>1.89</td>
<td>0.54</td>
</tr>
<tr>
<td>Haddow et al 99</td>
<td>HAD</td>
<td>5 (808)</td>
<td>74</td>
<td>55</td>
<td>1.64</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Notes: *Positive likelihood ratio = sensitivity/(1–specificity) and negative likelihood ratio = (1–sensitivity)/specificity. Zipursky et al did not present a pooled specificity rate; therefore, likelihood ratios could not be calculated. Abbreviations: HAD, HIV-associated dementia; HAND, HIV-associated neurocognitive disorder; HDS, HIV Dementia Scale; IHDS, International HIV Dementia Scale; MND, minor neurocognitive disorder; ND, not done.
International HIV Dementia Scale (IHDS) was developed as a “cross-cultural screening instrument” for resource-limited settings (Table 1).97 The IHDS was originally developed by analyzing two HIV+ patient samples, based in the US (n=66) and Uganda (n=81). Most subsequent studies of its accuracy have been conducted in low- and middle-income countries. It has been modified for use in several different languages.97–103

While most studies report sensitivity and specificity as their main results, these statistics only tell us how the result of the test is likely to turn out, given certain assumptions about the result of the gold standard diagnosis. In clinical practice, one usually wishes to know the likelihood of a patient truly having or not having a condition in the face of a positive or negative test. To derive this, two parameters are needed: the prior probability that the condition is present and either a positive LR (LR+) or a negative LR (LR−), depending on the result of the test (formula given in the footnote of Table 2). According to Jaeschke et al, tests with LR+ >10 and LR− <0.1 are needed to provide “convincing” diagnostic evidence.105

We can consider the situation of a case where the HDS or IHDS is used in a setting of high background prevalence, eg, where one-third of patients screened are considered to be affected by HAND (a probability of 33.3%; odds of “2:1 against” or 0.5), and apply rough estimates of LR+ and LR− derived from the reviews cited earlier. Given a positive HDS, the posttest odds are 1.5 or 3:2 in favor, so the probability of the patient having HAND is 60%. A positive IHDS results in the posttest odds of 0.8 or 5:4 against, a posttest probability of 44%. On the other hand, a negative HDS would lead to posttest odds of 4:1 against or a probability of 20%, and a negative IHDS would give the posttest odds of 0.3 or 10:3 against, a probability of 23%. Table 3 summarizes this scenario and the following two other scenarios: high clinical suspicion where the tests are used diagnostically and prior probability is assumed to be 66.7%, and screening in a lower prevalence setting. The extent to which either the HDS or IHDS advances the diagnostic process in these scenarios is debatable.

### Non-HIV-specific neurocognitive screening tools used to assess HIV+ patients

The mini–mental state Examination (MMSE) is a long-established tool for rapidly assessing cognitive function.106 Although frequently used in acute medicine and geriatric care, the tool has found little application in studies of HAND. An important study, albeit small in sample size (33 HIV+ and 13 seronegative participants), was that of Skinner et al, in which participants were tested with the MMSE, HDS,

| Table 3 Calculation of posttest probability of HAND using the HDS and IHDS, under three different fictional scenarios |
|---------------------------------------------------|--------------------|-----------------|-----------------|
| Scenario                                           | Assumed pretest probability (%) | Pretest odds     | Test and result* | Posttest odds | Posttest probability (%) |
| Higher prevalence screening                        | 33.3                | 0.5 (2:1 against)| HDS positive     | 1.5           | 60                        |
|                                                   |                     |                 | HDS positive     | 0.8           | 44.4                      |
|                                                   |                     |                 | HDS negative     | 0.25          | 20                        |
|                                                   |                     |                 | IHDS positive    | 0.3           | 23.1                      |
|                                                   |                     |                 | IHDS negative    | 1.0           | 50                        |
| Clinical suspicion                                 | 66.7               | 2.0 (2:1 on)    | HDS positive     | 6.0           | 85.7                      |
|                                                   |                     |                 | HDS positive     | 3.2           | 76.2                      |
|                                                   |                     |                 | HDS negative     | 1.2           | 54.5                      |
|                                                   |                     |                 | IHDS positive    | 0.333         | 25                        |
|                                                   |                     |                 | IHDS positive    | 0.178         | 15.1                      |
|                                                   |                     |                 | IHDS negative    | 0.0556        | 5.3                       |
|                                                   |                     |                 | IHDS negative    | 0.0667        | 6.2                       |
| Lower prevalence screening                         | 10                 | 0.111 (9:1 against) | HDS positive     | 0.0667        | 6.2                       |
|                                                   |                     |                 | IHDS positive    | 0.0556        | 5.3                       |
|                                                   |                     |                 | IHDS negative    | 0.0667        | 6.2                       |

*Assumptions in the table are that positive likelihood ratios are 3.0 for the HDS and 1.6 for the IHDS and that negative likelihood ratios are 0.5 for the HDS and 0.6 for the IHDS. These assumptions are approximations of the results of published meta-analyses,89–91 rounded off for easier arithmetic.

Abbreviations: HAND, HIV-associated neurocognitive disorder; HDS, HIV Dementia Scale; IHDS, International HIV Dementia Scale.
and IHDS. The MMSE showed a sensitivity of 46% and a specificity of 55% at the standard cut-off of <27, and the area under the receiver–operator curve was 0.48. These statistics suggest that the MMSE is little better than chance in its discriminatory power for detecting HAND.

There is an expanding literature on the Montreal Cognitive Assessment (MoCA) in the assessment of HIV+ patients (Table 4). This tool includes brief tests of construction (cube copy and clock face), picture naming, learning and recall of five words, digit span forward and backward, verbal fluency, abstraction, and attention and orientation in time and place. Although it may take longer to perform than the HDS and IHDS and requires a preprinted score card, it is highly standardized, easy to conduct, and freely available online (www.mocastest.org). The MoCA has a wider dynamic range of scores than the shorter screening tools, and different thresholds have been studied. In a study by Joska et al., the MoCA showed higher sensitivity and lower specificity than any of the other tests used in the study (including the HDS, IHDS, MMSE, and Simioni symptom questionnaire [SSQ]). A study in the US found results close to those of Joska, while work in the Netherlands showed similar high sensitivity and low specificity only when used at a cut-off of 28.5; at standard cut-off, both statistics were closer to 60%. In the Dutch study, the area under the receiver–operator curve suggested that the MoCA was diagnostically similar to the HDS. In contrast, several other studies, including one recruiting only participants aged >60 years and one using a locally modified version in Korea, have also reported lower sensitivity but higher specificity. Thus, there is a wide range of diagnostic accuracy estimates for the MoCA, similar to the HDS and IHDS. In all studies, the LR+ is <5 and the LR− is >0.2, implying that the test is limited in its diagnostic power.

The MoCA is also limited by its cultural specificity, as illustrated by work conducted in a Xhosa-speaking sample of HIV+ patients and seronegative controls in South Africa. In that study, the mean score was well below the standard cut-off for impairment in both groups. The study challenged assumptions about precisely which items in a screening tool may not have cross-cultural validity: in a picture-naming task, the rhinoceros, an animal indigenous to South Africa, was frequently misnamed as an elephant, buffalo, or hippopotamus. Also, the language tasks can be particularly difficult to transpose into non-Indo-European languages, such as Xhosa and Korean.

The MoCA has its limitations, but its main advantage over the HDS and IHDS may be that it provides more qualitative information and covers a wider range of cognitive tasks, thus contributing to the “narrative” around a patient’s neurocognitive symptoms and abilities.

### Short neuropsychological testing batteries used as screening tools

Another screening approach is the use of focused neuropsychological batteries. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) is a popular format in some centers, but there are few published results of its use specifically in HIV+ populations, and it lacks sufficient assessment of executive function. Other short batteries include the neurocognitive HIV study screen, the brief neurocognitive screen, the NeuroScreen, the high sensitivity cognitive screen, and multiple other short combinations of standardized tests, which have been reviewed systematically in two publications. Computerized tools, such as Cogstate and CalCAP, may be used as screening instruments or as more detailed batteries. Despite widespread use in clinical research, computerized assessment is infrequently encountered in routine practice.

There does not appear to be consensus on the best short battery to use in HIV. Most tools published have >70% sensitivity but lack specificity to be able to distinguish HAND from other conditions, and the overall quality of published diagnostic accuracy studies is not high. The main advantages of brief batteries may be their application in multicenter cohorts or trials and their potential role in task-shifting assessments away from neuropsychologists, either as a preliminary or

<table>
<thead>
<tr>
<th>Country and citation</th>
<th>Sample size</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive LR*</th>
<th>Negative LR*</th>
<th>DOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Africa and USA</td>
<td>156</td>
<td>89–100</td>
<td>22</td>
<td>1.14</td>
<td>0.50</td>
<td>2.28</td>
</tr>
<tr>
<td>USA</td>
<td>100</td>
<td>85</td>
<td>40</td>
<td>1.42</td>
<td>0.38</td>
<td>3.78</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>102</td>
<td>56</td>
<td>63</td>
<td>1.51</td>
<td>0.70</td>
<td>2.17</td>
</tr>
<tr>
<td>USA</td>
<td>200</td>
<td>63</td>
<td>71</td>
<td>2.17</td>
<td>0.52</td>
<td>4.17</td>
</tr>
<tr>
<td>USA</td>
<td>67</td>
<td>50</td>
<td>85</td>
<td>3.33</td>
<td>0.59</td>
<td>5.67</td>
</tr>
<tr>
<td>South Korea</td>
<td>194</td>
<td>53</td>
<td>73</td>
<td>1.99</td>
<td>0.64</td>
<td>3.10</td>
</tr>
</tbody>
</table>

Notes: *Range of sensitivity estimates relates to differing thresholds in the gold standard diagnosis. Positive likelihood ratio = sensitivity/(1–specificity) and negative likelihood ratio = (1–sensitivity)/specificity.

Abbreviations: DOR, diagnostic odds ratio; LR, likelihood ratio; MoCA, Montreal Cognitive Assessment.
as a full assessment. At this level of intensity, assessments occupy a place that is intermediate between screening and diagnostic testing.

**Neuropsychological symptom screening in HIV+ patients**

Perhaps the simplest approach to screening for HAND is through brief, structured symptom questionnaires. Arguably, screening that targets unrecognized symptomatic disease (not just true asymptomatic disease) is still screening.42 One widely used method is known as the SSQ, after the first author of study in which it was first published,12 although other short questionnaires in use include the four cognitive impairment symptom questions from the Medical Outcomes Study (MOS-HIV).120

Simioni et al used data from 50 Swiss patients who complained of symptoms affecting memory, concentration, or reasoning and 50 who were symptom free, all with undetectable viral loads. There was an unusually high prevalence of HAND (69%) and fairly poor accuracy for the SSQ (sensitivity 57% and specificity 69%).12 In unpublished work from a British and Irish study of HIV+ patients aged >50 years, one of the three questions of the SSQ, concerning memory, showed a statistically significant association with neurocognitive performance measured using Cogstate.121 Other works, such as the Spanish neurocognitive HIV study115,122 and analysis of a German-language version,123 have shown poor accuracy, probably no better than chance. The SSQ entered the European AIDS Clinical Society treatment guidelines in 2011 as a recommended tool for screening for NCI,58 but this has since been changed to a recommendation “to guide physician assessment” in 2015.124 British guidelines recommend the questionnaire as one possible screening tool for the routine identification of psychological support needs.125

The MOS-HIV cognitive impairment questions were initially validated by Knippels et al, who showed an association between questionnaire responses and neuropsychological test performance in most outcome cognitive measures (correlation with overall neuropsychological performance: $\beta$ coefficient $=0.252, P<0.01$).126 In a large four-nation study using a shorter computerized assessment, there was no association between the MOS-HIV questions and neuropsychological test performance on Cogstate.127

The relative lack of data on short symptom questionnaires and the volume of ongoing research in this area indicate that conclusions about their appropriate use in screening may be premature. And, it is unlikely that broad symptom questions will be in any way HIV specific, with respect to any other cause of NCI.

**Conclusion and perspectives on the practical implementation of screening for HIV-associated NCI**

If screening for NCI is indeed necessary in HIV+ patients, we clearly lack the evidence needed to implement it. Since the recommendation to screen passed into national and international guidelines, care providers have been grappling with questions around how to screen: how often, with which tools, and in whom. Furthermore, none of the tools have been assessed for practice effects, which have implications for periodic screening. Reassuringly, according to the limited results available, neurocognitive screening is acceptable to patients.128 However, while earlier European guidelines suggested specific screening intervals of 2 years in most patients and 1 year in higher risk patients,58 a blanket screening approach has been dropped in favor of active case finding.124 One possible strategy that has not been addressed in clinical research is selective screening, limited to patients at high risk of HAND, such as those with a lower nadir CD4 count or other risk factors.129,130

After many years of research, we still lack a suitable screening tool for HAND. The HDS lacks sensitivity, and the IHDS lacks specificity; neither has been tested in a wide enough cultural context. The MoCA gives more detail but is no more accurate than the HDS or IHDS, and available data are similarly heterogeneous. A diverse range of short batteries has been tried, but in general, each battery has only been tested in a single study, and none are significantly better than the shorter tools to justify the additional training and resource requirements. Symptom screening questions appear to be the least accurate and have the least published results of all the approaches, although they have the advantage of being easiest to administer in clinical practice. It may be that the diagnosis is too complex for a bedside tool. It is unlikely that a sufficiently low LR– will be achieved to remove the need for more detailed neuropsychological assessment or brain imaging in patients in whom there is a significant clinical suspicion. And with only a modest LR+, a positive screening test is unlikely to justify intensive investigation in patients with minimal symptoms or clinical suspicion of HAND.

It seems that, both from the available evidence and from clinical experience, symptoms suggestive of NCI are non-specific. In all likelihood, patients with symptoms affecting memory, concentration, or reasoning may have depression,
anxiety, substance misuse, stressful life events, or simply normal aging. Screening for these other psychological concerns is a laudable aim, and one should not be dissuaded from using short symptom questionnaires in routine practice. However, widespread use of screening may cause unnecessary anxiety among patients, and too great a focus on HAND may lead to hypervigilance for normal everyday memory lapses in some patients. Furthermore, given the aging of the HIV+ population and the omission of competing diagnoses from many studies of tools’ accuracy, we have little evidence that the HDS, IHDS, and other tests are able to distinguish HAND from other causes of NCI. Screening should be used responsibly and with appropriate counseling, bearing in mind the potential inaccuracy of the tools. Clinicians should not view them as filters solely to identify possible HAND. Instead, a positive result on tools such as the SSQ opens up wider lines of neurological and psychological enquiries, guiding clinicians to explore possible mood disorders, substance use, and so on in more detail.

Used correctly, screening tests are the first step in a clinical management pathway. They provide an opportunity to review patients’ mental health and any concerns that they may have relating to cognitive function. They should be viewed as a part of the assessment of a wide range of psychological concerns and not just as quick or preliminary diagnostic tests for HAND. An evidence-based clinical pathway may be a more useful end goal than further refinements to screening tools. Future studies should go beyond estimates of diagnostic accuracy and test the utility of a multistage assessment pathway in a representative clinical setting.

Disclosure

The authors report no conflicts of interest in this work.

References


