Published in final edited form as:

J Acquir Immune Defic Syndr. 2013 January 1; 62(1): 36–42. doi:10.1097/QAI.0b013e318278ffa4.

Concurrent Classification Accuracy of the HIV Dementia Scale for HIV-associated Neurocognitive Disorders in the CHARTER Cohort

Maiko Sakamoto, PhD¹, Thomas D. Marcotte, PhD¹, Anya Umlauf, MS¹, Donald Franklin Jr.¹, Robert K. Heaton, PhD¹, Ronald J. Ellis, MD, PhD², Scott Letendre, MD³, Terry Alexander, RN¹, J. Allen McCutchan, MD, MSc³, Erin E. Morgan, PhD¹, Steven Paul Woods, PsyD¹, Ann C. Collier, MD⁴, Christina M. Marra, MD⁴, David B. Clifford, MD⁵, Benjamin B. Gelman, MD, PhD⁶, Justin C. McArthur, MBBS⁷, Susan Morgello, MD⁸, David Simpson, MD⁹, Igor Grant, MD¹, and CHARTER group

¹Dept. of Psychiatry, University of California, San Diego, San Diego, CA

²Dept. of Neurosciences, University of California, San Diego, San Diego, CA

³Dept. of Medicine, University of California, San Diego, San Diego, CA

⁴Dept. of Medicine, University of Washington, Seattle, WA

⁵Dept. of Neurology, Washington University in St. Louis, St. Louis, MO

⁶Dept. of Pathology, University of Texas Medical Branch, Galveston, TX

⁷Dept. of Neurology, Johns Hopkins University, Baltimore, MD

⁸Dept. of Neurology, Neurosciences and Pathology, Mount Sinai School of Medicine, New York, NY

⁹Dept. of Neurology, Mount Sinai School of Medicine, New York, NY

Correspondence and reprint requests to: Thomas D. Marcotte, Ph.D., Department of Psychiatry, University of California, San Diego, 220 Dickinson Street, San Diego, CA 92103, USA. Phone number: (619) 543-5044, Fax number: (619) 543-1235. Electronic mail may be sent to tmarcotte@ucsd.edu..

*The CNS HIV Anti-Retroviral Therapy Effects Research (CHARTER) group is affiliated with the Johns Hopkins University, Mount Sinai School of Medicine, University of California, San Diego, University of Texas, Galveston, University of Washington, Seattle, Washington University, St. Louis and is headquartered at the University of California, San Diego and includes: Director: Igor Grant, M.D.; Co-Directors: Scott L. Letendre, M.D., Ronald J. Ellis, M.D., Ph.D., Thomas D. Marcotte, Ph.D.; Center Manager: Donald Franklin, Jr.; Neuromedical Component: Ronald J. Ellis, M.D., Ph.D. (P.I.), J. Allen McCutchan, M.D., Terry Alexander, R.N.; Laboratory, Pharmacology and Immunology Component: Scott Letendre, M.D. (P.I.), Edmund Capparelli, Pharm.D.; Neurobehavioral Component: Robert K. Heaton, Ph.D. (P.I.), J. Hampton Atkinson, M.D., Steven Paul Woods, Psy.D., Matthew Dawson; Virology Component: Davey M. Smith, M.D. (P.I.); Imaging Component: Christine Fennema-Notestine, Ph.D. (Co-P.I.), Michael J. Taylor, Ph.D. (Co-P.I.), Rebecca Theilmann, Ph.D.; Data Management Component: Anthony C. Gamst, Ph.D. (P.I.), Clint Cushman; Statistics Component: Ian Abramson, Ph.D. (P.I.), Florin Vaida, Ph.D.; Protocol Coordinating Component: Thomas D. Marcotte, Ph.D. (P.I.), Rodney von Jaeger, M.P.H.; Johns Hopkins University Site: Justin McArthur (P.I.), Kristen Lewis; Mount Sinai School of Medicine Site: Susan Morgello, M.D. (Co-P.I.) and David Simpson, M.D. (Co-P.I.), Letty Mintz, N.P.; University of California, San Diego Site: J. Allen McCutchan, M.D. (P.I.), Will Toperoff, N.P.; University of Washington, Seattle Site: Ann Collier, M.D. (Co-P.I.) and Christina Marra, M.D. (Co-P.I.), Trudy Jones, M.N., A.R.N.P.; University of Texas, Galveston Site: Benjamin Gelman, M.D., Ph.D. (P.I.), Eleanor Head, R.N., B.S.N.; and Washington University, St. Louis Site: David Clifford, M.D. (P.I.), Muhammad Al-Lozi, M.D., Mengesha Teshome, M.D. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Government.

Conflicts of Interest The authors declare that they have no conflict of interest.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Abstract

Background—The HIV Dementia Scale (HDS) was developed to screen for HIV-associated Neurocognitive Disorders (HAND), but concerns have persisted regarding its substandard sensitivity. This study aimed to examine the classification accuracy of the HDS using raw and norm-based cutpoints, and to evaluate the contribution of the HDS subtests to predicting HAND.

Methods—1,580 HIV-infected participants from 6 U.S. sites completed the HDS, and a gold standard neuropsychological battery, on which 51% of participants were impaired. Results: Sensitivity and specificity to HAND using the standard raw HDS cutpoint were 24% and 92%, respectively. The raw HDS subtests of attention, recall, and psychomotor speed significantly contributed to classification of HAND, while visuomotor construction contributed the least. A modified raw cutpoint of 14 yielded sensitivity of 66% and specificity of 61%, with cross-validation. Using norms also significantly improved sensitivity to 69% with a concomitant reduction of specificity to 56%, while the positive predictive value declined from 75% to 62% and negative predictive value improved from 54% to 64%. The HDS showed similarly modest rates of sensitivity and specificity among subpopulations of individuals with minimal comorbidity and successful viral suppression.

Conclusions—Findings indicate that while the HDS is a statistically significant predictor of HAND, particularly when adjusted for demographic factors, its relatively low diagnostic classification accuracy continues to hinder its clinical utility. A raw cutpoint of 14 greatly improved the sensitivity of the previously established raw cutscore, but may be subject to ceiling effects, particularly on repeat assessments.

Keywords

HIV; cognition; HIV-associated neurocognitive disorders; screening measures; HIV dementia scale

Introduction

HIV-associated neurocognitive disorders (HAND) are commonly observed among persons living with HIV infection^{1, 2}. With the introduction of combination antiretroviral therapy (cART) in the mid-1990s, the incidence of HIV-associated dementia (HAD), the most severe form of HAND, drastically decreased; however, milder forms of HAND remain quite common, affecting up to 50% of persons with HIV³⁻⁵.

The HIV Dementia Scale (HDS) was developed in 1995 as a brief screening tool for HAND. It was originally intended to assess a range of HAND severity, from mild deficits to frank dementia⁶⁻¹². The original study suggested that a raw cutpoint of 10 provided 80% of sensitivity and 91% of specificity for detection of HAD; however, concerns have been raised regarding the degree to which the HDS accurately identifies the milder neurocognitive impairments that are prevalent in the cART era⁵. Using a raw cutpoint of 10, studies found adequate-to-excellent specificity (e.g., 65 – 94%) but very poor-to-good sensitivity (e.g., 17 – 80%)^{7, 8, 10-12}. A recent study¹³ suggested that a raw score cutpoint of 14 would provide improved sensitivity to HAND of 83%, but specificity of only 63%.

Most of these studies used HDS raw scores; however, it is well-documented that performance on NP tests is related to demographic factors (e.g., age)¹⁴ and the HDS is no exception^{7, 15}. Morgan and colleagues developed normative standards, correcting for age and education¹⁶ and demonstrated improved diagnostic utility of the HDS. Although Morgan et al's study supports the value of the use of demographically-corrected norms, it has not been subjected to any large scale replication.

The goals of the current study were to 1) examine the classification accuracy of the HDS (raw and adjusted values) for HAND as determined by a comprehensive NP battery in a large group of HIV+ participants, 2) determine whether a revised raw cutpoint would improve the classification accuracy of HAND, and 3) evaluate the contribution of the HDS subtests to the classification of HAND.

Methods

Participants and Procedure

The 1,580 HIV-infected individuals in this study were drawn from six university research sites of the CNS HIV Anti-Retroviral Therapy Effects Research (CHARTER) study (Table 1). To increase representativeness, there were no formal exclusion criteria for participation in the CHARTER study. Table 2 shows the demographic, psychiatric, and medical characteristics of the study samples. All participants completed a neuromedical assessment, comprehensive neuropsychological testing, and a psychiatric interview. These procedures were approved by the Human Subjects Protection Committees of each participating institution. Written informed consent was obtained from all participants.

Neuromedical evaluation

HIV infection was diagnosed by enzyme linked immunosorbent assay and western blot confirmatory test. HIV RNA levels from plasma and CSF were measured centrally by reverse transcriptase PCR (Roche Amplicor, v. 1.5, lower limit of quantitation 50 copies/mL).

Comorbidity conditions

In order to address the role of non-HIV factors in cognition (e.g., traumatic brain injuries, epilepsy, etc), comorbidity conditions were reviewed and participants were assigned to one of three subgroups – incidental (mild), contributing (moderate), and confounding (severe) comorbid groups, based on the likelihood that condition(s) would significantly affect cognitive functioning^{1, 2}.

“Gold Standard” NP measures

The comprehensive NP test battery (administration time = 2-2.5 hours) served as a gold standard for determination of cognitive impairment. The battery was administered by trained and certified clinicians or psychometrists: Verbal Fluency [Controlled Oral Word Association Test (COWAT-FAS^{17, 18}) and Animal Fluency¹⁹]; Speed of Information Processing [WAIS-III Digit Symbol and Symbol Search subtests²⁰ and Trail Making Test (TMT) Part A^{14, 21}]; Executive Functioning [Wisconsin Card Sorting Test (WCST 64-item version²²) perseverative responses and completed categories, and TMT Part B^{14, 21}]; Working Memory (WAIS-III Letter-Number Sequencing²⁰, Paced Auditory Serial Addition Test-50²³); Learning (Story Learning¹⁴ and Figure Learning¹⁴); Memory (Story Loss¹⁴ and Figure Loss¹⁴); and Motor (Grooved Pegboard Test^{14, 24}).

Raw scores were converted into demographically-corrected T-scores (age-, education-, sex-, and ethnicity-corrected) to determine NP status (NP normal vs. NP impaired). T-scores of less than 40 were considered indicative of cognitive impairment for individual tests¹⁴. Impairment status was determined using a clinical ratings approach, in which performance in each cognitive domain was assigned a rating of one (above average) to nine (severe impairment); a score of five was indicative of mild NP impairment (see Woods et al., 2004). This approach conforms to the Frascati criteria for diagnosing HAND¹, and has been used in prior publications examining cognitive functioning in the CHARTER cohort^{2, 5}. Based on the gold standard, 51% of the entire sample was classified as NP impaired (Figure 1).

HIV Dementia Scale (HDS)

The HDS was administered and scored by trained clinicians or psychometrists. The four HDS subtests include antisaccadic errors (attention; range 0-4), timed written alphabet (psychomotor speed; range 0-6), recall of four items (memory; range 0-4), and timed cube copy (construction; range 0-2).

In the antisaccadic error task, patients were to alternately look at the examiner's nose and the index finger that was not moving alternately for 20 serial antisaccades. In the timed written alphabet task, participants were asked to write the entire alphabet as quickly as possible. The recall task was measured by giving patients four items and asking them to repeat those item names immediately, and then again in five minutes. The number of items that were recalled at a five-minute delay was scored. For the cube copy task, patients were asked to draw a 3-dimensional cube as precisely and quickly as possible⁶.

The raw HDS total score was the sum of the subtest scores (range 0-16). For the normative HDS score, the total HDS score was converted to a scaled score, to which we then applied a formula to obtain the age- and education-corrected T-score¹⁶. As with the NP scores, HDS T-scores < 40 were considered impaired^{14, 25}.

Data Analysis

For aim 1, we evaluated the predictive accuracy of the HDS for NP impairment in the entire study cohort using raw scores and T-scores. Receiver operating characteristic (ROC) curves were generated for both HDS scoring metrics, which allowed us to directly evaluate the comparability of their classification accuracy rates as measured by the area under the curve (AUC). Using the raw scores, participants were initially classified as NP impaired if the score was less or equal to 10⁶. The classification accuracy statistics including sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and odds ratio (OR) were calculated for both methods. McNemar's test was then conducted to determine if sensitivity and specificity of norms were significantly better than those of the original raw cutpoint.

To investigate the second hypothesis that impairment classification accuracy would improve by selecting different cutpoints, participant data were randomly split into two sets; a training subgroup (n= 801) and a validation subgroup (n= 779). The training subgroup was used to fit ROC curves to determine raw- and T-score cutpoints with the best balance between sensitivity and specificity. These cutpoints then were tested on the validation subgroup. McNemar's test was again conducted to determine whether sensitivity and specificity of norms and a new raw cutpoint significantly differed. Finally, multivariable logistic regression analyses were applied to determine the degree to which the HDS subtests contributed to the classification of NP impairment.

Results

HDS performance

The total average HDS raw score was 13.3 ± 2.9 . Twenty-seven percent (n= 431) of the entire cohort obtained full credit (16 points); on the gold standard NP battery, 68% of this group was NP normal, 17% had mild, 10% had mild to moderate, and 5% had moderate to severe impairment. Based on the HDS raw cutpoint of 10, only 17% (n= 262) of participants were classified as cognitively impaired (Figure 1). Participants' average score for attention was 3.5 ± 1.1 , written alphabet was 4.7 ± 1.8 , recall was 3.4 ± 0.9 , and construction was 1.7 ± 0.7 . Using T-scores, the group average was 39.7 ± 14.4 . Fifty-six

percent ($n = 887$) of participants were classified as impaired on the HDS with the standard cutoff of $T < 40$ (Figure 1).

The raw HDS total score was significantly associated with age ($\rho = -0.22$, $p < .0001$), education ($\rho = 0.28$, $p < .0001$), and ethnicity ($R^2 = .04$, $p < .0001$), but not with sex ($R^2 = .0001$, $p = .83$). A follow-up linear regression model showed that age, education, and ethnicity were independently associated with the HDS raw score (adjusted $R^2 = 0.15$, $p < .0001$). After applying the age and education-corrected norms, the associations between individual demographic factors and HDS performance using T scores were reduced, but remained statistically significant for age ($\rho = -0.10$, $p < .0001$), education ($\rho = -0.14$, $p < .0001$), and ethnicity ($R^2 = .03$, $p < .0001$), but not sex ($R^2 < .0001$, $p = .84$) in this very large sample. A follow-up linear regression model including age, education and ethnicity demonstrated that all three variables were independent predictors of the HDS T-scores (adjusted $R^2 = 0.06$, $p < .0001$).

Given prior evidence of a possible additive effect of HCV co-infection on HAND²⁶⁻²⁹, we examined the association between HCV and the HDS in the CHARTER cohort. HCV seropositivity was associated with significantly lower HDS raw scores ($R^2 = .009$, $p = .0003$), but not when using norms ($R^2 = .0003$, $p = .44$). Given that the HCV seropositive participants were significantly older and had obtained fewer years of formal education than HCV seronegatives ($ps < .0001$), the HDS T-score adjustments for age and education likely dampened the observed raw score HCV effect.

Classification accuracy

ROC analysis revealed that AUCs for T-scores and raw scores were both 0.68 (SE= 0.013 for both) and the overall accuracy for raw- and T-scores did not differ significantly ($z = 0.23$, $p = 0.82$). In order to confirm the commonly-used T-score cutpoint of < 40 maximized the sensitivity/specificity trade-off, a ROC curve was conducted, and it confirmed that the cutpoint of < 40 was indeed optimal (see Table 3). Accuracy statistics were then calculated for the raw total score and T-score (Figure 2). Sensitivity and specificity of the standard raw cutpoint (< 10) was 24% and 92%. Using the norms significantly improved sensitivity to 69% (McNemar's $\chi^2 = 355$, $p < .001$), with a concomitant decline in specificity to 56% (McNemar's $\chi^2 = 260$, $p < .001$). The PPV declined from 75% to 62% and NPV improved from 54% to 64%. Overall accuracy for the raw cutpoint was 57% while the T-score was 63%.

We repeated the analyses excluding those individuals with significant and contributing confounding conditions (e.g., severe brain injuries, epilepsy) that might affect CNS function. Within this subgroup of individuals with minimal confounding only ($n = 866$), sensitivity was 61%, specificity was 62%, and overall accuracy was 62%.

To determine whether the HDS is less sensitive to mild HAND compared to severe HAND, sensitivity using norms was calculated for groups with mild (clinical rating= 5, $n = 314$), mild-moderate (clinical rating 6, $n = 301$), and moderate-severe (clinical ratings 7-9, $n = 192$) NP impairment. The moderate-severe group showed the highest sensitivity of 77%, which was significantly higher than mild-moderate (65%) ($\chi^2 = 7.96$, $p = 0.005$) and mild (63%) ($\chi^2 = 10.8$, $p = 0.001$) groups. Sensitivity between the mild-moderate and mild groups did not differ ($\chi^2 = 0.28$, $p = 0.59$).

Within the subgroup of individuals who were successfully virologically suppressed on cART ($n = 605$), sensitivity was 66%, specificity was 55%, and overall accuracy was 61%.

Since correlations between ethnicity and HDS performance were found, we additionally examined classification accuracy for the Caucasian vs. African American groups. Using a raw cutpoint of < 14, there was a significantly higher false positive rate in African Americans (54%) than Caucasians (22%) ($p < 0.001$), and this held true after applying the age and education-corrected norms (50% vs. 32%, $p < 0.001$).

Selection of optimal cutpoints

Within a randomly selected training subgroup ($n = 801$), a raw score of 14 was the best cutpoint with the best balance between sensitivity and specificity, and yielded sensitivity (66%) and specificity (61%), which was confirmed in the validation subgroup (sensitivity = 66% and specificity = 60%) (see Table 3). Overall PPV, NPV, and OR were 64%, 63%, and 3.0, respectively. The sensitivity of the T-score (69%) was statistically significantly better than the raw cutpoint (66%; $p = 0.03$), and the specificity of the new raw cutpoint was statistically significantly better than the T-score ($p = 0.01$). Overall accuracy of the new raw cutpoint (< 14) and T-score (< 40) were both 63%.

Contributions of HDS subtests to overall NP impairment

Logistic regression analyses were executed to determine which HDS subtests were the most strongly associated with NP impairment. Attention (OR = 1.47/point, $p < 0.001$), speed (OR = 1.22/point, $p < 0.001$), and recall (OR = 1.30/point, $p < 0.001$) subtests were significant predictors while construction subtest was not significant (OR = 1.12 /point, $p = 0.17$) (Overall model: Likelihood χ^2 (4, $N = 1580$) = 159.3, $p < .0001$). Although age, education, and ethnicity were controlled for in the gold standard battery, they were not adjusted for in each HDS subtest raw score. Therefore, we included these three variables in the logistic regression model to evaluate whether these demographic factors would impact the association between HDS subtest scores and NP status. All four tests were significantly associated with the NP gold standard in this model (attention: OR = 1.54, $p < 0.001$, speed: OR = 1.28, $p < 0.001$, recall: OR = 1.39, $p < 0.001$, and construction: OR = 1.20, $p = 0.04$) (Overall model: Likelihood χ^2 (9, $N = 1580$) = 226.9, $p < .0001$).

Discussion

HIV-related neurocognitive impairment persists in the cART-era⁵, and there remains a need for a sensitive screening measure that can detect HAND, especially its milder forms. The HDS was specifically designed to detect HAND; however, the present study found that, within a large cohort, the originally-proposed raw cutpoint (< 10) resulted in very poor sensitivity (24%), such that 76% of individuals with HAND were inaccurately classified as neurocognitively intact. Furthermore, this study demonstrated that the HDS was most sensitive to moderate-to-severe HAND (clinical rating = 7-9) and least sensitive to mild HAND (clinical rating = 5). This suggests that individuals with especially mild HAND are likely missed and therefore would not be referred for additional evaluation or consideration of other interventions. This corroborates more recent studies demonstrating poor sensitivity of the HDS in various HIV cohorts^{7, 11, 16, 30}.

Additionally, we evaluated individuals with minimal comorbidities and found that sensitivity and specificity were both in low 60s. Within the subgroup of individuals who were successfully virologically suppressed on cART, sensitivity was slightly improved (66%) with compromised specificity (55%). Thus, although sensitivity and specificity values were somewhat variable across study subgroups, they nevertheless were broadly in the same range, which indicates that the HDS may be a suboptimal screener for HAND.

In order to improve diagnostic utility, we sought to determine a raw cutpoint that would maximize the balance of sensitivity and specificity. We found that a raw cutpoint of 14 yielded improved sensitivity (66%), while providing comparable specificity (61%), with a modest increase in overall accuracy (from 57% to 63%), as compared to the original raw cutpoint. This finding is consistent with a recent study by Simioni et al.¹³, although they found that a raw cutpoint of 14 demonstrated much better sensitivity (83%), with similar specificity (63%). As compared to the Simioni study, our study did not exclude anyone based on neuromedical or psychiatric confounds for generalizability reasons, so our sample was more heterogeneous. In addition, the NP battery in the Simioni study was relatively brief, and in order for participants to be identified as “impaired”, it is possible that they had more severe cognitive dysfunction than was captured by the more comprehensive battery in this study. In fact, the overall impairment rate reported in the Simioni study was 74%, whereas it was 51% in this study. Our prevalence rates, which are based on a multi-site US cohort, appear to more closely approximate current neuroepidemiological estimates in developed countries^{2, 31}. It is important to note that the impairment rates generally influence PPV and NPV and if the impairment rates and/or disease severity are high, predictive power of the test is improved. The higher rate of NP impairment in the Simoni study could inflate positive predictive power.

While a higher raw cutpoint improved the sensitivity to impairment, and ultimately yielded classification accuracy similar to T-scores, there is the risk of biasing the likelihood of classifying a person as impaired based upon demographic characteristics, and not acquired brain dysfunction. The current study showed that HDS raw scores were significantly associated with age and education and, to a lesser degree, ethnicity. Applying the published norms controlling for age and education¹⁶, improved sensitivity and overall accuracy, as compared to the original cutpoint. It is important to note that Morgan et al’s norms do not control for ethnicity because they did not find associations between HDS performance and ethnicity. In their study, 73% of the normative sample were Caucasian, whereas in the present study only 40% of participants were Caucasian and approximately 50% of the impaired subsample were African American. We found a weak, but significant, relationship between ethnicity and HDS performance. Of note, African Americans in our study had a significantly higher false positive rate using the HDS, suggesting that future norms should incorporate ethnicity adjustments, when possible.

While the higher raw cutpoint of 14 produced comparable classification rates for HAND relative to the demographically-adjusted T-scores, the use of this alternate raw score in clinic and research settings also carries with it a few psychometric concerns. A primary issue is the possible influence of ceiling effects, since the new cutpoint is only 2 raw scores removed from the HDS maximum of 16. This may be especially problematic upon repeated testing, since some individuals may show practice effects^{32, 33} on the HDS over multiple visits. As such, this raw cutpoint may be best suited for baseline evaluations, particularly for the clinical context when norms may be unavailable. Additionally, the raw cutpoint of 14 was empirically derived based on the balance between sensitivity and specificity in this study; however, clinicians may choose other cutscores to maximize sensitivity or specificity, depending on their particular screening needs.

We found that not all HDS subtests were equally useful in identifying NP impairment. The raw score from construction subtest was not a significant predictor of the NP impairment; however, it showed increased utility when demographic factors were included in the overall regression model. Individually, the subtests did not exhibit enough range in values to make demographic adjustments tenable. Since this leaves one with only one score (the total) to modify (increasing or decreasing a single cutpoint), there are inherent limitations in the degree to which the overall classification accuracy can be affected by applying norms.

The current study suggests that when identifying optimal cutpoints using either demographic adjustments or an increased raw cutpoint, the sensitivity of the HDS does not exceed 70%. This very brief screen has inherent limitations. For example, the range of subtest and summary scores are small; therefore, differentiation across a range of cognitive functioning is difficult. All four tests are relatively easy, which resulted in a ceiling effect: 27% of the entire sample, including 32% of those who were NP impaired, obtained a perfect score. Also, the HDS does not target the most vulnerable cognitive domains in HIV population, such as learning and executive functions. Other brief, well-validated measures (e.g., action fluency) might be considered for assessing these domains. Nevertheless pending the development of more sensitive measures, the HDS can still provide clinicians with useful information regarding the cognitive status of their patients, and offers an opportunity to discuss the issue of HAND.

The current study has some limitations. First, there were no formal exclusion criteria, in order to emphasize generalizability; as a result, many of our participants had comorbidities that might affect CNS functioning. Sensitivity and specificity slightly changed after excluding those who had severe confounds. However, clinicians may well want to be aware of cognitive impairments regardless of the etiology, and the source of dysfunction could be discerned during a more comprehensive evaluation. Secondly, a higher rate of detectable viral load compared to clinical trials was found in this study. It is, however, similar to other community samples³⁴⁻³⁶. Also, HIV-infected persons with comorbidities (e.g., substance use disorders), like those prevalent in the CHARTER cohort, are at greater risk for ART non-adherence and viremia³⁷⁻⁴⁰. Thirdly, the majority of our participants were male, which reflects the epidemiological trends of HIV in the U.S. Although sex was not a significant predictor of HDS performance, there are measures for which sex is related to outcomes (e.g., verbal learning). The results in the current study may not directly apply to HIV populations in other countries, especially when there are sex-related social, financial, and healthcare advantages or disadvantages for women^{41, 42}.

In summary, the HDS is a statistically significant predictor of HAND, especially when applying norms or an increased raw cutpoint. However, its relatively low diagnostic classification accuracy and potential ceiling effects hinder its clinical utility.

Acknowledgments

The CNS HIV Anti-Retroviral Therapy Effects Research (CHARTER) is supported by awards N01 MH22005, HHSN271201000036C, and HHSN271201000030C from the National Institutes of Health.

Source of Funding: The authors declare that they have no conflict of interest. The CNS HIV Anti-Retroviral Therapy Effects Research (CHARTER) is supported by awards N01 MH22005, HHSN271201000036C, and HHSN271201000030C from the National Institutes of Health.

References

1. Antinori A, Arendt G, Becker JT, et al. Updated research nosology for HIV-associated neurocognitive disorders. *Neurology*. Oct 30; 2007 69(18):1789–1799. [PubMed: 17914061]
2. Heaton RK, Clifford DB, Franklin DR Jr. et al. HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study. *Neurology*. Dec 7; 2010 75(23):2087–2096. [PubMed: 21135382]
3. McArthur JC. HIV dementia: an evolving disease. *J Neuroimmunol*. Dec; 2004 157(1-2):3–10. [PubMed: 15579274]
4. McArthur JC, Brew BJ. HIV-associated neurocognitive disorders: is there a hidden epidemic? *AIDS*. Jun 1; 2010 24(9):1367–1370. [PubMed: 20559041]

5. Heaton RK, Franklin DR, Ellis RJ, et al. HIV-associated neurocognitive disorders before and during the era of combination antiretroviral therapy: differences in rates, nature, and predictors. *J Neurovirol.* Feb; 2011 17(1):3–16. [PubMed: 21174240]
6. Power C, Selnes OA, Grim JA, McArthur JC. HIV Dementia Scale: a rapid screening test. *J Acquir Immune Defic Syndr Hum Retrovirol.* Mar 1; 1995 8(3):273–278. [PubMed: 7859139]
7. Berghuis JP, Uldall KK, Lalonde B. Validity of two scales in identifying HIV-associated dementia. *J Acquir Immune Defic Syndr.* Jun 1; 1999 21(2):134–140. [PubMed: 10360805]
8. Bottiggi KA, Chang JJ, Schmitt FA, et al. The HIV Dementia Scale: predictive power in mild dementia and HAART. *J Neurol Sci.* Sep 15; 2007 260(1-2):11–15. [PubMed: 17482212]
9. Dougherty RH, Skolasky RL Jr, McArthur JC. Progression of HIV-associated dementia treated with HAART. *AIDS Read.* Feb; 2002 12(2):69–74. [PubMed: 11905143]
10. Davis HF, Skolasky RL Jr, Selnes OA, Burgess DM, McArthur JC. Assessing HIV-associated dementia: modified HIV dementia scale versus the Grooved Pegboard. *AIDS Read.* Jan; 2002 12(1):29–31. 38. [PubMed: 11862658]
11. Smith CA, van Gorp WG, Ryan ER, Ferrando SJ, Rabkin J. Screening subtle HIV-related cognitive dysfunction: the clinical utility of the HIV dementia scale. *J Acquir Immune Defic Syndr.* May 1; 2003 33(1):116–118. [PubMed: 12792365]
12. Skinner S, Adewale AJ, DeBlock L, Gill MJ, Power C. Neurocognitive screening tools in HIV/AIDS: comparative performance among patients exposed to antiretroviral therapy. *HIV Med.* Apr; 2009 10(4):246–252. [PubMed: 19187172]
13. Simioni S, Cavassini M, Annoni JM, et al. Cognitive dysfunction in HIV patients despite long-standing suppression of viremia. *AIDS.* Jun 1; 2010 24(9):1243–1250. [PubMed: 19996937]
14. Heaton, RK.; Miller, SW.; Taylor, MJ.; Grant, I. Revised Comprehensive Norms for an Expanded Halstead-Reitan Battery: Demographically Adjusted Neuropsychological Norms for African American and Caucasian Adults Scoring Program. Psychological Assessment Resources, Inc.; Lutz, FL: 2004.
15. von Giesen HJ, Haslinger BA, Rohe S, Koller H, Arendt G. HIV Dementia Scale and psychomotor slowing--the best methods in screening for neuro-AIDS. *J Neuropsychiatry Clin Neurosci.* Spring; 2005 17(2):185–191. [PubMed: 15939972]
16. Morgan EE, Woods SP, Scott JC, et al. Predictive validity of demographically adjusted normative standards for the HIV Dementia Scale. *J Clin Exp Neuropsychol.* Jan; 2008 30(1):83–90. [PubMed: 17852582]
17. Benton, AL.; Hanmsher, KS.; Sivan, AB. Multilingual Aphasia Examination. 3rd ed. AJA Associates, Inc.; Iowa City: 1994.
18. Gladsjo JA, Schuman CC, Evans JD, Peavy GM, Miller SW, Heaton RK. Norms for letter and category fluency: demographic corrections for age, education, and ethnicity. *Assessment.* Jun; 1999 6(2):147–178. [PubMed: 10335019]
19. Strauss, E.; Sherman, EMS.; Spreen, O. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. 3rd ed. Oxford University Press; New York: 2006.
20. Psychological Corporation. Wechsler Adult Intelligence Scale. 3rd ed. Psychological Corporation; San Antonio, TX: 1997.
21. Reitan, R. Manual for Administration of Neuropsychological Test Batteries for Adults and Children. Neuropsychology Laboratory; Tuscon: 1979.
22. Kongs, SK.; Thompson, LL.; Iverson, GL.; Heaton, RK. Wisconsin Card Sorting Test - 64 card computerized version. Psychological Assessment Resources; Odessa, FL: 2000.
23. Diehr MC, Cherner M, Wolfson TJ, Miller SW, Grant I, Heaton RK. The 50 and 100-item short forms of the Paced Auditory Serial Addition Task (PASAT): demographically corrected norms and comparisons with the full PASAT in normal and clinical samples. *J Clin Exp Neuropsychol.* Jun; 2003 25(4):571–585. [PubMed: 12911108]
24. Kløve, H. Grooved pegboard. Lafayette Instruments; Lafayette, IN: 1963.
25. Woods SP, Rippeth JD, Frol AB, et al. Interrater reliability of clinical ratings and neurocognitive diagnoses in HIV. *J Clin Exp Neuropsychol.* Sep; 2004 26(6):759–778. [PubMed: 15370374]

26. Martin EM, Novak RM, Fendrich M, et al. Stroop performance in drug users classified by HIV and hepatitis C virus serostatus. *J Int Neuropsychol Soc.* Mar; 2004 10(2):298–300. [PubMed: 15012850]
27. von Giesen HJ, Heintges T, Abbasi-Boroudjeni N, et al. Psychomotor slowing in hepatitis C and HIV infection. *J Acquir Immune Defic Syndr.* Feb 1; 2004 35(2):131–137. [PubMed: 14722444]
28. Clifford DB, Evans SR, Yang Y, Gulick RM. The neuropsychological and neurological impact of hepatitis C virus co-infection in HIV-infected subjects. *AIDS.* Oct; 2005 19(Suppl 3):S64–71. [PubMed: 16251830]
29. Parsons TD, Tucker KA, Hall CD, et al. Neurocognitive functioning and HAART in HIV and hepatitis C virus co-infection. *AIDS.* Aug 1; 2006 20(12):1591–1595. [PubMed: 16868439]
30. Carey CL, Woods SP, Rippeth JD, et al. Initial validation of a screening battery for the detection of HIV-associated cognitive impairment. *Clin Neuropsychol.* May; 2004 18(2):234–248. [PubMed: 15587671]
31. Robertson KR, Su Z, Margolis DM, et al. Neurocognitive effects of treatment interruption in stable HIV-positive patients in an observational cohort. *Neurology.* Apr 20; 2010 74(16):1260–1266. [PubMed: 20237308]
32. Dikmen SS, Heaton RK, Grant I, Temkin NR. Test-retest reliability and practice effects of expanded Halstead-Reitan Neuropsychological Test Battery. *J Int Neuropsychol Soc.* May; 1999 5(4):346–356. [PubMed: 10349297]
33. Beglinger LJ, Gaydos B, Tangphao-Daniels O, et al. Practice effects and the use of alternate forms in serial neuropsychological testing. *Arch Clin Neuropsychol.* Jun; 2005 20(4):517–529. [PubMed: 15896564]
34. Palella FJ, Chmiel JS, Moorman AC, Holmberg SD, Investigators HOS. Durability and predictors of success of highly active antiretroviral therapy for ambulatory HIV-infected patients. *AIDS.* Aug 16; 2002 16(12):1617–1626. [PubMed: 12172083]
35. Paredes R, Mocroft A, Kirk O, et al. Predictors of virological success and ensuing failure in HIV-positive patients starting highly active antiretroviral therapy in Europe - Results from the EuroSIDA Study. *Archives of Internal Medicine.* Apr 24; 2000 160(8):1123–1132. [PubMed: 10789605]
36. Yehia BR, Fleishman JA, Metlay JP, Moore RD, Gebo KA. Sustained Viral Suppression in HIV-Infected Patients Receiving Antiretroviral Therapy. *Jama-J Am Med Assoc.* Jul 25; 2012 308(4):339–342.
37. Li X, Margolick JB, Conover CS, et al. Interruption and discontinuation of highly active antiretroviral therapy in the multicenter AIDS cohort study. *J Acquir Immune Defic Syndr.* Mar 1; 2005 38(3):320–328. [PubMed: 15735452]
38. Starace F, Bartoli L, Aloisi MS, et al. Cognitive and affective disorders associated to HIV infection in the HAART era: findings from the NeuroICONA study. Cognitive impairment and depression in HIV/AIDS. The NeuroICONA study. *Acta Psychiatr Scand.* Jul; 2002 106(1):20–26. [PubMed: 12100344]
39. Moore DJ, Blackstone K, Woods SP, et al. Methamphetamine use and neuropsychiatric factors are associated with antiretroviral non-adherence. *AIDS Care.* Apr 24; 2012
40. Carrico AW. Substance use and HIV disease progression in the HAART era: implications for the primary prevention of HIV. *Life Sci.* May 23; 2011 88(21-22):940–947. [PubMed: 20934437]
41. Hestad KA, Menon JA, Silalukey-Ngoma M, et al. Sex differences in neuropsychological performance as an effect of human immunodeficiency virus infection: a pilot study in Zambia, Africa. *J Nerv Ment Dis.* Apr; 2012 200(4):336–342. [PubMed: 22456588]
42. Gupta S, Vaida F, Riggs K, et al. Neuropsychological performance in mainland china: the effect of urban/rural residence and self-reported daily academic skill use. *J Int Neuropsychol Soc.* Jan; 2011 17(1):163–173. [PubMed: 21083967]

Figure 1. Distributions of NP status based on “Gold Standard” NP battery, HDS raw score and HDS T-score

Figure 2.

Classification accuracy (%): Raw cutpoints vs. T-score cutpoint

Note: PPV = positive predictive value, NPV = negative predictive value

Sens.: 10 vs. 14 ($p < .001$), 10 vs. T-scr ($p < .001$), 14 vs. T-scr ($p = .03$)

Spec.: 10 vs. 14 ($p < .001$), 10 vs. T-scr ($p < .001$), 14 vs. T-scr ($p = .01$) (*McNemar's tests)

Table 1

The CHARTER research sites

Location	Sample (n)
Johns Hopkins University (Baltimore, MD)	231
University of Washington (Seattle, WA)	262
Washington University (St. Louis, MO)	268
University of Texas Medical Branch (Galveston, TX)	261
Mt. Sinai School of Medicine (New York, NY)	270
University of California, San Diego (San Diego, CA)	288

Table 2

Demographic, psychiatric, and medical characteristics of the CHARTER participants

Variable	Total (N= 1580)	NP normal (n= 771)	NP impaired (n= 809)
Age (years)	43.1 (8.5)	42.6 (8.8)	43.4 (8.2)
Education (years)	12.5 (2.6)	12.6 (2.6)	12.5 (2.5)
Sex (Male)	1215 (77%)	609 (79%)	606 (75%)
Ethnicity	Caucasian 629 (40%) African American 762 (48%) Other 189 (12%)	Caucasian 304 (40%) African American 398 (51%) Other 69 (9%)	Caucasian 325 (40%) African American 364 (45%) Other 120 (15%)
AIDS diagnosis	978 (62%)	444 (58%)	534 (66%)
Estimated duration of infection (months)	118.5 (76.8)	117 (77.6)	119.9 (76.1)
Detectable HIV RNA plasma	928 (60%)	458 (61%)	470 (58%)
Nadir CD4 (cells/l) *	179 (IQR: 50, 307)	198 (IQR: 62, 347)	159 (IQR: 40, 288)
Current CD4 (cells/l) *	423 (IQR: 263, 607)	431 (IQR: 270, 610)	409 (IQR: 251, 607)
ART status (on)	1102 (70%)	503 (65%)	599 (74%)
Duration of current ART (months) *	11 (IQR: 4, 27)	10 (IQR: 3, 26)	12 (IQR: 4, 28)
HCV positive	404 (30%)	195 (26%)	209 (26%)
Life time depression	802 (51%)	401 (52%)	401 (50%)
Substance use disorders (last 12 months)	598 (38%)	273 (34%)	325 (42%)
Life time substance use disorders	1151 (74%)	583 (77%)	568 (70%)

* nadir CD4 count, current CD4 count, and current duration of ART are medians with interquartile ranges (IQR)

Table 3

Cutpoints for HDS summary score and T-score

		Sensitivity	Specificity	PPV	NPV	Accuracy	Odds ratio
Raw Scores	10	.24	.92	.75	.54	.57	3.4
	11	.33	.88	.73	.55	.60	3.4
	12	.41	.81	.69	.57	.61	2.9
	13	.53	.74	.68	.60	.63	3.2
	14	.66	.61	.64	.63	.63	3.0
	15	.80	.43	.59	.67	.62	2.9
	16	1.0	0.0	.51	N/A	.51	N/A
	35	.56	.70	.66	.60	.63	3.0
	36	.58	.67	.65	.60	.62	2.8
	37	.61	.63	.63	.61	.62	2.7
T-scores	38	.64	.61	.63	.62	.63	2.7
	39	.67	.58	.63	.63	.63	2.8
	40	.69	.56	.62	.64	.63	2.9
	41	.72	.53	.61	.64	.63	2.8
	42	.74	.51	.61	.65	.63	2.9
	43	.77	.48	.61	.66	.63	3.0
	44	.78	.46	.60	.67	.62	3.0
	45	.79	.44	.60	.67	.62	3.0