
David B. Clifford
Washington University School of Medicine in St. Louis
et al

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs

Recommended Citation
http://digitalcommons.wustl.edu/open_access_pubs/3295

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact engeszer@wustl.edu.
Diagnosing Symptomatic HIV-Associated Neurocognitive Disorders: Self-Report Versus Performance-Based Assessment of Everyday Functioning

1San Diego State University/University of California, San Diego Joint Doctoral Program in Clinical Psychology; San Diego, California, USA
2Department of Psychiatry, University of California, San Diego; San Diego, California, USA
3Department of Neurology, Washington University, St. Louis; St. Louis, Missouri, USA
4Department of Neurology, Mount Sinai School of Medicine; New York, New York, USA
5Departments of Neurology and Medicine (Infectious Diseases), University of Washington; Seattle, WA, USA
6Department of Pathology, University of Texas Medical Branch; Galveston, Texas, USA
7Department of Neurology, Johns Hopkins University; Baltimore, Maryland, USA
8Department of Neurosciences, University of California, San Diego; San Diego, California, USA
(RECEIVED May 17, 2011; FINAL REVISION September 12, 2011; ACCEPTED September 13, 2011)

Abstract

Three types of HIV-associated neurocognitive disorders (HAND) exist that are distinguished by presence and severity of impairment in cognitive and everyday functioning. Although well-validated neurocognitive measures exist, determining impairment in everyday functioning remains a challenge. We aim to determine whether Self-Report measures of everyday functioning are as effective in characterizing HAND as Performance-Based measures. We assessed 674 HIV-infected participants with a comprehensive neurocognitive battery; 233 met criteria for a HAND diagnosis by having at least mild neurocognitive impairment. Functional decline was measured via Self-Report and Performance-Based measures. HAND diagnoses were determined according to published criteria using three approaches to assess functional decline: (1) Self-Report measures only, (2) Performance-Based measures only, and (3) Dual-method combining Self-Report and Performance-Based measures. The Dual-method classified the most symptomatic HAND, compared to either singular method. Singular method classifications were 76% concordant with each other. Participants classified as Performance-Based functionally impaired were more likely to be unemployed and more immunosuppressed, whereas those classified as Self-Report functionally impaired had more depressive symptoms. Multimodal methods of assessing everyday functioning facilitate detection of symptomatic HAND. Singular Performance-Based classifications were associated with objective functional and disease-related factors; reliance on Self-Report classifications may be biased by depressive symptoms. (JINS, 2012, 18, 79–88)

Keywords: HIV/AIDS, Cognition disorders, Activities of daily living, Infectious disease, Self assessments, Employment

INTRODUCTION

Despite advances in antiretroviral therapy (ART), HIV-associated neurocognitive disorders (HAND) are still observed in close to half of the HIV-infected (HIV1) population (Heaton et al., 2010). In the context of assigning HAND diagnoses for HIV1 persons, guidelines were recently updated to include additional operationalization of the diagnostic criteria for daily functioning decline (Antinori et al., 2007). More specifically, the current nomenclature for diagnosing asymptomatic HAND (i.e., Asymptomatic Neurocognitive Impairment; ANI) requires neurocognitive deficits in at least two ability domains that are attributable to HIV-infection, but do not meaningfully influence daily functioning (Antinori et al., 2007). In contrast, “symptomatic” HAND diagnoses require significant HIV-associated neurocognitive deficits that interfere with functional capabilities at either a mild (i.e., Mild Neurocognitive Disorder, MND) or moderate-to-severe (i.e., HIV-associated Dementia, HAD) level. It is widely held that the most prevalent form of HAND is “asymptomatic,” or ANI (33% of the HIV+ population; Heaton et al., 2010), meaning that the observed neurocognitive impairment does not appear to affect daily functioning.
However, research reliably demonstrates that even mild HIV-associated neurocognitive deficits are significantly associated with impaired functional outcomes, ranging from poorer antiretroviral medication adherence (Albert et al., 1995; Benedict, Mezhir, Walsh, & Hewitt, 2000; Hinkin et al., 2002) and employment status (Benedict et al., 2000; Heaton et al., 1996) to general quality of life (Benedict et al., 2000) and even mortality (Mapou et al., 1993). Such data arguably run counter to the presumed predominance of “asymptomatic” HAND in the cognitively-impaired HIV + population and raise questions about the typical, self-report methods for assessing asymptomatic versus symptomatic status. Identifying the functional assessments that best distinguish between asymptomatic and symptomatic HAND is critical for understanding the etiology and progression of HAND in the neuroAIDS context.

Methodologically, well-validated neurocognitive assessments exist to adequately establish the presence or absence of neurocognitive impairment. The greater challenge for clinicians and researchers is in the determination of impairment in everyday functioning. Self-report measures of daily functioning have several advantages, including low cost, minimal participant burden, and high face validity (Simoni et al., 2006; Wagner & Miller, 2004). However, self-report is susceptible to social desirability and recall inaccuracies or bias, which may overestimate ability (Chesney et al., 2000; Thames, Kim, et al., 2010). For example, a recent study that examined self-report versus electronic medication monitoring of antiretroviral medications found that self-report significantly overestimated adherence rates (self-report up to 90% adherent vs. electronic monitoring 70% adherent; Lu et al., 2008). Additionally, self-report measures are susceptible to overestimation biases due to depressed mood; particularly in HIV+, studies have shown that depressive symptoms, not objective neuropsychological performance, accounts for a majority of the variance in cognitive and functional complaints (Rourke, Halman, & Bassel, 1999b; Thames, Becker, et al., 2010). Performance-based measures of daily functioning, on the other hand, can be time-intensive and require additional training and tools to administer (Moore, Palmer, Patterson, & Jeste, 2007), and because of their standardization may not capture differences in requirements of individual patients’ daily tasks and activities. Yet, performance-based measures used with HIV-infected patients have been shown to be objective and reliable in predicting “real life” outcomes such as employment status as well as medication and financial management (Heaton et al., 2004; Thames, Kim, et al., 2010).

Previous studies have established the importance of using multiple assessment methods (e.g., self, informant, performance, behavioral observation) to maximize sensitivity (Hunsley & Meyer, 2003; Meyer et al., 2001; Schwartz, Kozora, & Zeng, 1996) and increase the quality and usefulness of diagnostic information; yet, this approach is frequently applied in the context of HIV. Despite this, the current guidelines for assigning HAND diagnoses do not require both performance-based and self-report measures in determining level of daily functioning (Antinori et al., 2007). In practice, clinicians and researchers often rely exclusively on self-report measures due to their convenience (e.g., Woods et al., 2004). Examining the utility of performance-based measures in assessing functional status is important as it may add unique diagnostic information that is currently missing when only the self-report approach is applied.

Our study, therefore, aims to compare three functional assessment approaches to determine HAND: (1) Self-Report measures only; (2) Performance-Based measures only; and (3) a Dual method that combines self-report plus performance-based measures. Specifically, our goal was to assess the incremental validity of the Performance-Based approach, beyond the common practice of self-report, in determining functional status. In particular, we aim to examine how asymptomatic HAND classifications change with the addition of Performance-Based measures. We use the Dual (i.e., Self-Report plus Performance-Based) classifications as the benchmark to examine the diagnostic sensitivity and specificity of Performance-Based and Self-Report alone. We hypothesized that the Dual method would detect more symptomatic HAND (i.e., MND and HAD) as compared to the other approaches. Additionally, we expected that the Performance-Based approach would be less susceptible to depression-related reporting bias, and more strongly associated with objective indicators of disease status and functional outcomes.

METHODS

Participants
Six hundred seventy-four HIV-infected (HIV+) participants were drawn from the CNS HIV Antiretroviral Therapy Effects Research (CHARTER) study, a prospective cohort study conducted in HIV clinics at six academic centers: Johns Hopkins University (Baltimore, MD); Mt. Sinai School of Medicine (New York, NY); University of California at San Diego (San Diego, CA); University of Washington (Seattle, WA); University of Texas Medical Branch (Galveston, TX); Washington University (St. Louis, MO; see Heaton et al., 2010). All participants were HIV-infected and were excluded only if they could not complete the assessment at the time of evaluation. As the purpose of this study was to examine HAND diagnoses, only those individuals with neurocognitive impairment at entry were considered for inclusion in the analyses (44% of the cohort; 299/674). Of these, most (45%, 135/299) were classified as having none or minimal comorbidities (non-HIV-related factors that could affect cognition and functioning) and 33% (98/299) with moderate comorbidities. Those with severe comorbidities (22%, 66/299) that precluded a HAND diagnosis were excluded (see Figure 1). The demographic, psychiatric, and HIV disease and treatment characteristics of the study participants are summarized in Table 1.
Procedures

Standard approvals and participant consents

The Human Subjects Protection Committees of each participating institution approved the study procedures. Written informed consent was obtained from all study participants.

Laboratory assessment

HIV infection was diagnosed by enzyme linked immunosorbent assay with Western blot confirmation. Routine clinical chemistry panels, complete blood counts, rapid plasma reagin, hepatitis C virus antibody, and CD4+ T cells (flow cytometry) were performed at each site’s Clinical Laboratory Improvement Amendments (CLIA)-certified, or CLIA equivalent, medical center laboratory. HIV RNA levels were measured centrally in plasma and CSF by reverse transcriptase-polymerase chain reaction (Roche Amplicor, v. 1.5, lower limit of quantitation 50 copies/mL). AIDS was diagnosed using available clinical and immunologic data (defined as has having a CD4 Cell Count < 200 cells/µL or the presence of an AIDS-indicating clinical conditions using the CDC AIDS classification system).

Neurobehavioral examination

All participants completed a comprehensive neurocognitive test battery, covering seven ability domains commonly affected among HIV-infected persons: speed of information processing, learning, memory, verbal fluency, attention/working memory, executive functioning and motor (see Heaton et al., 2010, Table 1. Demographic and clinical characteristics (n = 233)

<table>
<thead>
<tr>
<th>Demographic variable</th>
<th>M, P, or Median</th>
<th>SD or IQR</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>45.2 ± 8.5</td>
<td>22–69</td>
<td></td>
</tr>
<tr>
<td>Education, years</td>
<td>13.3 ± 2.4</td>
<td>7–20</td>
<td></td>
</tr>
<tr>
<td>Gender (M)</td>
<td>77%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>44%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>African-American</td>
<td>39%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Hispanic</td>
<td>14%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Other</td>
<td>3%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Current CD4 (cells/µL)</td>
<td>451 ± 168–626</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Nadir CD4 (cells/µL)</td>
<td>150 ± 43–275</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>AIDS</td>
<td>64%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>% HIV CSF viral load detectable</td>
<td>25%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>% detectable if on ARVs</td>
<td>56%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>% HIV plasma viral load detectable</td>
<td>46%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>% detectable if on ARVs</td>
<td>73%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Hepatitis C virus co-infection</td>
<td>23%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Beck Depression Inventory-II</td>
<td>12.6 ± 11.1</td>
<td>0–56</td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>33%</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>LT substance abuse/dependence dx</td>
<td>67%</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

M = mean; P = percent; SD = standard deviation; IQR = Inter-quartile range; ARVs = antiretrovirals; LT substance abuse/dependence dx = DSM-IV diagnosis of lifetime substance abuse or dependence
for listing of specific tests). All participants completed three
eurocognitive testing visits (screening, baseline, and 6-month
follow-up). At the screening visit, participants were adminis-
tered a subset of the larger neuropsychological battery (i.e., Hopkins
Verbal Learning Test-Revised, WAIS-III Symbol Digit, and
Grooved Pegboard). Alternate forms of the Hopkins Verbal
Learning Test-Revised and Brief Visual Memory Test-Revised
were used at each visit. Raw test scores were converted
to demographically-corrected standard scores (T-scores).
Comprehensive normative standards were used, which correct
for effects of age, education, sex and ethnicity, as appropriate
(Heaton, Taylor, Manly & Tulsky, 2003; Heaton et al., 2004;
Norman et al., 2011). We used published practice effect
corrections to adjust for a single previous exposure to the
neurocognitive tests (Cysique et al., 2011). Specifically, the
median practice effect (in scaled score units) for the group
(individual baseline performance minus 6-month follow-up
performance) was subtracted from the scaled score at the 6-month follow-up, to estimate what the performance would
have been without practice (see Cysique, Franklin, et al., 2011,
Methods: “Computation of the practice effect”; see also further
validation of the method in a recent study of neuroAIDS in
China: Cysique, Letendre, et al., 2011). To classify presence and
severity of neuropsychological impairment, we applied a published
objective algorithm that has been shown to yield excellent
interrater reliability for HAND diagnoses in previous multisite
studies (see Woods et al., 2004). Briefly, neuropsychological
clinical ratings were assigned from demographically adjusted
T-scores for each of the seven major ability areas using a
nine-point scale (1 = above average functioning, T ≥ 55; 9 = severe impairment, T < 20) with a global rating of five or
above indicating definite mild neuropsychological impairment
(see Heaton et al., 2004 and Woods et al., 2004 for details).
Consistent with the most recent and widely accepted diagnostic
guidelines (i.e., Antinori et al., 2007), two neurocognitive
domains were required to be in the impaired range to assign an
impaired global rating (≥5).

Psychiatric examination
Psychiatric and substance abuse or dependence Diagnostic
and Statistical Manual-IV diagnoses were assessed by administering the Depression and Substance Use modules of
the computer-assisted Composite International Diagnostic Interview (CIDI; World Health Organization, 1997). Current
mood was assessed with the Beck Depression Inventory-II
(Heaton et al., 2004). Eleven items were included from this scale detailing the degree to which individuals independently function in the areas of Financial
Management, Home Repair, Medication Management, Laundry,
Transportation, Grocery Shopping, Shopping, Housekeeping
(Cleaning), Cooking, Work, and Telephone Use. For each
activity, the participant separately rates his/her current level of
independence and highest previous level of independence. The
total score is the total number of activities for which there is
currently a need for increased assistance (ranging from minimal
to complete assistance), with a range of zero (no change) to 11
(increased dependence in all activities).

Performance-based measures
Medication management was assessed via a revised version
of The Medication Management Test (MMT; Albert et al., 1999).
A full description of our modified version of the MMT
(i.e., MMT-R) can be found in Heaton et al. (2004). Briefly,
the MMT-R retains the pill-dispensing component in which
subjects must dispense a 1-day dosage of “medications” from
three standardized bottles labeled with dosing information. In
the medication inference component, there are seven questions
regarding the medications as well as an over-the-counter medica-
tion insert. The MMT-R takes approximately 20 minutes to
administer and the best score possible is 10 points.
Participants also completed standardized work samples
(MESA SF2) and the next generation COMPASS programs
(Valpar International Corporation, 1986, 1992). These
batteries consist of multi-modal, criterion-referenced instru-
ments designed to establish participant skill level in areas of
vocational functioning. The battery takes approximately 1 hr
and includes computerized subtests and noncomputerized
mechanical tasks that correspond to the Dictionary of Occupa-
tional Titles (DOT; U.S. Department of Labor, 1991) job levels.
Raw scores from these tests are converted into ability levels for
each of the DOT classifications using the commercial software
accompanying the MESA SF2 and COMPASS. A detailed
explanation of test development for the MESA and COMPASS
is beyond the scope of this study (see Valpar International
Corporation, 1986, 1992; Heaton et al., 2010).

Functional Impairment in Everyday Life
Self-report measures
The Patient’s Assessment of Own Functioning Inventory was
administered to determine perceived everyday functioning
impairments (PAOFI; Chelune, Heaton, & Lehman, 1986).
The PAOFI is a 41-item questionnaire in which the partici-
pant reports the frequency with which he/she has difficulties
with memory, language and communication, use of his/her
hands, sensory-perception, higher level cognitive and intel-
lectual functions, work, and recreation.
To assess dependence in performing instrumental activities
of daily living (IADLs), a modified version of the Lawton and
Brody scale was used (Heaton et al., 2004). Eleven items
were included from this scale detailing the degree to which
individuals independently function in the areas of Financial
Management, Home Repair, Medication Management, Laundry,
Transportation, Grocery Shopping, Shopping, Housekeeping
(Cleaning), Cooking, Work, and Telephone Use. For each
activity, the participant separately rates his/her current level of
independence and highest previous level of independence. The
total score is the total number of activities for which there is
currently a need for increased assistance (ranging from minimal
to complete assistance), with a range of zero (no change) to 11
(increased dependence in all activities).
the neuropsychologically normal cohort would be impaired at one standard deviation (cutoff scores: MMT-R < 5 and Valpar < 24) and 2% of the cohort would be impaired at two standard deviations (cutoff scores: MMT-R < 2 and Valpar < 17).

HAND classifications

For the purposes of this study, the authors created data-driven formulas to diagnose ANI, MND and HAD three distinct ways: (1) using only the Self-Report measures of daily functioning; (2) using only the Performance-Based measures of daily functioning; and (3) a Dual method that combines both the self- and performance-based measures of daily functioning. In all formulas, a neuropsychological global rating score of 5 or 6 defined mild-to-moderate cognitive impairment (i.e., minimum criteria needed for ANI or MND; see Figure 2) and a neuropsychological global rating score of 7 or above defined moderate-to-severe cognitive impairment (i.e., minimum criteria needed for HAD). Additionally, employment status that was associated with cognitive decline (i.e., participant endorsed both “I no longer am able to work” and “I feel that the difficulties I am having on the above tasks are due to: Primarily cognitive problems” or “Equally cognitive and physical problems” on the IADL) counted as one area of functional decline in all formulas in accordance with the Frascati criteria (Antinori et al., 2007). To determine functional decline in the Self-Report formula, scores on the PAOFI and IADL were examined. Specifically, impairment on the PAOFI was established according to the guidelines outlined by Woods et al. (2004) such that the presence of three or more items endorsed as “almost always,” “very often,” or “fairly often” indicated areas of functional impairment. To control for depression in self-report, previously defined criteria were used (Woods et al., 2004) in which subjects with elevated BDI scores (BDI ≥ 17) needed to exhibit a higher level of complaint on the PAOFI (PAOFI ≥ 10 complaints) to qualify for functional impairment on this measure. Scores on the IADL that show decline from “best” to “now” in two or more areas that were identified as being at least partially due to cognitive problems (vs. physical impairment) also qualified as one area of functional decline in the Self-Report formula (Woods et al., 2004).

In the Performance-Based formula, mild and major functional impairment were defined as scores one or two standard deviations below the mean, respectively, on the MMT-R and Valpar in line with the Frascati criteria. Therefore, MND was diagnosed (1) if both MMT-R and Valpar scores were one standard deviation below the mean; or (2) if one task was one standard deviation below the mean and the subject was unemployed. HAD was diagnosed: (1) if both MMT-R and Valpar scores were two standard deviations below the mean; (2) if scores on both tasks were one standard deviation below the mean and the subject was unemployed; or (3) if one task was two standard deviations below the mean and the subject was unemployed.

All diagnostic criteria for functional decline were included in the Dual diagnostic method. Measures included in each formula criterion are summarized in Table 2.

Discrepancy variable

 Discordant classifications between the Self-Report and Performance-Based methods were examined by creating a “discrepancy variable” with four levels: (1) agree: asymptomatic (i.e., ANI); (2) agree: symptomatic; (3) discrepant: functionally impaired by Self-Report only; and (4) discrepant: functionally impaired by Performance-Based only. The discrepancy variable was used to examine potential demographic, disease, psychiatric, and cognitive differences that may be associated with discordant diagnoses.

Statistical Analyses

The McNemar-Bowker nonparametric test for non-independent samples was conducted to compare HAND diagnosis frequencies...
Table 2. Criteria for measures included in each diagnostic formula

<table>
<thead>
<tr>
<th>Formula</th>
<th>NP Global Rating</th>
<th>PAOFI</th>
<th>IADL</th>
<th>MMT-R</th>
<th>Valpar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-Report</td>
<td>≥ 5 = cognitive impairment</td>
<td>≥ 3 elevated items</td>
<td>≥ 2 areas of decline</td>
<td>&lt; 5 = mild imp;</td>
<td>&lt; 24 = mild imp;</td>
</tr>
<tr>
<td>Performance-Based</td>
<td>≥ 5 = cognitive impairment</td>
<td>≥ 3 elevated items</td>
<td>≥ 2 areas of decline</td>
<td>&lt; 2 = mod imp</td>
<td>&lt; 17 = mod imp</td>
</tr>
<tr>
<td>Dual method</td>
<td>≥ 5 = cognitive impairment</td>
<td>≥ 3 elevated items</td>
<td>≥ 2 areas of decline</td>
<td>&lt; 5 = mild imp;</td>
<td>&lt; 24 = mild imp;</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt; 2 = mod imp</td>
<td>&lt; 17 = mod imp</td>
</tr>
</tbody>
</table>

NP = neuropsychological; PAOFI = Patient’s Assessment of Own Functioning Inventory; IADL = Instrumental Activities of Daily Living; MMT-R = Medication Management Test-Revised; mild imp = mild impairment (below 1 SD); mod imp = moderate impairment (below 2 SDs). Note: employment status was also included in the diagnostic formulas and additional modifications were made for individuals with significant depression, see text.

(i.e., ANI vs. MND vs. HAD) as well as “symptomatic” frequencies (i.e., MND or HAD) across each diagnostic method.

Sensitivity and specificity of the singular methods (i.e., Self-Report and Performance-Based) compared to the Dual method were calculated for each diagnostic level (i.e., ANI vs. MND+HAD; ANI+MND vs. HAD; and ANI vs. MND). \( \chi^2 \) analyses were conducted to compare the sensitivity and specificity of the Self-Report versus Performance-Based methods across each specified diagnostic level.

The discrepancy variable (described above) was explored by screening which disease and functional variables significantly predicted the discrepancy variable at a 10% significance level in a multivariable logistic regression. Only those variables remaining were again entered together in the multivariable logistic regression with discrepancy variable as the outcome. Each of the variables remaining in the model were then tested individually for differences between the “only Performance-Based functionally impaired” and the “only Self-Report functionally impaired” levels of the discrepancy variable.

Additional analyses explored the relationship between scores on the individual Self-Report and Performance-Based measures to self-reported depressive symptoms (i.e., BDI-II) using nonparametric Spearman’s correlations.

RESULTS

HAND Frequencies Across Self-Report, Performance-Based and Dual Methods

As shown in Figure 3, the Dual method yielded the lowest prevalence of ANI diagnoses at 59% as compared to 83% with the Performance-Based method (\( \chi^2 = 56; \ p < .001 \)) and 76% with the Self-Report approach (\( \chi^2 = 39; \ p < .001 \)). Additionally, the Performance-Based method classified more participants as ANI than the Self-Report approach (\( \chi^2 = 5.1; \ p = .02 \)). The Dual approach detected the most MND classifications at 36% as compared to 17% with the Performance-Based method (\( \chi^2 = 39.2; \ p < .001 \)), and 20% with the Self-Report method (\( \chi^2 = 38; \ p < .001 \)). The Performance-Based method yielded the least number of HAD diagnoses at 0.009% (\( n = 1 \)) compared to Self-Report (\( \chi^2 = 9; \ p = .003 \)) and the Dual method (\( \chi^2 = 10; \ p = .002 \)), which were comparable in their number of HAD diagnoses at 4% and 5%, respectively.

Among the symptomatic diagnoses, there was an additive effect between each of the singular diagnoses (Self-Report = 24% and Performance-Based = 17% symptomatic) to comprise the Dual diagnoses (i.e., 41%) suggesting that there are not many overlapping symptomatic diagnoses across each of these singular methods.

Sensitivity and Specificity of the Singular Methods Compared to the Dual Method

Using the Dual method as the “benchmark” to compare the singular methods against, the Performance-Based and Self-Report diagnoses had comparable sensitivity in detecting MND versus ANI classifications (42.7% vs. 57.3% respectively; \( \chi^2 = 2.2; \ p = .14 \)). The Performance-Based method was also comparably sensitive to the Self-Report method in determining symptomatic (i.e., MND+HAD) versus asymptomatic (i.e., ANI diagnoses) (41.2% vs. 58.8% respectively; \( \chi^2 = 3.8; \ p = .052 \)). However, the Self-Report method was more sensitive in detecting HAD versus ANI+MND (i.e., non-HAD) diagnoses than the Performance-Based method (90.9% vs. 9.1%; \( \chi^2 = 7.5; \ p = .006 \)). Specificity (i.e., proportion of individuals that are correctly identified as functionally unimpaired)

![Fig. 3. The Dual classification method yielded the lowest prevalence of ANI and largest prevalence of symptomatic diagnoses compared to either singular method. Each row represents the proportion of specific HAND diagnoses by assessment method among the participants with HIV-associated neurocognitive impairment. NCI = neurocognitive impairment; ANI = asymptomatic neurocognitive impairment; MND = mild neurocognitive disorder; HAD = HIV-associated dementia.](image-url)
Table 3. Multivariable logistic regression model predicting four-level discrepancy variable: 1) agree: asymptomatic; 2) agree: symptomatic; 3) discrepant: functionally impaired by Self-Report only; 4) discrepant: functionally impaired by Performance-Based only.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\chi^2$</th>
<th>$p$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: Multivariable logistic regression model.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall model: $R^2 = 0.27$, $\chi^2 = 102.9$, $p &lt; .001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age†</td>
<td>6.8</td>
<td>.08</td>
</tr>
<tr>
<td>Ethnicity (Caucasian vs. other)†</td>
<td>6.8</td>
<td>.08</td>
</tr>
<tr>
<td>Education†</td>
<td>6.9</td>
<td>.08</td>
</tr>
<tr>
<td>Gender</td>
<td>0.80</td>
<td>.85</td>
</tr>
<tr>
<td>Current CD4†</td>
<td>8.5</td>
<td>.04</td>
</tr>
<tr>
<td>AIDS</td>
<td>0.55</td>
<td>.91</td>
</tr>
<tr>
<td>Log_{10} HIV RNA (CSF)</td>
<td>2.9</td>
<td>.41</td>
</tr>
<tr>
<td>Log_{10} HIV RNA (plasma)</td>
<td>2.7</td>
<td>.43</td>
</tr>
<tr>
<td>HCV status†</td>
<td>6.8</td>
<td>.08</td>
</tr>
<tr>
<td>BDI-II†</td>
<td>27.9</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>LT substance abuse/dependence</td>
<td>4.4</td>
<td>.21</td>
</tr>
<tr>
<td>Employment†</td>
<td>20.3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Global NP T-score</td>
<td>0.37</td>
<td>.95</td>
</tr>
<tr>
<td>Step 2: Multivariable logistic regression model with significant ($p &lt; .10$) predictors.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall Model: $R^2 = 0.46$, $\chi^2 = 59.45$, $p &lt; .001$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age*</td>
<td>8.2</td>
<td>.04</td>
</tr>
<tr>
<td>Ethnicity*</td>
<td>9.1</td>
<td>.03</td>
</tr>
<tr>
<td>Education*</td>
<td>5.5</td>
<td>.14</td>
</tr>
<tr>
<td>Current CD4*</td>
<td>14.6</td>
<td>.002</td>
</tr>
<tr>
<td>HCV status*</td>
<td>6.30</td>
<td>.10</td>
</tr>
<tr>
<td>BDI-II*</td>
<td>28.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Employment*</td>
<td>18.0</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

$^\dagger p < .10; ^* p < .05$. HCV = Hepatitis C co-infection; BDI-II = Beck Depression Inventory-II; LT = lifetime; NP = neuropsychological.

Findings from this study indicate that multimodal assessment of functional status better detects symptomatic HAND as compared to singular methods of ascertainment. This is congruent with other research that has established the incremental value of incorporating information from multiple sources in determining diagnoses (Meyer et al., 2001; Schwartz et al., 1996). A moderate (76%) concordance existed between the Self-Report and Performance-Based methods in classifying functional impairment. Of the two methods, the asymptomatic agreement group endorsed fewer depressive symptoms ($\chi^2 = 17.3; p < .001$) and was more likely to be employed ($\chi^2 = 11.4; p < .001$) than the symptomatic agreement group, as expected.

**Depression**

As depressive symptoms have been shown to be highly associated with functional status and can particularly influence self-reported functioning (Rourke, Halman, & Bassel, 1999b; Thames, Becker, et al., 2010), we examined the relationship of Beck Depression Inventory-II scores to all of the functional measures. Both of the Self-Report measures were correlated to BDI-II scores (IADL: $p = 0.44$; $p < .001$; PAOFI: $p = 0.54$; $p < .001$). Neither of the Performance-Based measures was correlated to BDI-II scores (Valpar: $p = -0.01; p = .74$; MMT-R: $p = 0.007; p = .87$).

**DISCUSSION**

Across all of the different HAND levels was comparable between the Self-Report and Performance-Based methods (MND vs. ANI: 51.6% vs. 48.3%; and MND+HAD vs. ANI: 52.3% vs. 47.7%; HAD vs. ANI+MND: 51.0% vs. 49.0%; $p$’s > .05).

**Discrepant Singular Diagnoses**

When examining the discrepancy variable (i.e., discordant singular classifications) across demographic and clinical variables of interest, age, ethnicity (Caucasian vs. other), education, current CD4, HCV status, BDI-II, and employment status predicted the discrepancy variable at the $p < .10$ level. In a multivariable analysis, all of these predictor variables uniquely contributed to discrepant diagnoses except education and HCV status (see Table 3). Specifically, the Performance-Based functionally impaired participants were less likely to be Caucasian ($\chi^2 = 10.5; p = .001$), less likely to be employed ($\chi^2 = 9.1; p = .003$), endorsed fewer depressive symptoms ($\chi^2 = 7.6; p < .006$), and had lower current CD4 counts ($\chi^2 = 11.6; p < .001$) than the Self-Report functionally impaired individuals. See Figure 4 for discrepant diagnosis frequencies. Although age was a significant predictor of the discrepancy variables in the omnibus multivariable test, pairwise analyses did not indicate any group differences across the individual levels of the discrepancy variable ($p$’s > .05). Within the discordant diagnoses between the singular methods, the asymptomatic agreement group endorsed fewer depressive symptoms ($\chi^2 = 17.3; p < .001$) and was more likely to be employed ($\chi^2 = 11.4; p < .001$) than the symptomatic agreement group, as expected.

**Fig. 4.** Frequency of discrepant Self-Report versus Performance-Based classifications. Agree: No functional impairment = 67% (156/233); Agree: Functional impairment = 9% (20/233); Discrepancy: Impaired by Self-Report only = 16% (37/233); Discrepancy: Impaired by Performance-Based only = 9% (20/233). Abbreviations: NCI = neuropsychological impairment.
unemployed than those who were called functionally impaired only by Self-Report. In addition, participants impaired only on the Performance-Based approach had more immunosuppression. These findings support the validity of the Performance-Based measures in addition to the Self-Report measures in being able to detect “real life” constructs that are theoretically important to functional status.

Importantly, the prevalence of ANI changed from 76% using only Self-Report to 59% using the Dual Self-Report plus Performance-Based approach. This finding demonstrates that individuals classified as ANI via traditional Self-Report assessments may not be as “asymptomatic” as they are reporting, thereby inflating the prevalence rates of ANI. Instead, participants who were classified as functionally impaired by Self-Report only showed more depressive symptoms, which are known to be associated with negative self-image and a tendency to over-report functional impairment (Rourke et al., 1999b; Thames, Becker, et al., 2010). Depressive symptoms were to some extent accounted for in the Self-Report classification in that those individuals with elevated depressive symptom scores (BDI-II ≥ 17) need to have substantially more complaints (PAOFI ≥ 10) to meet HAND criteria. Therefore, despite a restricted BDI-II range, the effect of mild depression was still captured in the Self-Report diagnoses. The influence of depressive symptoms in the Self-Report classifications that is absent in the Performance-Based classifications illustrates the powerful interrelationship between depressive symptoms and complaints and, subsequently, how a belief of impairment can influence diagnosis. The potential impact of depressive symptoms in determining self-reported functional status highlights the importance of taking affective state into account when determining HAND classifications based on self-report.

On the other hand, since those individuals classified as functionally impaired only by Self-Report appeared to have more functionally complex lives (i.e., more likely to be employed and less disease severity), the importance of any decline may have impacted the number and severity of reported complaints relative to expectations in their lives. In other words, those with the most complicated lives had the most to lose functionally and were, therefore, more likely to complain of these changes. Self-report measures inherently control for variables relative to individual (e.g., perceived premorbid functioning and life expectations), and, therefore, reflect the individual’s inability to complete tasks specific to his or her own life. Although Performance-Based measures do not account for the subjective relevance of functional decline, they do standardize impairment so that individual performances can be equally compared to each other. In this manner, the Performance-Based approach provides a standard diagnostic of impairment for each HAND level thereby improving the ability for clinicians and researchers to communicate with common understanding.

An important limitation of the current study is the use of the Dual diagnostic method to compare the singular methods when the Dual method is just a composite of both of the singular methods together. As such, an association between the singular and dual methods is anticipated, and the true sensitivity and specificity of each of the singular methods to real life HAND diagnoses is not clear. Ideally, the Self-Report and Performance-Based methods should be compared to an independent measure of daily functioning, such as a clinician’s rating, informed proxy report, or an objective outcome measure (e.g., medication blood levels as an indicator of adherence). Additionally, it is important to note that each of the three formulas for HAND classifications were purely data-driven—clinical input is an integral aspect in determining HAND diagnoses and should always be incorporated in addition to the available data. Another potential limitation of our study is the use of the PAOFI to assess Self-Reported functional aptitude. The PAOFI addresses several cognitive skills in relation to daily functioning [e.g., “Do you have more difficulty now than you used to in calculating or working with numbers? (including managing finances, paying bills, etc.)”], and as such tend to emphasize patients' cognitive complaints rather than functional complaints; however, the PAOFI is a widely used assessment in assignment of HAND diagnoses as well as HIV-related functional declines in a large consortium of research (e.g., Heaton et al., 2004, 2011; Joska et al., 2011; Woods et al., 2004; Woods, Moore, Weber, & Grant, 2009). Lastly, the use of the Valpar/COMPASS program to evaluate everyday functioning may be considered a limitation of the current study given that several of the subtests included may not appear to be face valid for many jobs (e.g., threading a needle through a board); as such, these results should be interpreted accordingly. Nevertheless, the Valpar/COMPASS program as a whole has been designed to map onto Department of Labor determinations of requirements of jobs in the U.S. economy and has been used in previous studies to classify HIV-related functional impairment (e.g., Heaton et al., 2004; Twamley et al., 2006; Wright, Woo, Barclay, & Hinkin, 2009).

To summarize, findings from this study indicate that incorporating information from functional Performance-Based measures in addition to the traditional Self-Report approach detects more symptomatic impairment; using either Self-Report or Performance-Based measures alone classifies less functional impairment. In particular, use of only traditional functional Self-Report measures yields significantly higher rates of ANI compared to the Dual method. These findings support the commonly observed, and perhaps disproportionately, high levels of “asymptomatic” HAND reported in clinical studies when only self-report measures of functional impairment are used. Although Self-Report measures allow individuals to anchor their functional difficulties specific to the demands in their own daily lives, self-reported functional complaints appear to be significantly influenced by depressive symptoms. On the other hand, Performance-Based measures standardize deficits onto a comparable scale and, importantly, may be most predictive of true disease and functional status; this is supported by greater associations with employment and degree of current immunosuppression. Another benefit of the Performance-Based approach is the measures’ lack of relationship to depressive symptoms. As such, we suggest that Performance-Based measures are a useful and valid component
of functional assessments, which are likely to enhance the accuracy of symptomatic HAND classifications (beyond Self-Report only). Of note, however, Performance-Based measures do take more time and training than Self-Report which may not be practical in many clinical settings; nonetheless, clinicians and researchers should be aware that use of Self-Report only is likely to both underclassify symptomatic HAND and result in some false positive diagnoses due to depression-related biases in self-evaluations.

Overall, when assessing functional status, our findings support the time and resources necessary to incorporate Performance-Based measures in addition to the traditional Self-Report method in individuals with HIV. The current study uses a comprehensive approach in identifying HAND, which allows for objective evaluation of the best approach to define this phenomenon; however, the ultimate goal, particularly for clinicians, would be to establish brief, performance-based assessments that are able to capture those individuals identified as functionally impaired in a quicker manner. Our study helps identify a targeted battery for detecting HAND (i.e., inclusion of both Self-Report and Performance-Based functional measures), yet future studies are necessary to validate a brief battery including a combination of these assessments to be used clinically.

ACKNOWLEDGMENTS

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: J. Allen McCutchan, M.D., Ronald J. Ellis, M.D., Ph.D., Thomas D. Marcotte, Ph.D.; Center Manager: Donald Franklin, Jr.; Neurological 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: Joseph K. Wong, M.D. (P.I.), J. Hampton Atkinson, M.D., Steven Paul Woods, Psy.D., Matthew Dawson; Virology Component: Joseph K. Wong, M.D. (P.I.); Imaging Component: Terry Jermin, Ph.D. (Co-P.I.), Michael J. Taylor, Ph.D., Rebecca Theilmann, Ph.D.; Data Management Unit: Anthony C. Gamst, Ph.D. (P.I.), Clint Cushman (Data Systems Manager); Statistics Unit: Ian Abramson, Ph.D. (P.I.), Florin Vaida, Ph.D., Reena Deutsch, Ph.D., Anya Umlauf, M.S., Sanya Wolfson, M.A. This work was also supported by a T32 grant DA031098 from NIDA. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, nor the United States Government. Authors report no conflicts of interest affecting this article.

REFERENCES


