A comparison of screening batteries in the detection of neurocognitive impairment in HIV-infected Spanish speakers

Andrew J Levine1
Manuel Palomo1
Charles H Hinkin2,3
Miguel Valdes-Sueiras1,5
Enrique Lopez3,4
Glenn Mathisen1,5
Suzanne Donovan1,5
Elyse J Singer1

1National Neurological AIDS Bank, Department of Neurology, David Geffen School of Medicine at UCLA,
2Veterans Affairs Greater Los Angeles Healthcare System, 3Department of Psychiatry and Behavioral Sciences, David Geffen School of Medicine at UCLA, 4Department of Psychiatry and Behavioral Neuroscience, Cedars-Sinai Medical Center, Los Angeles; 5Olive View-UCLA Medical Center, Sylmar, CA, USA

Background: A substantial number of Spanish-speaking individuals from Mexico and Central America are now living in the United States. These individuals are at heightened risk for HIV infection and, due to late diagnosis and limited resources, for HIV-associated neurocognitive disorders (HAND). Early detection is key, yet adequate methods for detecting HAND in Spanish speakers, especially in resource-poor areas, remains problematic. Therefore, it is necessary to identify accurate yet efficient neurocognitive screening tools that are appropriate for use in resource-limited AIDS clinics serving Spanish-speaking patients.

Methods: Twenty-one Spanish-speaking, HIV-positive adults who migrated from Mexico or Central America underwent neuromedical and neurocognitive evaluation in Spanish. The concordance of three neurocognitive screening measures (the HIV Dementia Scale [HDS], the Mini-Mental State Examination [MMSE], and the NEUROPSI) with a comprehensive neuropsychological battery was examined. In addition, accuracy in detecting neurocognitive impairment using standard and alternative cutoff scores was examined.

Results: The HDS and the NEUROPSI showed high correlation with the comprehensive neuropsychological battery. The HDS and the NEUROPSI also had the highest sensitivity (67% and 75%, respectively) and specificity (50% and 38%, respectively). Both measures also showed greater sensitivity than the MMSE to very mild forms of HAND.

Conclusion: In this small sample of HIV-positive Spanish speakers from Mexico and Central America living in the United States, the HDS and the NEUROPSI demonstrated reasonable accuracy in detecting neurocognitive impairment, while the MMSE demonstrated very poor accuracy. The HDS and the NEUROPSI were equally sensitive in detecting mild HAND. Continued test development is required to capture this disorder, especially in resource-limited settings.

Keywords: HIV-associated neurocognitive disorders, Spanish speaking, NEUROPSI, HIV Dementia Scale, Mini-Mental State Examination

Introduction
The population of Latino immigrants is very high in the United States, especially in western states. For example, the 2000 US Census estimated that more than 8.7 million people living in California were of Mexican origin, with approximately 44% born in Mexico.1 Additionally, the second-largest Latino group in California is from Central America.2 Studies of primarily Mexican and Central American migrant workers indicate a high risk for HIV infection.3-5 This group includes young men who are away from their homes for prolonged periods of time, a situation that can result in a high number of sexual partners, sex with prostitutes, and sex between men. Further, the working conditions themselves may contribute to increased risk, including hazardous conditions, low pay, exploitation, and disruption of relationships in the country of origin.6
Background characteristics of Mexican migrant workers that exacerbate these risks include poor education, low literacy, limited English proficiency, traditional gender roles, and limited access to health and social services. Research has indicated that these factors are linked to survival strategies such as trading sex for money and illicit drugs including drugs abused intravenously, thus increasing risk for HIV infection.\(^7\) Compounding the problem further, diagnosis of HIV infection is often delayed in Latinos. For example, in the greater Los Angeles area, over 70% of Latinos with AIDS had their HIV infection detected very late in the progression of their illness.\(^2\) It can be expected that the increasing number of migrant laborers and their families originating from Mexico and Central America will be accompanied by an increase in HIV-positive individuals seeking health care services.

Because of the tendency toward late diagnosis and treatment of HIV among Mexican and Central American immigrants in the United States, these individuals are also at increased risk for HIV-associated neurocognitive disorders (HAND). Even in developed countries where there is greater access to combined antiretroviral therapy (cART), the prevalence of HAND has increased as a paradoxical consequence of decreased mortality.\(^11,12\) The prevalence of HAND in developing countries, where access to cART is limited, is likely to be considerably higher.\(^13\) In addition to HAND, individuals with HIV who are not receiving cART are also at risk of developing non-HIV-related neurocognitive problems, such as opportunistic infections of the brain. Thus, Mexican and Central American immigrant workers not only carry an increased risk for contracting HIV while in the United States but also they may come to this country with more advanced HIV infection and are more likely to evidence neurological symptoms when they first seek treatment. The consequence of these new realities is that US health care workers must be prepared for the detection and treatment of neurocognitive impairment among this growing segment of the HIV-positive population. However, many Latino immigrants are treated in general medical clinics and/or AIDS clinics without ready access to neurologists, neuropsychologists, or the like, so a screening battery appropriate for use with these individuals that could be administered by frontline clinicians would be extremely helpful in identifying those who need further evaluation and referral.

In this study, the authors compared the accuracy of three screening measures in detecting neurocognitive impairment in what was primarily a monolingual Spanish-speaking, HIV-positive sample from Mexico and neighboring Central American countries. The screening measures employed were two commonly used measures, the Mini-Mental State Examination (MMSE) and the HIV Dementia Scale (HDS), and a measure yet to be widely employed in the context of HIV, the NEUROPSI.\(^18\) The NEUROPSI is a neuropsychological screening measure developed in Mexico that provides normative data for a wide range of ages and education levels. However, as of yet, the psychometric properties of this battery have not been evaluated in a Spanish-speaking, HIV-positive population living in the United States.

**Materials and methods**

**Participants**

Twenty-one HIV-positive individuals (18 men and three women) were recruited from the National Neurological AIDS Bank (NNAB) study at the University of California in Los Angeles. The participants were aged 30–61 years \((Mean_{\text{age}} = 47.3\text{ years};\ \text{standard deviation } [SD] = 8.2)\) and years of formal education ranged from 2 to 16 years \((Mean_{\text{education}} = 8.7\text{ years}; SD = 4.5)\). The NNAB is a National Institutes of Health–funded longitudinal study investigating the neurobehavioral and neuropathologic effects of HIV infection. Eligible participants were offered monetary compensation to take the NEUROPSI in addition to the standard NNAB protocol, which included the HDS, MMSE, and comprehensive neuropsychological testing. Participants signed an informed consent form before any procedures were initiated.

Over an 18-month period, twelve participants from Mexico and nine from Central America (five from El Salvador, two from Guatemala, one from Nicaragua, and one from Honduras) were recruited. Median CD4\(^+\) cell count was 254 cells/mm\(^3\) \((SD = 152)\) and all participants were on a cART regimen.

The inclusion criteria were as follows: (1) primary Spanish speaker; (2) aged 18 years or older; (3) raised in Mexico or Central America; and (4) able to complete the neuropsychological screening measures. Twenty participants were monolingual Spanish speakers and one was considered bilingual (English and Spanish). All testing was conducted in Spanish. The study was carried out in accordance with human subjects protection guidelines of the University of California, Los Angeles and with permission of the Medical Institutional Review Board.

**Measures**

**NNAB comprehensive neuropsychological battery**

A bilingual psychometrist administered a Spanish translation of the standard NNAB neuropsychological battery to...
the participants under the supervision of a board-certified neuropsychologist (CHH). The battery consists of a number of measures that are widely used and psychometrically validated when used with English-speaking individuals from the United States. These tests were translated directly into Spanish for use with the Spanish-speaking participants. However, appropriate normative data for most measures do not exist for use with the Spanish-speaking participants. The NNAB Comprehensive Neuropsychological Battery (NNAB-NB) assesses seven neurocognitive domains (with constituent measures in parentheses): (1) processing speed (Digit Symbol and Symbol Search subtests from the Wechsler Adult Intelligence Scale – Third Edition14); (2) attention/working memory (Series 1 from the Paced Auditory Serial Addition Test15,16 and Letter Number Sequencing subtest of the Wechsler Adult Intelligence Scale – Third Edition14); (3) learning (total score from learning trials from the Hopkins Verbal Learning Test – Revised17 and Brief Visuospatial Memory Test – Revised18); (4) memory (recall total score from the Hopkins Verbal Learning Test – Revised17 and Brief Visuospatial Memory Test – Revised18); (5) verbal fluency (Controlled Oral Word Association Test19); (6) abstract/executive functioning (Wisconsin Card Sorting Test – 64-card version20 and Trail Making Test – Form B21); and (7) upper extremity motor speed (Grooved Pegboard22). Standardized T-scores are obtained based on published normative data. For the current study, a global measure of neurocognitive functioning (global T-score) was obtained by averaging all individual test T-scores. This battery takes approximately 2 hours to administer and score.

HDS
A board-certified, Spanish-speaking neurologist (MV-S) administered a Spanish translation of the HDS23 as part of a neuromedical examination. The HDS assesses learning, attention, psychomotor speed, memory, and construction. Interpretation of the HDS is quantitative, with a maximum score of 16. Scores of 10 points or less are considered indicative of cognitive impairment. Administration and scoring of the HDS takes 10–15 minutes.

MMSE
A board-certified, Spanish-speaking neuropsychologist (MV-S) administered a Spanish translation of the MMSE24 as part of a neuromedical examination. The MMSE is a widely used screening measure for neurocognitive deficits and dementia. A score of 23 or less was considered indicative of cognitive impairment. The MMSE takes approximately 5–10 minutes to administer and score.

NEUROPSI
The NEUROPSI25 was administered in Spanish by a trained psychometrist under the supervision of a board-certified neuropsychologist (AJL). The NEUROPSI assesses a wide spectrum of cognitive functions, including orientation, attention, memory, language, visuospatial abilities, and executive functions. The NEUROPSI is composed of items that are relevant for Spanish-speaking communities. In developing the NEUROPSI, normative data were collected from 800 Spanish-speaking individuals from five Mexican states who were between the ages of 16 and 85 years. In the NEUROPSI, normative data is stratified into four age groups: (1) 16–30 years; (2) 31–50 years; (3) 51–65 years; and (4) 66–85 years. Data is also stratified into four education levels: (1) illiterates (0 years of schooling); (2) 1–4 years of schooling; (3) 5–9 years of schooling; and (4) 10 or more years of formal education. For the current study, the authors examined both the raw score and a categorical designation based on normative data stratified for age and education level (no impairment, mild, moderate, and severe impairment). Psychometrically, the NEUROPSI has been shown to have good reliability.25 Administration time is 15–20 minutes, as well as approximately 5 minutes for scoring.

Neurocognitive diagnosis
Diagnosis of neurocognitive impairment was determined via consensus agreement between the examining study neurologist and a board-certified neuropsychologist, with consideration of the neuromedical examination, laboratory results (eg, viral load and CD4 count), neuroimaging (when available), and results of the NNAB-NB testing. Those individuals who were diagnosed as neurologically normal were classified as “unimpaired.” Those who were found to have impairments ranging from subsyndromal HIV-related deficits up to HIV-associated dementia were classified as “impaired.” Individuals with neurocognitive impairment due to other reasons (ie, non-HIV related) were also included in the impaired group. A consensus in neurocognitive diagnosis was reached for all 21 participants and was based on established criteria.26 Based on the consensus diagnosis, eight participants were classified as neurocognitively normal, two with subsyndromal HIV-related neurocognitive impairment (herein referred to as asymptomatic neurocognitive impairment [ANI]), one with minor cognitive/motor disorder,
four with HIV-associated dementia, and six with neurocognitive impairment due to other reasons.

**Statistical analyses**

Two general analyses were conducted. In the first, the authors sought to determine the correlation between the three screening measures and the global T-score from the NNAB-NB. This was based on the assumption that the NNAB-NB represents the best psychometric standard for detecting true neurocognitive impairment. Pearson product-moment correlation was employed for this analysis. In the second analysis, the authors sought to determine the diagnostic accuracy of the three screening measures when compared with the multidisciplinary diagnosis, which was considered the most suitable criterion for assessing accuracy. All cases were designated as impaired or unimpaired, based on their neurocognitive diagnosis as determined via multidisciplinary consensus. Standard cutoff scores were then used for each of the three screening measures and the global T-score from the NNAB-NB. Therefore, based on 20 cases, the NEUROPSI and the HDS had correlation coefficients of 0.636 (P = 0.003) and 0.72 (P < 0.001), respectively. The MMSE was not significantly correlated with the global T-score did not change markedly. As shown in Table 1, correlations between the three screening measures and the NNAB-NB did not statistically significantly correlate with the global T-score (−0.249, P = 0.289). Also shown in Table 1 are impairment rates (percent classified as neurocognitively impaired based on standard cutoff scores) for the three screening measures.

Area under the curve, sensitivity, and specificity for the three measures are reported in Table 2. At the standard cutoff scores, the MMSE showed a very low sensitivity of 8% but a high specificity of 88%. For the HDS, sensitivity was 67% and specificity was 50% at the standard cutoff score. The NEUROPSI cutoff score varied depending on age and education strata. When all impairment groups were collapsed, sensitivity of 75% and specificity of 38% were observed.

Alternate cutoff scores were explored. As shown in Table 2, when compared with the multidisciplinary consensus diagnosis, the MMSE sensitivity was modestly acceptable (50%) only at an almost errorless score of 29. Specificity was modest (60%) at this cutoff score. A cutoff score of 28 had even lower sensitivity (25%) with a similar specificity of 59%. For the HDS, at the standard cutoff score of ≤10 points the sensitivity was 67% and specificity was 50%. At a cutoff score of 12 points, sensitivity was very good (75%), with a sharp drop in specificity (19%). Finally, the optimal NEUROPSI cutoff score appeared to be 90 points, resulting in a modest sensitivity of 75% and specificity of 61%.

Due to the inherent difficulty in diagnosing ANI, all of the analyses were performed without the two individuals with this diagnosis. As displayed in Table 3, correlations with global T-score did not change markedly. As shown in Table 4, at the standard cutoff scores, sensitivities for all three screening measures improved slightly, whereas specificities remained unchanged. This indicates that, using the standard cutoff scores, individuals diagnosed with ANI were generally misclassified as neurocognitively normal. Indeed, the MMSE missed both ANI cases as normal, whereas the HDS and the NEUROPSI classified 50% as normal.

To examine if region of origin affected test performance or diagnosis, two additional analyses were conducted. First, the frequencies of impairment, as defined, were examined. Fifty-eight percent of those from Mexico were considered impaired, whereas 63% of those from other Central American countries were considered impaired. Mean scores on the three screening measures and the NNAB-NB did not statistically differ between the two groups.

**Discussion**

The identification of neurocognitive impairment among HIV-positive Latino immigrants visiting medical clinics is complicated by a frequent lack of on-site specialist resources

<table>
<thead>
<tr>
<th>Measure</th>
<th>Participants (n)</th>
<th>Mean score</th>
<th>SD</th>
<th>Impairment rate (‘%’ n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>21</td>
<td>27.8</td>
<td>2.4</td>
<td>(14.3%) 3</td>
</tr>
<tr>
<td>HDS</td>
<td>21</td>
<td>9.7</td>
<td>3.1</td>
<td>(61.9%) 13</td>
</tr>
<tr>
<td>NEUROPSI</td>
<td>21</td>
<td>86.4</td>
<td>9.9</td>
<td>(71.4%) 15</td>
</tr>
<tr>
<td>Global T-score</td>
<td>20</td>
<td>38.1</td>
<td>6.3</td>
<td>N/A</td>
</tr>
</tbody>
</table>

**Note:** Percent classified as neurocognitively impaired based on standard cutoff scores.

**Abbreviations:** HDS, HIV Dementia Scale; MMSE, Mini-Mental State Examination; N/A, not applicable; SD, standard deviation.
and by language and cultural and educational issues. This is further compounded by the fact that many such individuals have limited access to health care resources. This necessitates the use of effective screening batteries and is especially important because early and accurate detection of HAND allows for early intervention in clinical settings. 

In comparing the utility of three neurocognitive screening measures, the authors found that the HDS had the highest correlation with a more comprehensive neuropsychological battery (NNAB-NB), with the NEUROPSI a close second. It is not unusual that the MMSE correlated poorly with the NNAB-NB, as the MMSE has been found to be more useful for detecting cortical dementias, such as in Alzheimer’s disease, than the predominantly subcortical deficits attributed to HIV infection. Moreover, education level has a marked impact on MMSE scores. Ostrosky-Solis et al previously examined the utility of the MMSE among Spanish speaking individuals. Their findings indicated that those individuals with no formal education perform similarly to individuals with severe dementia, while those with minimal (1–4 years) schooling perform similarly to those with mild dementia. As such, the MMSE may not be useful for assessing cognitive dysfunction in people with little or no education. However, in the sample, even individuals with very low levels of education and moderate HAND or other neurocognitive impairment did well on the MMSE, indicating that it lacks sensitivity in this population. Not surprisingly, the MMSE misclassified both ANI cases as normal, and when these cases were not included it still had an unacceptably low sensitivity of 8%.

The HDS was found to correlate strongly with the NNAB-NB. Moreover, in English-speaking samples, the HDS has been shown to be superior to the MMSE in detecting severe HAND. However, it has been reported that the HDS is a poor tool for detecting mild to moderate impairment, such as in minor cognitive and motor disorder and ANI. This is especially alarming as the prevalence of milder forms of HAND is increasing while severe impairment (ie, HIV-associated dementia) has decreased. However, in a study by Morgan et al, when the HDS was used in conjunction with age and education normative data, sensitivity increased from 17.2% to 70.7%. While specificity dropped slightly, to 73.7%, the drop was tolerable, given the benefit of the sizeable sensitivity increase. However, despite this increase in sensitivity, the HDS may still not be sufficiently sensitive in capturing milder forms of HAND, as it identified only 50% of ANI cases in Morgan et al’s sample. The latter results are concordant with the results of the current study, where the HDS correctly identified 50% of the ANI cases in the sample.

The NEUROPSI yielded sensitivity of 75% and specificity of 38% in detecting HAND and other neurocognitive impairment in the sample. Like the HDS, it was able to accurately classify 50% of the ANI cases. This is promising given the difficulty in detecting these milder forms of HAND. While the NEUROPSI was not developed with the goal of detecting HIV-related neurocognitive impairments, it assesses many of the domains found to be affected by HIV infection. Considering its high sensitivity and modest accuracy in detecting ANI, it may be an attractive alternative to current screening measures.
The assessment of the monolingual Spanish speaker continues to be a challenging task for a variety of reasons. For example, Spanish speakers in the United States are not a homogenous group; rather, they are made up of peoples from different regions and cultures with differing levels of acculturation, all factors that have been shown to affect neuropsychological testing. For example, Razani et al. examined a group of Anglo-American/monolingual English speakers and an ethnically diverse group fluent in English and found that as acculturation levels increased in the ethnically diverse group, so did their scores. While the development of culture-specific normative data may help control for factors such as acculturation, normative studies are a long and expensive process; some argue that developing normative data for minority groups may also mask true impairment and may negate some patients’ much-needed access to certain services. Further, normative data cannot be used indefinitely and are time limited, necessitating data to be regularly recalibrated at future points.

The scarcity of tests developed specifically for Spanish-speaking groups often results in clinicians turning to direct translations from their English-language counterparts. For instance, the HDS has not been validated in Spanish. Indeed, in a study by Wojna et al., translating the HDS into Spanish was found to result in item-bias and this threatened the construct validity of the test overall as a consequence. Other biases inherent in directly translating English tests into Spanish have been described in detail elsewhere. In the case of the NEUROPSI, normative data were derived from Spanish speakers and stratified by age and education level, including those with no education. This makes the NEUROPSI more suitable for Spanish speakers, especially those originating from Mexico, and perhaps those from surrounding Central American countries as well. Its utility is also practical in the United States, given the large number of Mexican and Central American immigrants currently residing in many western states who have limited spoken English. Nonetheless, the utility of the NEUROPSI should be further determined in other HIV-positive Spanish-speaking groups.

### Conclusion

Clearly, there is an increasing need to implement effective neurocognitive screening measures among the growing Spanish-speaking population in the United States in order to identify individuals with neurocognitive impairment and to provide them with the necessary treatment. While the NEUROPSI appears to be a promising tool for the assessment of Spanish speakers residing in resource-poor areas of the United States, it does take longer to administer than both the MMSE and the HDS and it also requires scoring. As such, it may be less suitable for general health clinics than the HDS, the other measure examined that showed modest accuracy. Other cognitive screening measures such as the International HIV Dementia Scale may be helpful in detecting mild forms of HAND among these disenfranchised Spanish-speaking populations. However, given the lack of culturally valid neuropsychological tests and normative data for Spanish speakers, the NEUROPSI is potentially a powerful tool for the detection of mild HAND among Mexican and Central American immigrants residing in resource-limited settings.

The results of this study should be considered with the following caveats. First, while the NEUROPSI was developed for Spanish speakers originating from Mexico, the sample also included individuals from neighboring Central American countries. As such, cultural and language differences may have impacted performances on the NEUROPSI. Second, individuals with neurocognitive impairment due to non-HIV-related causes were included. However, considering the high frequency of such diagnoses among those with HIV, this should serve to improve the study’s generalizability. Finally, the study

### Table 4 Sensitivity and specificity for standard and alternate cutoff scores excluding asymptomatic neurocognitive impairment cases (n = 19)

<table>
<thead>
<tr>
<th>Measure</th>
<th>AUC</th>
<th>Standard cutoff</th>
<th>Alternate cutoff</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE</td>
<td>0.57</td>
<td>N/A</td>
<td>24</td>
<td>10%</td>
<td>88%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>28</td>
<td>25%</td>
<td>54%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>29</td>
<td>75%</td>
<td>27%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>69%</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>75%</td>
<td>10%</td>
</tr>
<tr>
<td>HDS</td>
<td>0.65</td>
<td>N/A</td>
<td>10</td>
<td>70%</td>
<td>50%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>69%</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td>75%</td>
<td>10%</td>
</tr>
<tr>
<td>NEUROPSI*</td>
<td>0.65</td>
<td>Varied</td>
<td>N/A</td>
<td>80%</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>90</td>
<td>75%</td>
<td>27%</td>
</tr>
</tbody>
</table>

**Notes:** NEUROPSI ordinal score (normal, mild, moderate, severe impairment). Cutoff scores are based on demographic strata. Sensitivity and specificity for standard cutoff scores based on collapsing all impaired strata. Alternative cutoff scores based on unstratified raw score for entire sample.

**Abbreviations:** AUC, area under the curve; HDS, HIV Dementia Scale; MMSE, Mini-Mental State Examination; N/A, not applicable.
was completed with a very small sample. As such, the findings should be considered preliminary in nature, until larger cohorts can further explicate the utility of these measures. Despite these limitations, the authors have reported the first comparative study of the NEUROPSI and neurocognitive screening measures in HIV-positive Spanish speakers, with findings that are likely to be of use in resource-limited areas.

Acknowledgments
This study was made possible through the following funding sources: UCLA AIDS Institute and the UCLA Center for AIDS Research (AI28697 [AJL]) and the NNAB (U01-MH08021 and R24-NS38841 [EJS]).

Disclosure
The authors report no conflicts of interest in this work.

References


