OPEN

A better screening tool for HIV-associated neurocognitive disorders: is it what clinicians need?

Marie-J. Brouillette[a,b,c], Nancy Mayo[c,d], Lesley K. Fellows[e], Elena Lebedeva[c], Johanne Higgins[f,g], Edgar T. Overton[h], Beau M. Ances[i] and Lisa Koski[j,k]

Objective: Existing screening tools for HIV-Associated Neurocognitive Disorders (HAND) may lack the accuracy required for clinical use. We hypothesized that the diagnostic accuracy of the Montreal Cognitive Assessment (MoCA) as a screening tool for HAND might be improved with a stronger scoring methodology.

Design: Two hundred HIV-positive participants aged 18–65 years completed the MoCA and a battery of neuropsychological tests.

Methods: HAND diagnosis was established according to the Frascati criteria, and an NPZ-8 score was also calculated. Rasch analysis was applied to the MoCA items to create a quantitative score.

Results: The optimal cut-off on the quantitative MoCA for detecting impairment as per Frascati criteria yielded a sensitivity of 0.74 and a specificity of 0.68. Overall accuracy was 0.79 (95% CI: 0.73–0.85), an improvement over standard scoring methods. However, whether cognition was quantified with the quantitative MoCA or with NPZ-8, there was substantial overlap between diagnostic categories; several individuals categorized as impaired had better overall cognitive function as assessed by NPZ-8 or quantitative MoCA than those classified as normal using standard criteria.

Conclusion: Quantifying performance on MoCA items through Rasch analysis improves its accuracy as a screening tool for HAND, and demonstrates that cognition can be measured as a unidimensional construct in HIV, at least at the level of precision of bedside testing. However, the current categorical diagnostic approach to HAND is poorly aligned with summary measures of cognitive ability. Measuring cognition as a quasi-continuous construct may be more relevant than conventional HAND diagnostic categories for many clinical purposes.

AIDS 2015, **29**:895–902

Keywords: cognition, cognition disorders, HIV, HIV-associated neurocognitive disorder, Montreal Cognitive Assessment, neuropsychological tests

[a]Department of Psychiatry, Faculty of Medicine, McGill University, [b]Chronic Viral Illness Service, McGill University Health Centre, [c]Research Institute of the McGill University Health Centre, [d]Division of Clinical Epidemiology, Department of Medicine, McGill University, [e]Department of Neurology and Neurosurgery, Montreal Neurological Institute, Faculty of Medicine, McGill University, [f]School of Rehabilitation, Université de Montréal, [g]Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal, Montreal, Quebec, Canada, [h]Division of Infectious Diseases, University of Alabama at Birmingham, Alabama, [i]Department of Neurology, University of Washington, St. Louis, Missouri, USA, [j]Department of Neurology and Neurosurgery, Faculty of Medicine, and [k]Department of Psychology, McGill University, Montreal, Quebec, Canada.

Correspondence to Marie-Josée Brouillette, McGill University Health Centre, Allan Memorial Institute, 1025 Pine Ave West, Montreal, QC, Canada, H3A 1A1.

E-mail: marie-josee.brouillette@mcgill.ca
Received: 13 August 2013; revised: 13 November 2013; accepted: 13 November 2013.

DOI:10.1097/QAD.0000000000000152

ISSN 0269-9370

Introduction

The prevalence of HIV-Associated Neurocognitive Disorder (HAND) in adults is estimated to range from 30 to 50% [1–4]. These neurocognitive deficits can lead to meaningful changes in everyday life, compromising occupational function and medication adherence [5–8]. Accurate diagnosis of HAND according to consensus Frascati criteria requires neuropsychological testing [9]. However, neuropsychological evaluation is not readily available in the majority of settings in which patients receive their care. One strategy to overcome this problem is to administer a brief screening tool to identify those who would benefit from full diagnostic evaluation. However, existing screening tools have poor diagnostic accuracy compared with the consensus criteria, with relatively high sensitivity coming only at the expense of low specificity [3,10–12]. For example, in one recent study, the four-item HIV Dementia Scale, with a cut off 14 or less (out of 16), had a sensitivity ranging from 83 to 88% and specificity ranging from 63 to 76% [3].

HIV-specific cognitive screening tests were initially developed to detect dementia rather than the mild impairment that is of increasing concern. Recent work has focused on the potential of generic screening tests to identify milder cognitive impairment. The Montreal Cognitive Assessment (MoCA), a quick and free tool developed to screen for mild cognitive impairment in geriatric populations [13,14], has been studied in this regard. When the threshold for classifying impairment was set very high ($\leq$27/30), the sensitivity to HAND was 90%, but this was associated with unacceptably low specificity (43%) [15]. When the threshold for impairment was lowered ($\leq$23/30), fewer people with HAND were detected (sensitivity 38%) but more of the cognitively normal individuals were correctly classified (specificity 95%). No threshold yielded an overall diagnostic accuracy of at least 70%, considered acceptable for clinical use [16].

The MoCA has reasonable face validity for HAND, with items testing a broad range of cognitive domains that are commonly affected in this condition, including memory, attention and frontal-executive functions [1]. Given this apparent validity, its reported accuracy is disappointingly low. One possibility is that MoCA items are appropriate, but the method of assigning an overall score is suboptimal. Although summing items to reach a total score has intuitive appeal, in fact this does not yield a true 'quantity' that relates in a systematic, mathematical way to cognitive ability [17]. Modern measurement theory, Rasch Measurement Theory in particular, provides an evidence base for the extent to which a set of items form a real measure of a single construct (in this case, cognition), the relationship of individual items to each other and an estimate of the 'distance' between items, allow the mathematical quantification of the construct [18]. We recently showed in a convenience sample of HIV-positive individuals [19] that the items of the MoCA fit the Rasch model creating a linear scoring system.

In the current study, we extend the analysis to test whether the diagnostic accuracy for HAND of the MoCA might be improved with a scoring methodology that produces linearized units of cognitive ability, a scoring feature of true 'measures'. We first undertook to replicate our prior finding that MoCA items could yield a meaningful 'quantity' of cognition in a new sample gathered in a tertiary-care HIV clinic setting in the United States. We then examined the relationship of cognitive ability quantified with this novel, Rasch-informed measure against gold standard diagnostic classification based on neuropsychological testing [15].

Methods

Study population

The study from which the data are derived has been described elsewhere [15]. Briefly, 200 HIV-infected participants aged 18–65 years, without confounding neurological or psychiatric conditions, recent AIDS-defining opportunistic infection or active substance abuse were recruited from the outpatient Infectious Disease Clinic at Washington University in St. Louis (WUSTL). At the same testing session, each participant completed the MoCA as well as a battery of neuropsychological tests: Timed Gait, Grooved Pegboard (motor skills); Hopkins Verbal Learning Test Revised (memory-learning and recall); Trail A, CalCAP choice and sequential reaction times, Symbol Digit, Stroop colour and words (speed of information processing); Trail B, Stroop Interference (executive functions). The Human Research Protection Office at WUSTL approved the study; written informed consent was obtained from all participants.

Measures

According to the 2007 Frascati definition, HAND comprises three distinct conditions: asymptomatic neurocognitive impairment (ANI), mild neurocognitive disorder (MND) and HIV-associated dementia (HAD). These conditions are defined by two parameters: performance on neuropsychological testing and functional impairment [9]. In this study, no documentation of functional impairment was acquired, so it was not possible to distinguish between the three categories. Therefore, the primary analysis used a two-category variable, normal or impaired, with the impaired category (performance $\geq$1.0 SD below appropriate norms in two or more domains) encompassing all three HAND categories.

Raw scores from each test were standardized using demographic (age, sex, race, education) adjusted normative means [1] and a standardized z-score was calculated. Timed Gait was excluded from the determination of

neuropsychological impairment. Although its inclusion is permitted by the *Frascati* classification, performance was abnormal ($\geq$1 SD below norms) in 62% of the participants, inflating the number of participants meeting HAND criteria, despite being of questionable relevance to a diagnosis of cognitive impairment.

This analysis used a slightly different neuropsychological scoring method than the previously published report from the same sample [15], selecting the test scores suggested by Woods [20] whenever applicable; this approach has been shown to improve inter-rater reliability. This led to a small subset of patients being re-classified compared with the previously reported study [15]. An analysis carried out with the previously reported classification showed very similar results. For simplicity, we report here only the more reliable classification.

Two secondary analyses were performed. In the first analysis, as suggested by Gisslen, we modified the cut-off that defines impairment to performance 1.5 SD or more below norms [20]. In a second analysis, the impaired group was split into 'mild impairment' (1.0 SD or more below norms in two or more domains) and 'severe impairment' (2.0 SD or more below norms in two or more domains), following the neuropsychological test classification of the Frascati criteria. Sufficient data were available for most participants to also calculate a widely used summary score, the NPZ-8 [21], by averaging performance on eight NP tasks.

Data analysis

Rasch analysis was carried out. Items were assessed for fit to the underlying model. People and items are aligned on the same scale, such that more difficult items are passed by fewer individuals and those individuals are the ones who have successfully answered more of the questions. The end result is a set of items representing a unidimensional construct, with responses on the items arranged hierarchically along a calibrated scale. Applied here, Rasch analysis transforms the ordinal MoCA score into a quantitative score, producing a 'measure' in the strict, quantitative sense. Units on this calibrated scale are in logits with a mean of 0 (probability of passing is 50%) and a SD of 1. Items that most people can pass are easy items and people who cannot pass them have less ability; items that few people pass are harder items and people who pass them have more ability. In the current analysis, higher logit scores represent better cognitive ability. Rasch analysis was performed using RUMM 2030.

To test the diagnostic accuracy of the MoCA scored according to published guidelines and this novel, quantitative MoCA scoring, sensitivity and specificity were calculated for each observed value against the presence or absence of HAND (i.e. normal or impaired). A Receiver Operator Characteristic (ROC) curve and c-statistic (the probability of predicting an outcome, considered reasonable when higher than 0.7 and excellent when higher than 0.8) were generated to establish the diagnostic accuracy of both MoCA scores. Standard metrics for comparing measurement properties of the two MoCA scoring systems were also calculated including coefficient of variation, percentage at ceiling and kurtosis (deviation from normal distribution).

Results

Characteristics of the sample were as follows: mean age (SD) 43 (11) years old, 72% were men, 67% African–Americans, mean (SD) educational level was 13(3) years. All participants were virologically suppressed on HAART ($<$400 copies/ml). The median (IQR) baseline and nadir CD4 counts (cells/μl) were 538 (361–695) and 191 (70–300), respectively. Baseline characteristics are described in more detail in [15]. Of the 200 participants, full neuropsychological results to calculate the NPZ-8 were available in 193, and a Rasch MoCA score could be calculated in 198 people. Based on neuropsychological test results, 74% met criteria for a diagnosis of HAND (ANI, MND and HAD combined).

Rasch analysis of the Montreal Cognitive Assessment items

A step-by-step process was followed to determine which MoCA items fit a single measure, to estimate their position in this measure, and to compare measurement features with the standard MoCA. Our prior work showed that scoring of both the clock and cube items on the MoCA may vary by rater. To avoid this potential problem, these items were all re-scored by a single rater. The repetition of five words and delayed memory for the words based on cued recall or multiple-choice formats are not scored in the standard MoCA. We created four new items (repeat 1, repeat 2, recall with category cue, and recall multiple choice) to determine if this additional information improved the measure.

In the first step of the Rasch analysis, the 'City' item was deleted as an extreme item (i.e. was answered correctly by all respondents, and therefore does not contribute to measurement). In the next step, all multiple response-category items (serial 7s, repeat and recall) with disordered thresholds were rescored taking into account the item's category probability curve. Recall with category cue was considered to be a misfitting item (fit residual < -2.5) and was eliminated from the data set. The Lion naming item showed a high residual correlation with responses on Date and Month (0.72), which means that a correct response to one of these items is associated with an increased likelihood of succeeding on the other item; this violates one of the rules of Rasch measurement, namely that items must be independent from one another. After excluding both Date and Month from the model, no other significant dependency was observed.

Unidimensionality of the model was confirmed, as no components of the Principal Component Analysis (PCA) of the residuals accounted for more than 8% of the total variance (acceptable level usually set as $< 10\%$). The Person Separation Index (PSI) of 0.68 for this quantitative MoCA scoring indicated that the power to discriminate among the patients with different levels of cognitive ability is relatively low (with 0.70 and 0.80 considered respectively acceptable and good). Finally, we examined all the items for differential item function (DIF), selecting consecutively one of four person factors: age ($\leq 45 / > 45$ years), race (Caucasian/African–American), sex (male/female), and education ($\leq 12 / > 12$ years). The Watch/Ruler abstraction item had uniform DIF for the sex and age factors; men and individuals more than 45 years performed better on this item in each of the three class intervals. Nonuniform DIF was found on Watch/Ruler abstraction (race factor), Recall Multiple (age factor), Trail (race factor) and Drawing the Hands of the Clock (sex factor). Neither splitting by DIF analysis (for Watch/Ruler abstraction), nor deleting items with nonuniform DIF improved model fit statistics or PSI, so these items were retained. Overall fit to the Rasch model was good, with a nonsignificant χ^2 probability value ($P = 0.43$). Figure 1 shows the distribution of people (upper part) and items (lower part) along the spectrum of global cognitive ability. The quantitative MoCA items cover a broad range of the construct of cognitive ability, ranging from −3.74 (Place-easiest item) to +3.20 logits (free recall of all 5 words-most difficult item). A considerable difference in the mean locations for persons (2.382 ± 1.039 logits) and for items (0.00 ± 1.77 logits) suggests that the quantitative MoCA is too easy for the persons who underwent testing, despite the fact that they had on average only 13 years of education. Six individuals had a perfect score on the Rasch MoCA scoring (3% ceiling effect), compared with 10 (5%) on the standard MoCA scoring.

Table 1. Statistical features of the MoCA scored according to published instructions and the Rasch quantitative MoCA.

	Ordinal MoCA	Quantitative MoCA
Mean (SD)	24.37 (3.69)	2.382 (1.039)
Coefficient of variation (SD/mean)	0.15	0.43
Median (IQR)	25 (4)	2.31 (1.51)
Kurtosis	0.54 (0.34)	0.10 (0.34)
Min/Max	12/30	−0.32/4.80
% at ceiling	10 (5%)	6 (3%)

Measurement characteristics

Table 1 shows that the quantitative (Rasch) scoring of the MoCA items results in better measurement features than the standard (ordinal) MoCA score, with a larger coefficient of variation, a kurtosis closer to 0 and fewer people at ceiling.

Distribution of the quantitative MoCA score for those categorized as normal or impaired (HAND) is presented

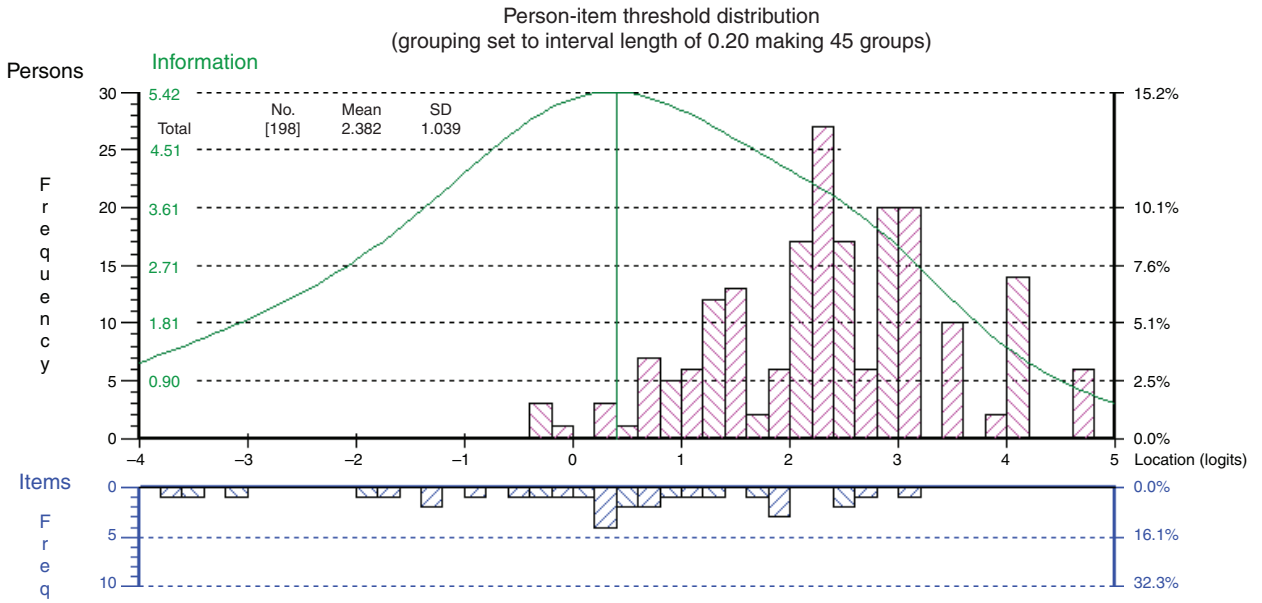

Fig. 1. Distribution of patients and item thresholds along the spectrum of cognitive ability. The hierarchy of cognitive ability is spread out along the *x*-axis, with 0 anchored to the test item of middle difficulty (50% chance of success on this item), easier items spreading to the left and more difficult items to the right. Bars descending below the *x*-axis show the number of items thresholds representing each level of cognitive ability. Bars ascending above the *x*-axis represent the frequency of people at each level of cognitive ability. Mean score on the quantitative MoCA (SD) is indicated on the top left part of the graph. The line plotted above the *x*-axis reflects the precision of measurement that can be obtained at each point along the continuum of cognitive ability. Here, best precision is obtained at a level much below the mean ability of the people, indicating that the items are too easy.

in Fig. 2a. We expected to see very little overlap of the two curves, consistent with distinct categories. On the contrary, substantial overlap was present with many individuals meeting criteria for HAND having better scores on the quantitative MoCA than many classified as normal. To assess whether this unexpected finding was specific to the MoCA, we also plotted the distribution of NPZ-8 scores for these two diagnostic categories (Fig. 2b). A similar degree of overlap across categories was again seen. This overlap was not resolved by modifying the cut-off that defines impairment to 1.5 SD or more below norms rather than 1 SD (Fig. 2c and 2d).

We next asked whether the discrimination between groups would be better if we compared the normal group and the severely impaired group only. We repeated the analysis, splitting the impaired into two groups: mild impairment [performance 1.0 SD or more below norms

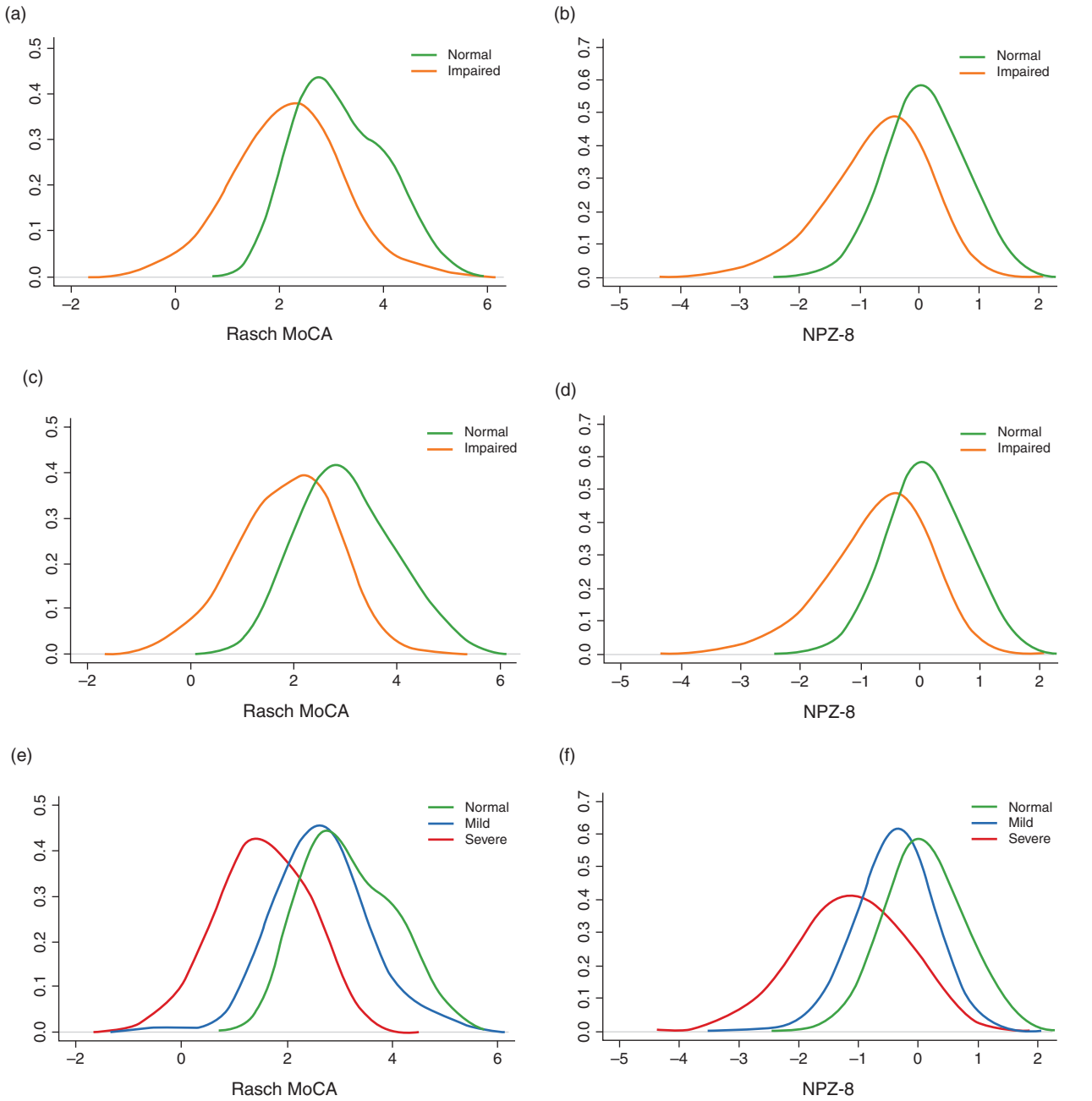

Fig. 2. Distribution of quantitative MoCA scores and NPZ-8 according to NP severity rating. (a) Distribution of quantitative MoCA scores for normal and all impaired (b) distribution of NPZ-8 scores for normal and all impaired (c) distribution of the quantitative MoCA scores for normal and all impaired, using $\leq$ 1.5 SD (d) distribution of NPZ-8 scores for normal and all impaired, using $\leq$ 1.5 SD (e) distribution of quantitative MoCA scores for normal, mildly impaired and severely impaired (f) distribution of NPZ-8 scores for normal, mildly impaired and severely impaired.

in two or more domains; $N = 83$ (42% of the whole sample)] and severe impairment [performance 2.0 SD or more below norms in two or more domains; $N = 66$ (33%)]. Significant overlap remained evident in the distribution of the quantitative MoCA (Fig. 2e) and the NPZ-8 scores (Fig. 2f) for these three groups. To confirm that the overlap in NPZ-8 was not driven by the Timed Gait score, which was included in the NPZ-8 but not used in the diagnostic classification, we repeated all analyses excluding the Timed Gait from the composite NPZ score. The degree of overlap did not change (data not shown). Hence, there is a sub-optimal relationship at the individual level between Frascati-defined categories and cognitive ability indexed by either the NPZ-8 or the quantitative MoCA.

Nonetheless, we constructed an ROC curve to evaluate the performance of the quantitative MoCA as a screening test for HAND (Fig. 3). For comparison with the standard MoCA scoring, we also included a curve for the cut-off with the highest accuracy (see Table 2 and Fig. 3). Evaluation using the ROC curve established that the optimal cut-off score on the quantitative MoCA for differentiating those meeting criteria for HAND from normal was 2.82 logits, which yielded a sensitivity of 0.74 and specificity 0.68. The predictive value of a positive test (PPV) on the quantitative MoCA was 0.87, the negative predictive value (NPV) was 0.47 and overall accuracy 0.79 (95% CI: 0.73–0.85). The ROC curves also show that, for a specific level of sensitivity, the quantitative MoCA scoring results in a higher specificity than the standard MoCA scoring. For example, when sensitivity was 0.49 for both scoring algorithms, the specificity for the standard MoCA was 0.86 and for the quantitative MoCA was 0.90; when sensitivity was 0.72 for both measures, specificity was 0.57 and 0.68, respectively.

Discussion

This study confirmed that Rasch analysis can provide an improved scoring algorithm for MoCA items, yielding a quantitative measure of cognitive ability with better measurement properties than the standard, ordinalMoCA score. We show that there is a consistent relationship between MoCA items assessing a variety of cognitive domains (memory, attention, etc.), allowing them to be ordered by level of difficulty to produce a single, ruler-like measure of cognitive ability. This finding in a sample drawn from an inner-city US clinic replicates our prior results in a demographically different Canadian sample, supporting the potential generalizability of this approach [19]. However, as in our initial study, we observed that MoCA items were sub-optimally targeted to the individuals being tested, with too few items of sufficient difficulty to measure the full spectrum of cognitive ability.

Can the quantitative MoCA score be used to screen for HAND? The optimal cut-off score yielded an accuracy of 0.79 (95% CI: 0.73–0.85) for predicting HAND, which is an improvement over that obtained using the standard MoCA scoring, and would be acceptable for clinical use. However, there was considerable overlap in quantitative MoCA performance between diagnostic categories: many people meeting criteria for HAND had better cognitive ability as measured by the quantitative MoCA

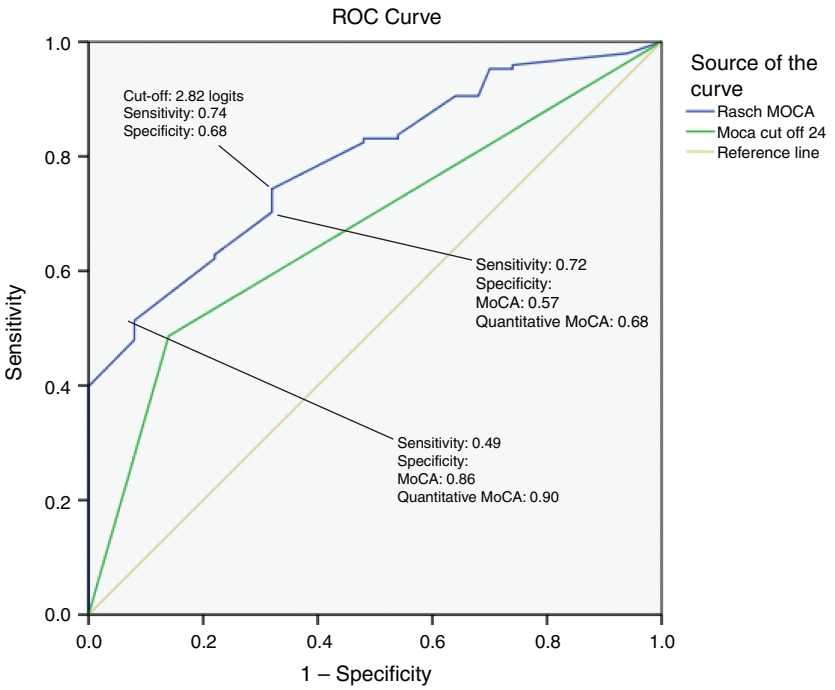

Fig. 3. Receiver operating characteristic (ROC) curve of the quantitative MoCA scores.

Table 2. Statistical features of different cut-off scores on the MoCA scored according to published instructions and the Rasch quantitative MoCA.

	Threshold for impairment by MoCA					
	≤27	≤26	≤25	≤24	≤23	≤2.82[a]
Number identified as impaired	156	129	101	80	52	126
Sensitivity	85.9%	71.8%	59.1%	49.0%	33.6%	74.3%
Specificity	45.1%	56.9%	74.5%	86.3%	96.1%	68.0%
PPV	82.1%	82.9%	87.1%	91.3%	96.2%	87.3%
NPV	52.3%	40.8%	38.4%	36.7%	33.1%	47.2%
Area under ROC	0.66	0.64	0.67	0.68	0.65	0.79
(95% CI)	(0.56–0.75)	(0.55–0.73)	(0.58–0.75)	(0.60–0.76)	(0.57–0.73)	(0.73–0.85)

NPV, negative predictive value; PPV, positive predictive value; ROC, receiver operating characteristic.
[a]Quantitative MoCA score, score in logits.

than people classified as normal. A similar pattern was seen within the HAND group, split by neuropsychological-defined severity of impairment. This overlap is not restricted to the quantitative MoCA: similar discordance was found between diagnostic categories and cognition as indexed by the NPZ-8.

This discordance arises from an important difference in the approaches taken for diagnosis of neurocognitive disorders on the one hand, and for measurement of cognition on the other. For the purpose of diagnostic classification, expert consensus established that, in light of the evidence currently available, cognitive impairment is best defined by abnormal performance on neuropsychological testing in two or more domains [9]. In contrast, the measurement approach, often applied in intervention trials and illustrated here by the NPZ-8, includes performance on all items, normal or impaired, aiming to capture 'global' cognitive performance; the MoCA uses a similar approach. Here, we found that diagnostic and measurement approaches do not align very well. Given the different scoring algorithms, this finding is not unexpected [22]; however, it suggests that it may be difficult to develop new screening tools for HAND that further improve diagnostic accuracy when the total score on such tools is obtained by adding the score of individual items, or when the items can be considered unidimensional rather than reflecting multiple, independent domains.

The Rasch approach we applied here to the MoCA items empirically determines which items can be combined to produce a single measure, and the 'distance' in level of difficulty between items. The calibrated score thus derived is a measure of a unidimensional construct, cognitive ability. This approach has major advantages for interpreting scores cross-sectionally (the same score means the same level of ability, which is not necessarily true for the standard MoCA), and for interpreting changes over time (a similar change in score represents the same quantity of cognitive ability irrespective of the starting point). Although the MoCA items need to be supplemented with more difficult items (such as measures of reaction times) for optimal measurement in this population, this can be accomplished without increasing test burden:The Rasch framework sets the stage for adaptive testing in which as few as eight items, selected from a larger item bank, are administered to a given person, based on their particular level of ability, without sacrificing measurement precision. The possibility of identifying meaningful change (or stability) in cognitive ability over time with a brief instrument could have important impact in many clinical settings.

Our study has limitations. First, the absence of information on functional impact of cognitive difficulties prevents full diagnostic classification. However, our conclusions do not hinge on these diagnostic distinctions. Second, although we use the term 'gold standard' to describe the neuropsychological testing here, in fact neuropsychological batteries vary across studies. The battery administered here is one that is commonly used. However, selection of different tests, a different number of tests, use of different norms or replication in a different population could lead to different results. This issue is not so much a limitation of our study, but rather a central problem for the current diagnostic approach.

Should clinicians use the diagnostic approach or the measurement approach? This depends on their goals. We suspect that in many clinical settings, a brief measure to track cognitive ability over time may be particularly useful, as deterioration could be identified before performance reaches the threshold of impairment in two domains required for HAND classification. This could be of particular importance in individuals with high premorbid cognitive ability, when significant decline needs to occur before reaching the impaired range. In addition, a focus on trajectory avoids the potential misclassification that can result from interpretation of neuropsychological testing in the absence of suitable norms [23], as is often the case in multicultural contexts. Of course, the goals of diagnosis and measurement are not mutually exclusive.

In conclusion, it is possible to develop a better screening method for HAND by applying Rasch Measurement Theory to the MoCA items. However, we should not be satisfied with this result in isolation. Lack of access to neuropsychological testing required for diagnostic

confirmation in those who screen positive currently constitutes a major impediment to the effective management of cognitive difficulties. This problem will not be solved by the availability of a better screening tool. Rasch analysis of MoCA items yields a quantitative score; application of this approach could lead to the development of measurement tools that we believe will be a breakthrough for clinical care. Such tools probably will not align with the current diagnostic classification: as this is not their aim, this should not be seen as a pitfall. A better screening tool for HAND may not be what is more urgently needed. Alternative approaches to the identification of cognitive decline that can be easily transferred to the clinic setting, such as the one we propose, merit further exploration.

Acknowledgements

We want to thank Dr Robert Heaton for his helpful comments on a previous version of the article.

M.J.B. and N.M. conceived of the study, interpreted the data, and drafted the article.

L.K.F. conceived of the study and contributed to the article.

L.K. designed and carried out the analysis, contributed to data interpretation, and contributed to the article.

J.H. and E.L. carried out the analysis and contributed to data interpretation.

B.M.A. collected data and reviewed the article.

T.O. collected data and reviewed the article.

This work was supported by the National Institute of Mental Health (K23MH081786 to B.M.A.); National Institute of Nursing Research (R01NR012907, R01NR012657, and R01NR014449 to B.M.A.), and by a CIHR Co-morbidity Team grant TCO-125272 (M.J.B., N.M., L.K.F., L.K.). E.T.O. is supported by the University of Alabama at Birmingham Center for AIDS Research (P30 AI027767).

Conflicts of interest

The authors report no conflicts of interest in this work.

References

1. Heaton RK, Franklin DR, Ellis RJ, McCutchan JA, Letendre SL, Leblanc S, *et al*. **HIV-associated neurocognitive disorders before and during the era of combination antiretroviral therapy: differences in rates, nature, and predictors.** *J Neurovirol* 2011; **17**:3–16.
2. Ciccarelli N, Fabbiani M, Colafigli M, Trecarichi EM, Silveri MC, Cauda R, *et al*. **Revised central nervous system neuropenetration-effectiveness score is associated with cognitive disorders in HIV-infected patients with controlled plasma viraemia.** *Antivir Ther* 2013; **18**:153–160.
3. Simioni S, Cavassini M, Annoni JM, Rimbault Abraham A, Bourquin I, Schiffer V, *et al*. **Cognitive dysfunction in HIV patients despite long-standing suppression of viremia.** *AIDS* 2010; **24**:1243–1250.
4. Heaton RK, Clifford DB, Franklin DR Jr, Woods SP, Ake C, Vaida F, *et al*. **HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study.** *Neurology* 2010; **75**:2087–2096.
5. Heaton RK, Marcotte TD, Mindt MR, Sadek J, Moore DJ, Bentley H, *et al*. **The impact of HIV-associated neuropsychological impairment on everyday functioning.** *J Int Neuropsychol Soc* 2004; **10**:317–331.
6. Gandhi NS, Skolasky RL, Peters KB, Moxley RT 4th, Creighton J, Roosa HV, *et al*. **A comparison of performance-based measures of function in HIV-associated neurocognitive disorders.** *J Neurovirol* 2011; **17**:159–165.
7. Thames AD, Kim MS, Becker BW, Foley JM, Hines LJ, Singer EJ, *et al*. **Medication and finance management among HIV-infected adults: the impact of age and cognition.** *J Clin Exp Neuropsychol* 2011; **33**:200–209.
8. Gorman AA, Foley JM, Ettenhofer ML, Hinkin CH, van Gorp WG. **Functional consequences of HIV-associated neuropsychological impairment.** *Neuropsychol Rev* 2009; **19**:186–203.
9. Antinori A, Arendt G, Becker JT, Brew BJ, Byrd DA, Cherner M, *et al*. **Updated research nosology for HIV-associated neurocognitive disorders.** *Neurology* 2007; **69**:1789–1799.
10. 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* 2009; **10**:246–252.
11. Valcour V, Paul R, Chiao S, Wendelken LA, Miller B. **Screening for cognitive impairment in human immunodeficiency virus.** *Clin Infect Dis* 2011; **53**:836–842.
12. Zipursky AR, Gogolishvili D, Rueda S, Brunetta J, Carvalhal A, McCombe JA, *et al*. **Evaluation of brief screening tools for neurocognitive impairment in HIV/AIDS: a systematic review of the literature.** *AIDS* 2013; **27**:2385–2401.
13. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, *et al*. **The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment.** *J Am Geriatr Soc* 2005; **53**:695–699.
14. Smith T, Gildeh N, Holmes C. **The Montreal Cognitive Assessment: validity and utility in a memory clinic setting.** *Can J Psychiatry* 2007; **52**:329–332.
15. Overton ET, Azad TD, Parker N, Demarco Shaw D, Frain J, Spitz T, *et al*. **The Alzheimer's disease-8 and Montreal Cognitive Assessment as screening tools for neurocognitive impairment in HIV-infected persons.** *J Neurovirol* 2013; **19**:109–116.
16. Hosmer D Jr, Lemeshow S, Sturdivant RX. *Applied logistic regression*. 3rd ed. Hoboken, New Jersey: John Wiley & Sons; 2013.
17. Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. **Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations.** *Lancet Neurol* 2007; **6**:1094–1105.
18. Andrich D. **Rating scales and Rasch measurement.** *Expert Rev Pharmacoecon Outcomes Res* 2011; **11**:571–585.
19. Koski L, Brouillette MJ, Lalonde R, Hello B, Wong E, Tsuchida A, *et al*. **Computerized testing augments pencil-and-paper tasks in measuring HIV-associated mild cognitive impairment(∗).** *HIV Med* 2011; **12**:472–480.
20. Gisslen M, Price RW, Nilsson S. **The definition of HIV-associated neurocognitive disorders: are we overestimating the real prevalence?** *BMC Infect Dis* 2011; **11**:356.
21. Clifford DB, McArthur JC, Schifitto G, Kieburtz K, McDermott MP, Letendre S, *et al*. **A randomized clinical trial of CPI-1189 for HIV-associated cognitive-motor impairment.** *Neurology* 2002; **59**:1568–1573.
22. Blackstone K, Moore DJ, Franklin DR, Clifford DB, Collier AC, Marra CM, *et al*. **Defining neurocognitive impairment in HIV: deficit scores versus clinical ratings.** *Clin Neuropsychol* 2012; **26**:894–908.
23. Winston A, Arenas-Pinto A, Stohr W, Fisher M, Orkin CM, Aderogba K, *et al*. **Neurocognitive function in HIV infected patients on antiretroviral therapy.** *PLoS One* 2013; **8**:e61949.