The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS): Utility in Detection and Characterization of Mild Cognitive Impairment due to Alzheimer’s Disease†

Stella Karantzoulis1,*, Julia Novitski2, Michael Gold3, Christopher Randolph4

1Department of Neurology, New York University Langone Medical Center, New York, NY, USA
2Department of Psychology, Rosalind Franklin University of Medicine and Science, North Chicago, IL, USA
3Alton Therapeutics, Vancouver, British Columbia, Canada
4Department of Neurology, Loyola University Medical Center, Maywood, IL, USA

*Corresponding author at: Department of Neurology, New York University Langone Medical Center, 145 East 32nd Street, 2nd floor, New York, NY 10016, USA.
Tel.: +1-212-263-3210; fax: +1-212-263-3273.
E-mail address: stella.karantzoulis@nyumc.org (S. Karantzoulis).

Accepted 25 June 2013

Abstract

Current diagnostic criteria for mild cognitive impairment (MCI) due to Alzheimer’s disease (AD) require standardized tests that are capable of measuring a range of neurocognitive abilities in healthy elderly individuals and sensitive to detect change over time. There currently is no clearly-established “gold standard” for this purpose. The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) is a widely used neuropsychological test battery for the clinical diagnosis/tracking of dementia also recently incorporated into clinical trials of new investigational medications for AD treatment. The RBANS has a number of design features that suggest possible utility in diagnosis/tracking of MCI. Eighty-one patients with MCI completed the RBANS and their scores were compared with 81 demographically matched healthy controls. RBANS Total Scale scores in both groups were normally distributed, demonstrating no floor/ceiling effects. The MCI group was most impaired on the Delayed Memory Index (DMI). Receiver operating characteristic analyses reflected good discrimination, with an area under the curve of 0.88 for the Total Scale score and 0.90 for the DMI score. The profile of performance for the MCI group was similar to that previously reported for mild AD patients. The RBANS may be a suitable neurocognitive battery for the detection and tracking of MCI presumed to be due to AD.

Keywords: RBANS; MCI; Alzheimer’s disease; Diagnosis; Clinical trials

Introduction

The revised National Institute on Aging and Alzheimer’s Association (NIA-AA) Workgroup criteria for a clinical diagnosis of mild cognitive impairment (MCI) due to Alzheimer’s disease (AD) require both the observation of a decline in cognition (typically in the domain of anterograde memory) and the documentation of impairment in cognition, preferably via standardized neurocognitive testing (Albert et al., 2011). In addition, patients should not be exhibiting impairments of a magnitude sufficient to constitute a dementia, which is defined by “significant” impairment in social or occupational functioning. Evidence of a profile of cognitive impairments that is similar to that observed in dementia due to AD is supportive of the core clinical diagnosis of MCI, as is evidence of progression of impairments over time.

There are, however, a number of impediments in arriving at the clinical diagnosis of MCI, as well as in defining MCI for the purposes of research (e.g., clinical trials). This is clearly illustrated by the fact that in many prospective studies, only a small percentage of patients initially diagnosed as having MCI progress to dementia on an annual basis. For example, in the InDDEXX clinical
trial of rivastigmine for MCI, only approximately 5% of patients in the placebo arm of the study progressed to AD on an annual basis (Feldman et al., 2007). Progression (or conversion) rates in some community-based studies are even lower than this, and in some studies, a larger percentage of subjects diagnosed with MCI actually improve over time and “normalize” rather than progress to AD (Palmer, Wang, Backman, Winblad, & Fratiglioni, 2002; Tyas et al., 2007). In clinic-based populations, conversion rates are somewhat higher, typically in the range of 10%–20% per year (Bruscoli & Lovestone, 2002). In the recent AD Neuroimaging Initiative (ADNI), approximately 400 subjects with MCI were recruited for a prospective study. The conversion rate to AD after 1 year was 16.5% (Petersen et al., 2010), with a very similar conversion rate observed in year 2 (Gomar, Bobes-Bascarán, Conejero-Goldberg, Davies, & Goldberg, 2011).

The apparent imprecision of the clinical diagnosis of MCI has led many researchers to advocate the use of biomarkers to “enrich” samples of MCI patients with individuals who are more likely to have prodromal AD. This was even emphasized in the NIA-AA revised guidelines referenced above (see Albert et al., 2011). In these revised guidelines, the likelihood of MCI being due to AD is classified as “unlikely,” “intermediate,” or “high,” based upon biomarker evidence. The predictive value of biomarkers to identify individuals who are likely to progress to dementia over the near future has not yet been confirmed. Direct comparison of neuropsychological and biomarker data has suggested that neuropsychological data may actually be superior to biomarker data in predicting conversion from MCI to AD (Gomar et al., 2011; Schmand, Eikelenboom, & van Gool, 2012). Furthermore, modeling of biomarker enrichment strategies for clinical trials has concluded that these may not significantly improve enrollment of appropriate patients and that cost–benefit ratios may be unacceptably high (Lorenzi et al., 2010; Schneider, Kennedy, & Cutter, 2010).

Although clinical assessments therefore are likely to remain critical for diagnosing MCI (and possibly for predicting progression/conversion), there is no consensus regarding which tests, or battery of tests, are best suited to this purpose. Multinational clinical trials have attempted to use the AD Assessment Scale-Cognitive Section (ADAS-Cog; Rosen, Mohs, & Davis, 1984), but this scale is highly insensitive to MCI. In a clinical trial involving 2,000 patients with MCI, Winblad and colleagues (2008) reported that the majority of subjects had baseline scores of 0 (no impairment) on 9 of the 11 ADAS-Cog subtests. In addition, the placebo group in this study did not exhibit any decline on the ADAS-Cog over the course of the study. In contrast, they actually “improved” on the ADAS-Cog over a 2-year period, despite worsening on the Clinical Dementia Rating scale (a widely used measure of functional status). This apparent practice effect was also reported in a 48-week trial of donepezil in MCI (Doody et al., 2009), where patients in the placebo arm improved slightly over the duration of the study on the ADAS-Cog.

In contrast to clinical trials research, clinical neuropsychological assessment of MCI typically involves the administration of multiple standardized measures of memory, language, attention, visuospatial, and other neurocognitive domains. Commonly used tests include measures known to be sensitive to MCI/mild AD, including measures of anterograde memory involving delayed recall and recognition, measures of semantic fluency, and measures of cognitive processing speed (De Jager, Hogervorst, Combrinck, & Budge, 2003; Devanand, Folz, Gorlyn, Moeller, & Stern, 1997; Howieson et al., 2008; Wilson, Leurgans, Boyle, & Bennett, 2011). The tests employed in routine assessments of this nature are, however, selected by the examining neuropsychologist and there is no “standard” battery that is administered across locations. In addition, these tests are for the most part not co-standardized or co-normed, and the interpretation of results requires a good deal of expertise. As these tests were typically designed for diagnostic purposes, and not for tracking/outcome measurement, most also lack alternate forms. Alternate forms are necessary to prevent practice effects, which can obscure underlying disease progression and prevent the detection of true decline. Finally, a typical clinical neuropsychological test battery used in the diagnosis of MCI does not generate a single global score, and global scores are useful for peer-to-peer communications about disease severity, for characterization of samples for research, and for tracking outcome in clinical trials.

The Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998) is a clinical tool that was specifically designed for both diagnostic purposes and for tracking outcome. One of the key design goals of the battery was to detect and characterize very mild dementia. The RBANS has the additional advantages of being relatively brief (~25 min) to administer, it has four equivalent alternate forms, and there are currently over 25 linguistically and culturally validated translations. It is also portable, requiring only a stimulus booklet and record form to administer, and it does not require the subject to be literate, increasing cross-cultural applicability and generalizability of results. The RBANS is currently widely used for clinical diagnostic purposes and has recently been deployed in multinational clinical trials of investigational compounds for AD. The RBANS generates index scores for five neurocognitive domains as well as a Total Scale Index score. These design features of the RBANS suggest that it may be a suitable tool for diagnosing and tracking patients with MCI.

Additional data supporting the potential utility of the RBANS in this context comes from several reports that patients with AD have a distinct profile of impairment across RBANS index scores, in contrast to dementias of other etiologies (Beatty et al., 2003; Randolph, Tierney, Mohr, & Chase, 1998). The utility of the RBANS in diagnosing AD dementia has been well-established (Duff et al., 2008), and the RBANS is predictive of functional capacity in patients with AD (Freilich & Hyer, 2007). The RBANS has also been reported to be predictive of functional capacity (i.e., driving) in patients with MCI (Badenes Guia, Casas Hernanz, Cejudo
Bolivar, & Aguilar Barbera, 2008). The RBANS correlates with all six domains of the Clinical Dementia Rating Scale in patients with MCI and AD (Hobson, Hall, Humphreys-Clark, Schrimsher, & O’Bryant, 2010), and the RBANS has been shown to correlate with AD biomarkers, including functional neuroimaging data in patients with AD (Forster et al., 2010; Wilson et al., 2009). Finally, the RBANS has been shown to be sensitive to detecting cognitive enhancement in computer training designed to boost memory functioning in non-demented older adults (Mahncke et al., 2006) and in a small clinical trial in MCI (Kotani et al., 2006). Therefore, it seems that the RBANS may meet many of the criteria necessary for a “gold standard” neurocognitive battery for diagnosis, tracking, and clinical trial outcome measure in MCI.

Although the utility of the RBANS in the diagnosis of AD is well-established, there are less data on the sensitivity/specificity of the RBANS in MCI. Duff, Hobson, Beglinger, and O’Bryant (2010) recently published a paper on this topic suggesting that the specificity of the RBANS was very good, but that the sensitivity was only moderate. These authors reported receiver operating characteristics (ROC) analyses with area under the curve (AUC) values of 0.78 for both the RBANS Delayed Memory Index (DMI) and the Total Scale score. Specificity ranged from 0.8 to 0.986 for various cutoffs on each of these measures, but sensitivity was no higher than 0.566 (with corresponding specificity of 0.829).

The subjects for the study by Duff were recruited from the community, however, and they were selected on the basis of performance on telephone screening tasks that was suggestive of MCI. They were subsequently classified as either MCI or “normal” on the basis of performance on the delayed recall trials of the Hopkins Verbal Learning Test-Revised (HVLT-R; Brandt & Benedict, 2001) and Brief Visuospatial Memory Test-Revised (BVMT-R; Benedict, 1997). If a subject obtained a performance at or below the 7th percentile (≥1.5 SD) of normal on the average of these two scores and also had a subjective complaint (either subject, collateral source, or both) of memory problems, that subject would be classified as MCI. If a subject scored above the 7th percentile on this score, that subject was classified as normal, regardless of whether or not that subject had memory complaints.

The statistical analyses in this study were carefully done and the sample sizes were fairly large (MCI = 72, controls = 71), but this method of subject selection does not reflect the way in which MCI is diagnosed clinically and introduces an ascertainment bias. None of the subjects had been referred for clinical evaluation of suspected memory loss and none of them were clinically diagnosed with MCI. Reliance upon a poor score composed of the average of just two subtest scores as a classification technique is questionable and also suggests the possibility of regression to the mean on other measures of memory, including the RBANS. In fact, the “MCI” group’s mean performance on the RBANS DMI (SS = 92.4) was in the normal low average range and does not suggest substantial impairment, despite being significantly below that of the “normal” group (SS = 101.4) in this study. Since the RBANS DMI is a composite score from four subtests and is scaled based upon population-based norming (unlike either the HVLT-R or BVMT-R), scores on the RBANS are more likely reflective of “true” memory ability. Other studies that have included clinic-based subjects have reported mean RBANS DMI scores that were more in line with expectations for a memory-disordered population (e.g., Hobson et al., 2010; Kotani et al., 2006).

The focus of the current study was to examine RBANS performance in a group of patients who were clinically diagnosed with MCI and to compare them with a group of subjects from the RBANS normative database, with precise demographic matching of the two groups. Although we are unable to track change using this cross-sectional approach, this approach was considered to be a more suitable methodology for examining the sensitivity/specificity of the RBANS in this patient population. We hypothesized that the RBANS scores in this clinic sample would be substantially lower than in the Duff and colleague study and more consistent with other studies that have reported on clinic-based samples.

Methods

Participants

Patients were identified from prospectively evaluated cases over 2 years from the pooled databases of two academically affiliated clinical centers. All patients were diagnosed with MCI due to AD (or “amnestic MCI” prior to the publication of the revised guidelines) on the basis of (a) observed decline by an informant in memory over the preceding year; (b) objective evidence of performance below expected premorbid levels in anterograde memory on standardized neuropsychological batteries in the judgment of the examining neuropsychologist (SK and CR); and (c) relative preservation of independence with respect to most activities of daily living. At each center, the typical standard of care also involved review of neuroimaging data and standard laboratories necessary to rule out structural, metabolic, or infectious etiologies for cognitive impairment, although those data were not collated for reporting here. This methodology differs somewhat from the clinical diagnostic approach used in some research samples, as informant-based observation was required (as opposed to relying upon subject self-report as equivalent), and there was no defined cutoff score for establishing “impairment.” In addition, the diagnosis was made by a clinical neuropsychologist, who ruled out other potential etiologies (e.g., depression, somatoform disorder) for the reported complaints and observed impairments.
While this approach is a little more stringent than required by the revised NIA-AA guidelines, this was the standard protocol for diagnosis at each center during this timeframe.

As evident in Table 1, this was a relatively high-functioning sample of patients, with mean general intellectual skills (as assessed by the Wechsler Abbreviated Scale of Intelligence [WASI] or the Short-Form of the Wechsler Scale of Intelligence [WAIS-III]) falling solidly within the average range. Performance on all other non-memory measures was also within normal limits (Table 1).

Healthy controls (HCs) were selected from the RBANS standardization sample and matched to the MCI patients on the basis of gender, race, age, and years of education. This study was approved by the institutional review board at New York University Langone Medical Center and Loyola University Medical Center.

Results

Group discrimination

All statistical analyses were conducted via SPSS 15.0 with nominal significance set at 0.05. The two groups did not differ significantly across any of the demographic variables (Table 1). Despite relatively intact general intellectual ability, patients with MCI performed significantly below the controls on all six RBANS index scores. Consistent with expectations, the largest difference was observed on the RBANS DMI, with a difference between the two group means of nearly 30 points (Table 2). This is in contrast to the 9-point difference reported by Duff and colleagues (2010) on this Index score.

The distributions of the RBANS Total Scale score for the HC and MCI groups were separately tested for normality, and both appear to be normally distributed (Kolmogorov–Smirnov and Shapiro–Wilk statistics both non-significant for both groups). The distributions of the RBANS Total Scale score by group are depicted in Figure 1.

Receiver operating curves were generated for both the RBANS DMI and the Total Scale Index score. The AUC for the DMI was 0.90 and for the Total Scale Index the AUC was 0.88 (Fig. 2).

Sensitivity and specificity were calculated for scores of 1.0 and 1.5 SD below the population-based means for the RBANS. Since the RBANS has population-based norms, this was seen as a more generalizable approach than simply relying upon the HC sample distributions in this study. Cutoffs were therefore set for index scores below 85 and below 78 for both the DMI score and the Total Scale Index score. The results including positive predictive and negative predictive values for each cutoff score are contained in Table 3.

Table 1. Participant demographics

<table>
<thead>
<tr>
<th></th>
<th>MCI</th>
<th>HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>81</td>
<td>81</td>
</tr>
<tr>
<td>Age (years)</td>
<td>77.3 (6.8)</td>
<td>76.8 (6.4)</td>
</tr>
<tr>
<td>Gender (F:M)</td>
<td>47:34</td>
<td>47:34</td>
</tr>
<tr>
<td>Yrs Ed</td>
<td>15.4 (2.7)</td>
<td>15.4 (3.0)</td>
</tr>
</tbody>
</table>

Neuropsychological tests

<table>
<thead>
<tr>
<th>Test</th>
<th>MCI</th>
<th>HC</th>
</tr>
</thead>
<tbody>
<tr>
<td>FSIQ</td>
<td>106.5 (13.6)</td>
<td>–</td>
</tr>
<tr>
<td>WTAR Predicted FSIQ</td>
<td>108.5 (9.6)</td>
<td>–</td>
</tr>
<tr>
<td>Rey Complex Figure Test-Copy (n = 42)</td>
<td>30.0 (5.8)</td>
<td>–</td>
</tr>
<tr>
<td>BNT</td>
<td>46.6 (8.2)</td>
<td>–</td>
</tr>
<tr>
<td>COWAT</td>
<td>34.5 (12.1)</td>
<td>–</td>
</tr>
<tr>
<td>VCAT (n = 39)</td>
<td>14.03 (4.5)</td>
<td>–</td>
</tr>
<tr>
<td>GDS</td>
<td>7.9 (4.9)</td>
<td>–</td>
</tr>
<tr>
<td>DRS-2 Attention (n = 37)</td>
<td>36.3 (15.2)</td>
<td>–</td>
</tr>
<tr>
<td>DRS-2 Initiation/Perseveration (n = 37)</td>
<td>33.0 (4.7)</td>
<td>–</td>
</tr>
<tr>
<td>DRS-2 Construction (n = 37)</td>
<td>5.78 (0.5)</td>
<td>–</td>
</tr>
<tr>
<td>DRS-2 Conceptualization (n = 37)</td>
<td>37.16 (2.3)</td>
<td>–</td>
</tr>
<tr>
<td>DRS-2 Memory (n = 37)</td>
<td>20.2 (4.2)</td>
<td>–</td>
</tr>
<tr>
<td>DRS-2 Total Score (n = 37)</td>
<td>131.0 (10.3)</td>
<td>–</td>
</tr>
</tbody>
</table>

Notes: MCI = mild cognitive impairment; HC = healthy control; Yrs Ed = years of education; WTAR = Wechsler Test of Adult Reading; BNT = Boston Naming Test; COWAT = Controlled Oral Word Association Test; VCAT = Verbal Concept Attainment Test; GDS = Geriatric Depression Scale; DRS-2 = Dementia Rating Scale-2; FSIQ = full-scale IQ. Numeric data are the means (SD).
The RBANS has also been used to profile cognitive impairment in dementia, differentiating between the “cortical” profile of AD and the “subcortical” profile of Huntington’s disease, Parkinson’s disease, and vascular dementia (e.g., Beatty et al., 2003; Randolph, 1998; Randolph et al., 1998). A subcortical–cortical deviation score is calculated by taking the mean of the Visuospatial-Constructional and Attention Index scores and subtracting the mean of the Language and DMI scores. Positive values reflect a “cortical” or more AD-like pattern of performance.

Subcortical–cortical deviation scores were calculated for all subjects, and the groups were compared via one-way ANOVAs. The MCI group had significantly higher deviation scores than the HC group (\(F = 8.3, p < .005\)). This suggests that the MCI group had an AD-like pattern of impairments.

Discussion

The present study involved a retrospective review of RBANS data from a sample of clinically diagnosed patients from two academically affiliated practice settings. RBANS index scores from this sample were compared with scores from a HC sample extracted from the RBANS normative database to match the clinical sample on relevant demographic characteristics. The RBANS proved to be sensitive to a full range of performance in both groups, discriminated between the groups at a fairly high rate of accuracy, and the MCI subjects exhibited a profile that was similar to that observed in AD subjects. These findings further establish the utility of the RBANS as a suitable instrument for diagnostic evaluations of patients with suspected MCI.

The RBANS also has four equivalent forms, allowing for repeat testing over fairly short time intervals without incurring significant practice effects. In conjunction with the brevity of the test, this is a feature that is valuable and necessary for clinical trial applications. The clinical validity of the test with respect to the assessment of AD and other dementing illnesses has been...
well-established, and there is no other standardized neurocognitive battery with population-based norms that has these design features (e.g., brief, multiple fully-alternate forms, designed to detect and characterize mild dementia).

In addition to being able to discriminate between patients with clinically diagnosed MCI and HCs; however, a battery useful for both clinical diagnostic and clinical trial purposes should prove to have some utility in identifying individuals who are truly in the prodromal phase of AD. Future studies with the RBANS and other candidate instruments for this purpose should involve exploration of the extent to which global performance or specific patterns of performance are predictive of subsequent cognitive decline and conversion to AD.

As noted earlier, preliminary data comparing various neuropsychological measures with biomarkers in such predictions have tended to favor the neuropsychological measures. This may not be all that surprising when considering the results of some recent data from large community-based prospective studies that have involved clinical—pathological correlations, such as the Rush Memory and Aging Project and Religious Orders Study. These large longitudinal studies have yielded data, indicating that a high percentage of individuals who are cognitively normal at death actually meet neuropathological criteria for AD and that the percentage of individuals diagnosed with AD who actually have “pure” AD on autopsy is not much different, in the range of 30%—40% for each group (Bennett et al., 2006; Schneider, Arvanitakis, Bang, & Bennett, 2007). This does not bode well for biomarker prediction of cognitive decline and probably explains why studies of CSF markers and brain amyloid imaging typically report a great deal of overlap between groups of subjects who are clinically classified as cognitively normal, MCI, and AD.

Limitations of this study include the fact that the RBANS was part of the larger neuropsychological battery used in part to make the diagnosis of MCI. In each center, however, the RBANS made up no more than 1/3 of the neurocognitive test session (in terms of administration time), and additional measures of attention, language, memory, visuospatial, and executive functions were given during these diagnostic evaluations. Although the RBANS results might therefore have influenced the diagnosis to some extent, there were a number of other measures of the same neurocognitive domains that were also relied upon for diagnostic purposes.

Table 3. Sensitivity and specificity for cutoff scores

<table>
<thead>
<tr>
<th>Score</th>
<th>Cutoff</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>PPV (%)</th>
<th>NPV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delayed Memory Index</td>
<td>&lt;85</td>
<td>71.6</td>
<td>89.9</td>
<td>86.6</td>
<td>75.8</td>
</tr>
<tr>
<td>Delayed Memory Index</td>
<td>&lt;78</td>
<td>55</td>
<td>95.6</td>
<td>91.8</td>
<td>68.1</td>
</tr>
<tr>
<td>Total Scale Index</td>
<td>&lt;85</td>
<td>64.2</td>
<td>90.1</td>
<td>86.7</td>
<td>71.6</td>
</tr>
<tr>
<td>Total Scale Index</td>
<td>&lt;78</td>
<td>32.1</td>
<td>96.3</td>
<td>89.7</td>
<td>58.6</td>
</tr>
</tbody>
</table>

Notes: PPV = positive predictive value; NPV = negative predictive value.
These varied to some extent across the two centers, and since our control sample came from the RBANS standardization data set (which did not include any other measures given to the normative sample), we could not apply ROC analyses or examine sensitivity/specificity for any measures other than the RBANS. We also currently lack sufficient follow-up data to explore the predictive value of the RBANS in identifying which MCI patients in this sample worsened over time or converted to AD, although those data are currently being collected.

Overall, for a relatively brief battery, the RBANS demonstrated good sensitivity and specificity in the identification of patients with MCI, and the magnitude of the difference between the control and MCI groups was substantial (e.g., nearly 30 points, or 2 SD, on the DMI), despite the fact that the MCI group still had a mean full-scale IQ in the normal above-average range. The psychometrics of the RBANS is well-established, as is the clinical validity of the test in the assessment of AD. It would be advantageous for the field of neuropsychology to establish a “gold standard” for the neurocognitive assessment of MCI, at least for research applications, and the RBANS appears to be a viable candidate for this purpose, although further research is necessary. The existence of such a neurocognitive gold standard battery would be useful in challenging the incremental utility of biomarker data in diagnosis/prognosis involving this patient group and to serve as a benchmark against which to measure alternative measurement tools. The direct comparison of various candidate instruments is particularly important in determining the relative utility of a given instrument. Differences in sample selection and control groups across studies can result in different conclusions regarding an instrument’s sensitivity to impairments, ability to profile deficits across different disorders, and ability to predict functional capacity. This is likely the source of the differences noted between the results of the present study and those reported by Duff and colleagues (2010). Additional research should involve head-to-head comparisons of the RBANS with other candidate batteries, using a common control group. The addition of measures of functional instrumental activities of daily living scales and even biomarker measurements would be useful in determining the relative utility of different assessments in differential diagnosis and prediction of functional impairments.

Funding

No portions of this research were funded by an outside organization or government agency.

Conflict of Interest

CR is the author of the RBANS and receives royalty payments from the publisher and copyright holder of the test, Pearson, Inc.

References


