BMC
Neurology

ORIGINAL PAPER

Open Access

Brief International Cognitive Assessment for MS (BICAMS): international standards for validation

Ralph HB Benedict[*], Maria Pia Amato, Jan Boringa, Bruno Brochet, Fred Foley, Stan Fredrikson, Paivi Hamalainen, Hans Hartung, Lauren Krupp, Iris Penner, Anthony T Reder and Dawn Langdon

Abstract

An international expert consensus committee recently recommended a brief battery of tests for cognitive evaluation in multiple sclerosis. The Brief International Cognitive Assessment for MS (BICAMS) battery includes tests of mental processing speed and memory. Recognizing that resources for validation will vary internationally, the committee identified validation priorities, to facilitate international acceptance of BICAMS. Practical matters pertaining to implementation across different languages and countries were discussed. Five steps to achieve optimal psychometric validation were proposed. In Step 1, test stimuli should be standardized for the target culture or language under consideration. In Step 2, examiner instructions must be standardized and translated, including all information from manuals necessary for administration and interpretation. In Step 3, samples of at least 65 healthy persons should be studied for normalization, matched to patients on demographics such as age, gender and education. The objective of Step 4 is test-retest reliability, which can be investigated in a small sample of MS and/or healthy volunteers over 1–3 weeks. Finally, in Step 5, criterion validity should be established by comparing MS and healthy controls. At this time, preliminary studies are underway in a number of countries as we move forward with this international assessment tool for cognition in MS.

Background

Multiple sclerosis (MS) is an inflammatory disease of the central nervous system, causing demyelination and neuro-degeneration in most patients [1,2]. As would be expected in such a disease with prominent cerebral pathology, a substantial number [3-5] of MS patients are compromised neuropsychologically. In recently diagnosed or benign course patients, the incidence of cognitive impairment ranges from 20-40% [5,6]. In clinic based samples where secondary progressive course is more common, roughly 50-60% of patients are affected [4].

Neuropsychological (NP) testing provides quantification of cognition, and is used clinically to diagnose impairment and to inform medical and behavioral treatment decisions [7]. Two descriptors, psychometric or neuropsychological tests, are often used inter-changeably to describe the cognitive testing procedures used with MS patients [8]. *Psychometric tests* are standardized, behavioral measures of mental phenomena. They measure many domains of

mental function, including psychomotor speed and dexterity, personality or psychopathology [via standardized questionnaires or surveys], intelligence, memory and other aspects of cognitive processing [eg attention, language, executive function].

The term *neuropsychological test* conveys the idea that the psychometric test result is relevant for conclusions pertaining to cerebral function. Neuropsychological tests are used to examine brain-injured patients or to study hypotheses in neuroscience. A deficient neuropsychological test value is often judged to be indicative of cerebral dysfunction. The Wechsler Adult Intelligence Scale (WAIS) [9], is a classic example of a psychometric test of intelligence. It has carefully standardized instructions, scoring criteria, extensive age-based normative data, and information derived from extensive research concerning reliability and validity. It can also be construed as a neuropsychological test [especially nonverbal components] because there are extensive data that show its relationship with cognitive aging, dementia, and other changes in cerebral status.

Psychometric data regarding normal performance, test reliability and the validity of test interpretation are

* Correspondence: benedict@buffalo.edu
Buffalo General Medical Center, Department of Neurology, Suite E2, 100 High Street, Buffalo, NY 14203

necessary for accurate application of NP testing [10]. Most tests in common use are carefully standardized such that the same instructions, stimuli and marking criteria are used by all examiners. Most often, high test-retest reliability is emphasized in order to avoid error in repeat testing circumstances. Unfortunately, alternate test versions and normative data are not available for some NP tests in all languages and cultures. In addition, many centers lack expertise in psychometrics and NP test interpretation.

The Brief International Cognitive Assessment for MS (BICAMS) initiative was undertaken to recommend a brief, cognitive assessment for MS that is optimized for small centers, with perhaps one or few staff members, who may not have NP training [11]. BICAMS was particularly focused on international use, to facilitate comparison across settings. An expert committee of twelve neurologists and neuropsychologists representing the main cultural groups that have so far contributed extensive data about cognitive dysfunction in MS was convened. The opinions generated from the meeting are published elsewhere [11]. In brief, the panel recommended one particular test with high reliability and good sensitivity, the Rao [12] adaptation of the Symbol Digit Modalities Test (SDMT) [13]. Consensus was also achieved on optimal measures for learning and memory in MS patients, time permitting: the initial learning trials of the second edition of the California Verbal Learning Test (CVLT2) [14] and the revised Brief Visuospatial Memory Test (BVMTR) [15].

In order to facilitate international implementation of the BICAMS assessment, multiple translations are needed, as well as psychometric research to insure the reliability and validity of new test forms. With this in mind, a second conference was held to develop consensus on a BICAMS validation protocol.

Brief International Cognitive Assessment for MS (BICAMS) description

The SDMT [13] presents a series of nine symbols, each paired with a single digit in a key at the top of a standard sheet of paper. An adapted version of the test is presented in Figure 1. Patients are asked to voice the digit associated with each symbol as rapidly as possible for 90 sec. There is a single outcome measure – the number correct over the 90 sec time span.

The auditory/verbal learning test is the CVLT2 [16]. The test begins with the examiner reading a list of 16 words [Figure 2]. Patients listen to the list and report as many of the items as possible. There is no instruction as to the order in which items are recalled. After recall is recorded, the entire list is read again followed by a second attempt at recall. Altogether, there are five learning trials. The reader will note that the 16-item list [see faux example in Figure 2] has words that conform to four semantic categories, in this case sports, vegetables, clothes, and tools. Some subjects will recall items in a grouped fashion, and others may recall the list in serial order. There are many variables of recall available in the CVLT2, as a second list is presented, and after 25 min there is a delayed recall trial as well as a yes/no recognition memory task. The BICAMS panel noted that few studies have shown incremental
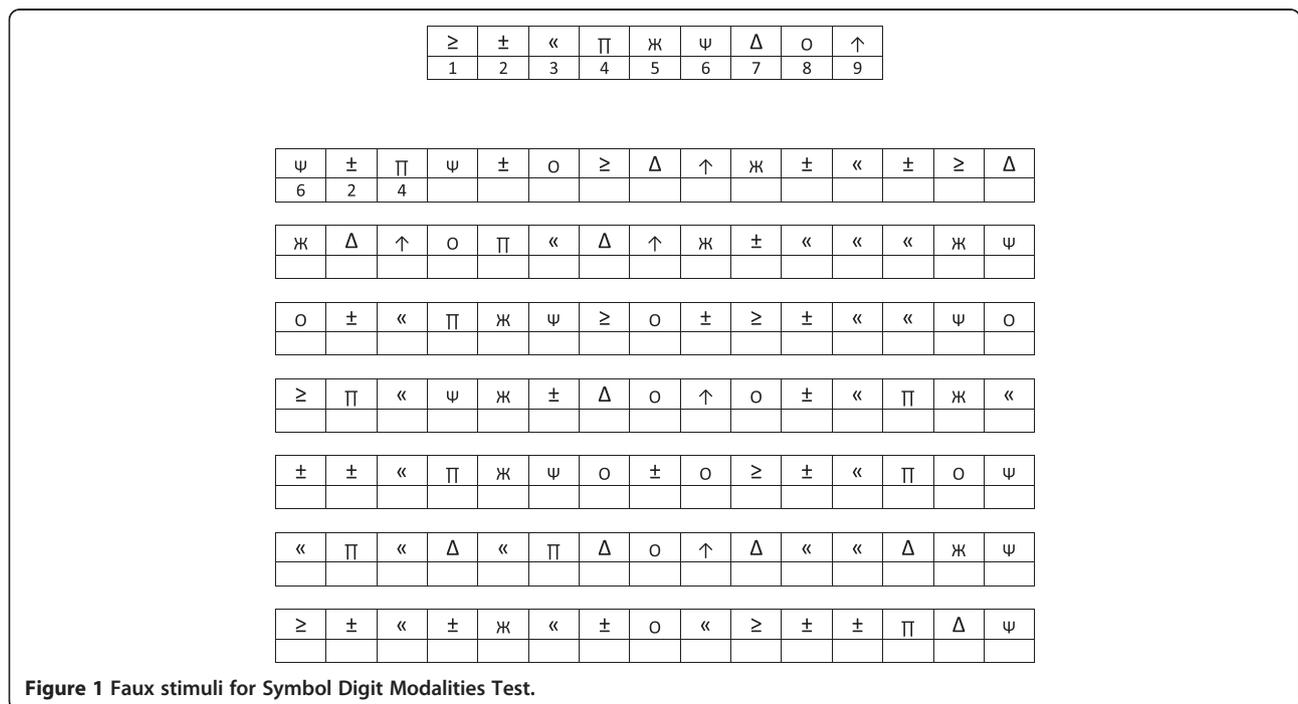
Figure 1 Faux stimuli for Symbol Digit Modalities Test.

	Trial 1 Recall	Trial 2 Recall	Trial 3 Recall	Trial 4 Recall	Trial 5 Recall
football					
notebook					
island					
billiards					
paper					
river					
tennis					
cake					
folder					
boxing					
mountain					
pie					
candy					
envelope					
valley					
ice cream					
	__/16	__/16	__/16	__/16	__/16

Total Learning   ___/80]

Figure 2 Faux stimuli for California Verbal Learning Test Second Edition.

validity with these measures, as the total number of recalled items over the five learning trials is most sensitive [17].

Visual/spatial memory is assessed in BICAMS using the BVMTR [18]. In this test, six abstract designs [Figure 3] are presented for 10 sec. The display is removed from view and patients render the stimuli via pencil on paper manual responses. Each design receives from 0 to 2 points representing accuracy and location. Thus, scores range from 0 to 12. There are three learning trials, and the primary outcome measure is the total number of points earned over the three learning trials. Because there is little evidence that the delayed recall trial adds to discriminant validity in MS [4,19], as in the MATRICS consensus battery [20], only the initial learning trials are recommended for BICAMS.

Conference process

Having already decided on the core [SDMT] and ancillary memory tests [CVLT2, BVMTR], the focus of the conference was on a BICAMS psychometric validation protocol. The committee reviewed basic psychometric standards from the literature [10,21-26] that are widely accepted for validation of behavioral or psychological outcome measures. Recognizing that economic resources for validation will vary across country and region, the committee discussed priorities for validation, that is, those aspects of research design that would enable empirical appraisal of core psychometrics that may engender confidence and wide application of BICAMS. Following consensus on these central components, it was noted that most of the psychometric evidence relating to the BICAMS tests relies

on US samples and discussion shifted to practical matters pertaining to implementation across different languages and countries.

Figure 3 Faux stimuli for Brief Visuospatial Memory Test Revised.
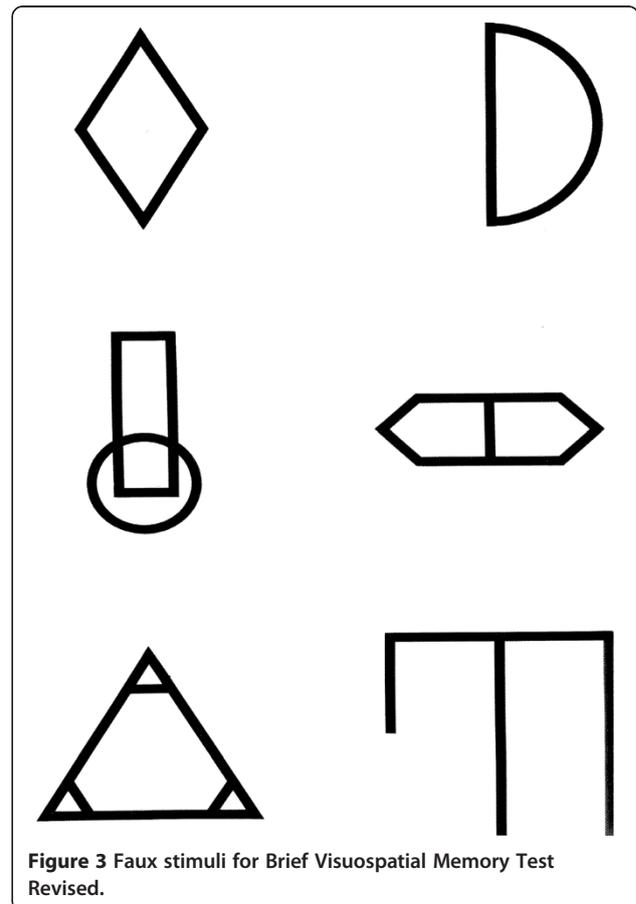
Psychometric standards
Standardization
The first step in the development of any test is to ensure that test stimuli and administration procedures have good face validity and consistent stimulus presentation. While this can be a painstaking process during the initial phases of psychometric test development, the work is already done for the tests that were selected for BICAMS. The SDMT, CVLT2 and BVMTR stimuli are well established and are readily mass produced using digital and print technology. Examiners can be easily trained to use standardized instructions, more or less verbatim, to enhance reliability across settings. The BICAMS tests are adequate in this regard.

Normalization of raw scores
Normative data are of course essential for the clinical application of psychometric, neuropsychological tests. Acquiring normative data can be an expensive endeavor. For example, the recently revised US normative data for BVMTR has a sample size of 588 [27]. These data were fairly recently acquired and are judged by the BICAMS committee to be current.

Normalization data for SDMT are more complicated. The manual based norms date to 1982, raising the spectre of cohort effects [ie gradual shift upward over time] or poor generalizablity to patients and controls in the present day. Benedict et al published normative data on the MAC-FIMS battery, which includes the SDMT, in 2006 [4] and 2010 [28], using US healthy samples numbering 56 and 120 respectively. These are controlled research studies with applicability largely restricted to clinicians treating MS patients. A potential problem is that normative data from one region [North Eastern USA in this case] may result in interpretive error when applied to raw test scores derived from a different culture, language, region or country.

Reliability
By reliability we mean the degree to which there is error when using the same instrument across settings, examiners, etc. It is perhaps the most critical psychometric criterion - if the test is unreliable, there is little confidence in the validity of the outcome. Of the various forms of reliability, the panel decided that test-retest reliability has the highest priority and is most relevant for future BICAMS validation. The coefficient of variation can be used in very small samples to determine the extent to which changes in mean values outweigh the variance in test scores [29]. A more valid measure of test-retest reliability is the Pearson correlation coefficient [30]. Most commonly, a brief test-retest interval ranging from 1–3 weeks is employed. For most purposes, r values for test-retest correlation are considered adequate if >0.70 and good if >0.80 [23].

The SDMT has particularly high test-retest reliability. In one US study of MS patients the test-retest r value was 0.97. In a US study repeating the test over six monthly sessions, r values approximated 0.80 for healthy controls and 0.90 for MS patients. Acceptable test-retest reliability [CVLT2 = 0.78; BVMTR = 0.91] was found in a well controlled investigation with US MS patients [31].

Validity
Rather than accuracy, validity refers to the meaning of a test score. There are multiple aspects of validity in psychometric science. Does a low score, for example, represent the presence of neuropsychological dysfunction, a particular disease state such as MS, a high likelihood of brain atrophy or some other marker of cerebral involvement, or perhaps an increased risk of failing at work? Correlations between test scores and other measures [ie validity coefficient r] and comparing differences in the performance from specified samples [eg, MS vs controls; employed vs disabled MS patients] are common methods of investigation.

The BICAMS committee decided that the most important aspect of validity for clinical purposes is criterion-related validity, most notably differentiating MS patients from healthy controls. All of the BICAMS tests discriminate well with SDMT most often the most sensitive measure in NP batteries [4,19].

A more difficult endeavor is to establish the ecological or predictive validity of a psychometric test [32]. Neuropsychological testing is correlated with a wide range of activities of daily living in MS [33-40], as well as work disability [41-44]. The BICAMS tests are correlated with vocational outcomes [4,45-47] and recently job loss was associated with specific decline of 4–5 points on SDMT [48]. More such research is needed in order to clearly establish thresholds for clinically meaningful changes for the BICAMS measures.

Alternate test forms
When NP tests are repeated in healthy volunteers or stable MS patients, performance often improves for two reasons: item-specific or task-specific learning [49]. The former refers to the learning of, or memory for, specific test stimuli. For example, on the CVLT2, one might remember specific words from one testing session to the next. Indeed, on a similar test, the 1st trial recall was similar to the delayed recall trial of the same test administered two weeks earlier, but only when the same list was repeated [49]. Test- or task-specific learning refers to the benefit of performing the same behavioral procedure successively, even if the to-be-remembered stimuli are altered. For example, one could possibly learn to label BVMTR figures verbally, a strategy that may carry over to the next session, despite presenting different

figures. While both item- and task-specific practice likely play a role in retest effects, in MS, we [31] have found that changing stimuli reduces practice effects on memory tests such as the CVLT2 and BVMTR.

The SDMT was originally published with one test form. Alternate forms were generated by Rao and colleagues [3,12], but in the only work examining inter-form equivalence [50], there was little support for the forms being equivalent. Recently, two new forms were created and found to be equivalent to the standard form [51]. There are two forms for the CVLT2. In the CVLT2 test manual, the normative data are very similar for each form, and the forms yielded similar data in a test-retest within-subjects design [31]. There is strong support for inter-form reliability for the BVMTR [20,31,49,52-54].

Consensus opinion: the BICAMS validation protocol

The above list of psychometric criteria is not intended to be comprehensive, but the discussed items are essential in the test development process. As we move forward with implementing BICAMS internationally, each of these criteria may be difficult to achieve in other languages and cultures. In this section, we describe the suggested, core validation process, highlighting special considerations for each of the BICAMS measures.

The BICAMS tests were selected, in part, due to extant validation findings and thus it is not surprising that they hold up well to psychometric scrutiny. Table 1 summarizes what we know about the current English versions of BICAMS. Note that the tests are good on the most primary criteria, involving standardization, normalization, test-retest reliability and criterion-related validity. More variable data are available pertaining to alternate forms and predicting clinically meaning changes over time.

Looking forward, as summarized in **Appendix 1** below, we envisage five steps in future validation protocols in populations for whom English is not the first language.

Preliminary work in Step 1 will be needed to maximize standardization while remaining true to the meaning of the original version, where possible. The extant SDMT stimuli are deemed adequate for international use, at least for cultures where Arabic numerals are in common use.

One consideration is the pronounciation of numbers which may vary from monosyllabic to polysyllabic utterances [eg one in English and nueve in Spanish], or be simply longer in others [eg üheksa in Estonian]. Rarely, the meaning of the SDMT and BVMTR symbols could become important. These stimuli have little semantic meaning in English but could conceivably have meaning in some cultures. Like the SDMT stimuli, the BVMTR test stimuli are adequate for international application.

The CVLT2 is of course entirely another matter. Here, precise translation necessitates as close approximation of the English words as possible, while maintaining word frequency in the target language, semantic relationships among the target words, orthography, and alike. As noted above, in some languages this could mean a very arduous process. The BICAMS committee agreed that in some countries, another, simple, auditory word-list learning test could replace the CVLT2, provided that the procedure is in the common format – that is reading the list on each learning trial and including at least three learning trials. Some English language examples are the Rey Auditory Verbal Learning Test [55] and the revised Hopkins Verbal Learning Test [56,57].

All BICAMS tests must re-standardize the administration and scoring instructions in the new language. The time required for Step 2 will depend on the specific test and technical support available. The patient instructions for SDMT are brief, and there is minimal instruction necessary for scoring the test in the standardized manner. In contrast, the scoring aspects of the BVMTR manual are quite detailed. Patient responses could be delivered to another party and scored blindly in lieu of translating the entire scoring sections of the manual.

In Step 3, a sample of at least 65 healthy volunteers must be studied with the new BICAMS to develop normative data in the native language. This minimum sample size should provide enough power to detect a medium effect size in a two-group [eg MS vs controls] comparison. Unless a larger sample is available, the normalization sample should be group matched to population studies of MS patients in terms of demographic characteristics. Linear regression approaches can be employed to extend the applicability of the data to demographics that are not fully represented in the database.

Table 1 Manner in which BICAMS measures meet psychometric criteria in samples with English as a first language

	SDMT	CVLT2	BVMTR
Standardization	Smith 1982 Teat Manual	Delis 2000 Test Manual	Benedict 1997 Test Manual
Normalization	Parmenter 2010. Peer review journal article	Delis 2000 Test Manual	Benedict 2005 Test Manual
Reliability I: Test-Retest	Benedict 2005 r = 0.91	Benedict 2005 r = 0.80	Benedict 2005 r = 0.91
Reliability II: Alternate Form	Rao 1991. Benedict 2012. Good	Delis 2000. Fair	Benedict 1996. Good
Validity I: Criterion Related	Many Studies. Good	Many Studies. Good	Many Studies. Good
Validity II: Clinically Meaningful Change	Morrow 2010. Fair	No or Little Data	No or Little Data

Test-retest reliability can be assessed in both patients and controls by a repeat testing session 1–3 weeks after baseline [Step 4]. While both samples are of interest, the panel believes that reliability in MS is more important than in healthy volunteers. In order to assess criterion-related validity [Step 5], the controls must be compared to MS patients, with control for demographics. If Steps 1 and 2 have been completed effectively, all BICAMS variables should discriminate the groups significantly, with d values >0.5. Studies to determine the validity of BICAMS in distinguishing MS populations from healthy controls will need to be carefully constructed, because the criteria of diagnosis of MS adopted may vary among countries [58]. It may be necessary for published data to be segmented to allow comparison with MS samples from other language groups and the BICAMS committee will facilitate this wherever possible.

Finally, other psychometric considerations include inter-rater reliability, alternate forms, and various forms of convergent and discriminant validity. These are not deemed essential, but potentially valuable. For example, does BICAMS predict vocational outcomes? Are low BICAMS scores associated with brain atrophy? Alternate form reliability work has already been commenced by some members of the BICAMS committee.

Conclusion

The Brief International Cognitive Assessment for MS (BICAMS) initiative was undertaken to recommend a brief, cognitive assessment for MS that can be utilized internationally, in small centers, with perhaps one or few staff members, who may or may not have formal neuropsychological training. Consensus was earlier achieved regarding the BICAMS tests, with special consideration for SDMT, and supplementation by CVLT2 and BVMTR, time permitting. Research is needed to validate BICAMS where English is not the first language. In this article, we have summarized a second consensus opinion which offers a process by which BICAMS can be validated in other languages. Research projects pursuing some of the aims described herein are underway.

Appendix 1 Recommended Step-by-Step Protocol for BICAMS Validation

▪Step 1, Standardization and Translation of Test Stimuli. For visual stimuli, determine if there are any semantic associations to stimuli in the culture or language under consideration. For CVLT2 must match new words on word frequency and appropriate similarity of meaning. If these parameters cannot be applied scientifically, then expert review and performance on test by appropriate participants will be utilized to assess translation.

▪Step 2, Standardization and Translation of Test Instructions. All information from the test manual necessary for administration and interpretation must be translated, back translated, and checked for errors. Where possible the translated instructions should be validated against expected participant performance in terms of accuracy and error profile. Step 4 will also contribute to the accuracy of the test instructions.

▪Step 3, Normalization. Large samples of 150 or more healthy persons are needed for data applicable to persons of all ages and diverse ethnicity. The minimum sample size is 65 healthy volunteers, provided they are group matched on demographics to either a concurrent MS sample, or matched to samples in other published descriptive MS studies. Where possible, the distribution of test scores and error profile of the normalization sample should be examined and compared to published distributions from other language groups.

▪Step 4, Test-Retest Reliability. Assessment of this criterion can be achieved by evaluating an MS and/or healthy volunteer sample on two occasions separated by 1–3 weeks. This is the gold standard separation where the question is only test reliability, controlling for maturation effects. A Pearson's correlation coefficient >0.70 will usually be required.

▪Step 5, Criterion-Related Validity. This step can be pursued in conjunction with Step 3, in that an MS sample can be compared to a healthy control group that also serves for normalization. To determine if a new Italian BVMTR is sensitive to MS disease state, for example, compare 50 patients to the healthy controls in Step 3. After the study, the investigator adds another 35 healthy volunteers to round out the normalization sample.

References
1. Hauser SL, Oksenberg JR, Hauser SL, Oksenberg JR: **The neurobiology of multiple sclerosis: genes, inflammation, and neurodegeneration.** *Neuron* 2006, **52**:61–76.
2. Trapp BD, Nave KA: **Multiple sclerosis: an immune or neurodegenerative disorder?** *Annu Rev Neurosci* 2008, **31**:247–269.
3. Rao SM, Leo GJ, Bernardin L, Unverzagt F: **Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction.** *Neurology* 1991, **41**:685–691.
4. Benedict RHB, Cookfair D, Gavett R, *et al*: **Validity of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS).** *J Int Neuropsychol Soc* 2006, **12**:549–558.
5. Patti F, Amato M, Trojano M, *et al*: **Cognitive impairment and its relation with disease measures in mildly disabled patients with relapsing-remitting multiple sclerosis: baseline results from the Cognitive Impairment in Multiple Sclerosis (COGIMUS) study.** *Mult Scler* 2009, **15**:779–788.
6. Glanz BI, Holland CM, Gauthier SA, *et al*: **Cognitive dysfunction in patients with clinically isolated syndromes or newly diagnosed multiple sclerosis.** *Mult Scler* 2007, **13**:1004–1010.
7. Benedict RHB: *Integrating cognitive function screening and assessment into the routine care of multiple sclerosis patients.*: CNS Spectrums; 2005. in press.
8. Benedict RH, Zivadinov R: **Risk factors for and management of cognitive dysfunction in multiple sclerosis.** *Nat Rev Neurol* 2011.
9. Wechsler D: *Wechsler Adult Intelligence Scale.* 4th edition. New York: Pearson; 2008.
10. Lezak M: *Neuropsychological Assessment.* 2nd edition. New York: Oxford University Press; 1995.
11. Langdon DW, Amato MP, Boringa J, *et al*: **Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS).** *Multiple Sclerosis Journal*, in press.
12. Rao SM: *A Manual for the Brief, Repeatable Battery of Neuropsychological Tests in Multiple Sclerosis: National Multiple Sclerosis Society*; 1991.
13. Smith A: *Symbol digit modalities test: Manual.* Los Angeles: Western Psychological Services; 1982.
14. Delis DC, Kramer JH, Kaplan E, Ober BA: *California Verbal Learning Test Manual: Second Edition, Adult Version.* San Antonio, TX: Psychological Corporation; 2000.
15. Benedict RHB: *Brief Visuospatial Memory Test - Revised: Professional Manual.* Odessa, Floriday: Psychological Assessment Resources, Inc.; 1997.
16. Benedict RHB, Schretlen DS, Groninger L, Dobraski M. *Revision of the brief Visuospatial Memory Test: Studies of normal performance, reliability and validity.* Psychological Assessment 1996, **8**:145–153.
17. Stegen S, Stepanov I, Cookfair D, *et al*: **Validity of the California Verbal Learning Test-II in Multiple Sclerosis.** *Clin Neuropsychol* 2010, **24**:189–202.
18. Benedict R: *Brief Visuospatial Memory Test-Revised: Professional manual.* Odessa, Florida: Psychological Assessment Resources; 1997.
19. Strober L, Englert J, Munschauer F, Weinstock-Guttman B, Rao S, Benedict RH: **Sensitivity of conventional memory tests in multiple sclerosis: comparing the Rao Brief Repeatable Neuropsychological Battery and the Minimal Assessment of Cognitive Function in MS.** *Mult Scler* 2009, **15**:1077–1084.
20. Buchanan RW, Davis M, Goff D, *et al*: **A Summary of the FDA-NIMH-MATRICS Workshop on Clinical Trial Design for Neurocognitive Drugs for Schizophrenia.** *Schizophr Bull* 2005, **31**:5–19.
21. Benton AL, Sivan AB, Hamsher K, Varney NR, Spreen O: *Contributions to Neuropsychological Assessment.* 2nd edition. New York: Oxford University Press; 1994.
22. Spreen O, Strauss E: *A compendium of neuropsychological tests: administration, norms, and commentary.* 2nd edition. New York: Oxford University Press; 1998.
23. Anastasi A: *Psychological Testing.* 6th edition.: Macmillan Publishing Company; 1988.
24. Benedict RHB, Fischer JS, Archibald CJ, *et al*: **Minimal neuropsychological assessment of ms patients: a consensus approach.** *Clin Neuropsychol* 2002, **16**:381–397.
25. Amato MP, Zipoli V, Portaccio E: **Mulitple-sclerosis related cognitive changes: a review of cross-sectional and longitudinal studies.** *J Neurol Sci* 2006, **245**:41–46.
26. Camp SJ, Thompson AJ, Langdon DW: **A new test of memory for multiple sclerosis I: format development and stimuli design.** *Mult Scler* 2001, **7**:255–262.
27. Benedict RHB, Brandt J: *Hopkins Verbal Learning Test Revised and Brief Visuospatial Memory Test Revised: Professional Manual Supplement.* Odessa, Fl: Psychological Assessment Resources, Inc; 2007.
28. Parmenter BA, Testa SM, Schretlen DJ, Weinstock-Guttman B, Benedict RHB: **The utility of regression-based norms in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS).** *J Int Neuropsychol Soc* 2010, **16**:6–16.
29. Horsfield MA, Rovaris M, Rocca MA, *et al*: **Whole-brain atrophy in multiple sclerosis measured by two segmentation processes from various MRI sequences.** *J Neurol Sci* 2003, **216**:169–177.
30. Cohen J, Cohen P: *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* 2nd edition. Hillsdale, NJ: Lawrence Erlbaum; 1983.
31. Benedict RHB: **Effects of using same vs. alternate form memory tests in short-interval, repeated assessment in multiple sclerosis.** *J Int Neuropsychol Soc* 2005, **11**:727–736.
32. King MT: **A point of minimal important difference (MID): a critique of terminology and methods.** *Expert Rev Pharmacoecon Outcomes Res* 2011, **11**:171–184.
33. Langdon DW, Thompson AJ: **Multiple sclerosis: a preliminary study of selected variables affecting rehabilitation outcome.** *Mult Scler* 1999, **5**:94–100.
34. Amato MP, Ponziani G, Pracucci G, Bracco L, Siracusa G, Amaducci L: **Cognitive impairment in early-onset multiple sclerosis: Pattern, predictors, and impact on everyday life in a 4-year follow-up.** *Arch Neurol* 1995, **52**:168–172.
35. Jean VM, Paul RH, Beatty WW: **Psychological and neuropsychological predictors of coping patterns by patients with multiple sclerosis.** *J Clin Psychol* 1999, **55**:21–26.
36. Schultheis MT, Garay E, Millis SR, DeLuca J: **Motor vehicle crashes and violations among drivers with multiple sclerosis.** *Arch Phys Med Rehabil* 2002, **83**:1175–1178.
37. Goverover Y, Kalmar J, Gaudino-Goering E, *et al*: **The relation between subjective and objective measures of everyday life activities in persons with multiple sclerosis.** *Arch Phys Med Rehabil* 2005, **86**:2303–2308.
38. Nortvedt MW, Riise T, Myhr KM, Nyland HI: **Quality of life in multiple sclerosis: measuring the disease effects more broadly.** *Neurology* 1999, **53**:1098–1103.
39. Goverover Y, Genova HM, Hillary FG, DeLuca J: **The relationship between neuropsychological measures and the Timed Instrumental Activities of Daily Living task in multiple sclerosis.** *Mult Scler* 2007, **13**:636–644.
40. Schultheis MT, Weisser V, Ang J, *et al*: **Examining the relationship between cognition and driving performance in multiple sclerosis.** *Arch Phys Med Rehabil* 2010, **91**:465–473.

41. Beatty WW, Goretti B, Siracusa G, Zipoli V, Portaccio E, Amato MP: **Changes in neuropsychological test performance over the workday in multiple sclerosis.** *Clin Neuropsychol* 2003, **17**:551–560.

42. Goretti B, Portaccio E, Zipoli V, Razzolini L, Amato MP: **Coping strategies, cognitive impairment, psychological variables and their relationship with quality of life in multiple sclerosis.** *Neurol Sci* 2010, **31**:S227–S230.

43. Rao SM, Leo GJ, Ellington L, Nauertz T, Bernardin L, Unveragt F: **Cognitive dysfunction in multiple sclerosis. II. Impact on employment and social functioning.** *Neurology* 1991, **41**:692–696.

44. LaRocca NG, Kalb RC, Gregg K: **A program to facilitate retention of employment among persons with multiple sclerosis.** *Work* 1996, **7**:37–46.

45. Benedict RHB, Wahlig E, Bakshi R, et al: **Predicting quality of life in multiple sclerosis: Accounting for physical disability, fatigue, cognition, mood disorder, personality, and behavior change.** *J Neurol Sci* 2005, **231**:29–34.

46. Benedict RHB, Zivadinov R: **Predicting neuropsychological abnormalities in multiple sclerosis.** *J Neurol Sci* 2006, **245**:67–72.

47. Parmenter BA, Zivadinov R, Kerenyi L, et al: **Validity of the Wisconsin Card Sorting and Delis-Kaplan Executive Function System (DKEFS) sorting tests in multiple sclerosis.** *J Clin Exp Neuropsychol* 2007, **29**:215–223.

48. Morrow SA, Drake AS, Zivadinov R, Munschauer F, Weinstock-Guttman B, Benedict RHB: **Predicting loss of employment over three years in multiple sclerosis: Clinically meaningful cognitive decline.** *Clin Neuropsychol*, . in press.

49. Benedict RHB, Zgaljardic DJ: **Practice effects during repeated administrations of memory tests with and without alternate forms.** *J Clin Exp Neuropsychol* 1998, **20**:339–352.

50. Boringa JB, Lazeron RH, Reuling IE, et al: **The brief repeatable battery of neuropsychological tests: normative values allow application in multiple sclerosis clinical practice.** *Mult Scler* 2001, **7**:263–267.

51. Benedict RH, Smerbeck A, Parikh R, Rodgers J, Cadavid D, Erlanger D: **Reliability and equivalence of alternate forms for the Symbol Digit Modalities Test: implications for multiple sclerosis clinical trials.** *Mult Scler* 2012, in press.

52. Benedict RHB, Groninger L: **Preliminary standardization and validation of a new visuospatial memory test with six alternate forms.** *Clin Neuropsychol* 1995, **9**:11–16.

53. Benedict RHB, Schretlen D, Groninger L, Dobraski M, Shpritz B: **Revision of the brief visuospatial memory test: studies of normal performance, reliability, and validity.** *Psychol Assess* 1996, **8**:145–153.

54. Green MF, Nuechterlein KH: **The MATRICS initiative: developing a consensus cognitive battery for clinical trials.** *Schizophr Res* 2004, **72**:1–3.

55. Rey A: *L'Examen Clinique en Psychologie.* Paris: Press Universitaire de France; 1964.

56. Benedict RHB, Schretlen D, Brandt J, Groninger L: **Revision of the Hopkins verbal learning test: reliability and normative data.** *Clin Neuropsychol* 1998, **12**:43–55.

57. Shapiro AM, Benedict RHB, Schretlen D, Brandt J: **Construct and concurrent validity of the Hopkins Verbal Learning Test - Revised.** *Clin Neuropsychol* 1999, **13**:348–358.

58. Trojano M: **Multiple sclerosis: the evolving diagnostic criteria for multiple sclerosis.** *Nat Rev Neurol* 2011, **7**:251–252.