Brief International Cognitive Assessment for MS (BICAMS): International Standards for Validation

Ralph H B Benedict,[1] Maria P Amato,[2] Jan Boringa,[3] Bruno Brochet,[4] Frederik Foley,[5] Sten Fredrikson,[6] Päivi Hämäläinen,[7] Hans-Peter Hartung,[8] Lauren Krupp,[9] Iris-Katharina Penner,[10] Anthony Reder,[11] Dawn W Langdon.[12]

[1]Jacobs Neurological Institute, New York, NY, USA [2]Careggi University Hospital, Florence, Italy [3]Meander Medisch Centrum, Amersfoort, The Netherlands [4]Universite de Bordeaux, Bordeaux cedex, France [5]Yeshiva University, Bronx, NY, USA and Holy Name Hospital Multiple Sclerosis Center, Teaneck, NJ, USA [6]Karolinska Institute, Huddinge University Hospital, Stockholm, Sweden [7]Masku Neurological Rehabilitation Centre, Masku, Finland [8]Heinrich-Heine-Universität, Düsseldorf, Germany [9]Albert Einstein College of Medicine New York, East Setauket, NY, USA [10]University of Basel, Basel, Switzerland [11]University of Chicago, Chicago, IL, USA [12]Royal Holloway University of London, London, UK

Abstract

Background. The Brief Cognitive Assessment for MS (BICAMS) comprises the Symbol Digit Modalities Test, first five recall trials of the California Verbal Learning Test –II and first three recall trials of the Brief Visuospatial Memory Test – Revised.

Goal. To produce international standards for BICAMS psychometric validation.

Method. The international expert consensus committee reviewed basic psychometric validation standards from the literature. Recognizing that resources for validation will vary internationally, the committee identified validation priorities, to engender international acceptance of BICAMS. It was noted that most of the psychometric evidence relating to the BICAMS tests relies on US samples. Practical matters pertaining to implementation across different languages and countries were discussed.

Results. Step 1. Standardisation 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 verbal memory, 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.

Step 3. Normalisation. Samples of 150 or more healthy persons are needed for data applicable to persons of all ages. The minimum sample size is 65 healthy volunteers, matched on demographics to either a concurrent MS sample, or to samples in published descriptive MS studies.

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 >.80 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.

Discussion

Preliminary plans are being made to validate BICAMS in a number of countries, which will constitute important steps towards an international assessment tool of cognition in MS.

Introduction and purpose

Psychometric data regarding normal performance, test reliability and the validity of test interpretation are necessary for accurate application of NP testing.[1]

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. BICAMS was particularly focused on international use, permitting better comparability 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 presented elsewhere (Langdon *et al.*, submitted).

In circumstances allowing for only a few minutes of assessment, the committee recommended the Rao[2] adaptation of the Symbol Digit Modalities Test (SDMT),[3] the initial learning trials of the second edition of the California Verbal Learning Test (CVLT2).[4] and the revised Brief Visuospatial Memory Test (BVMTR).[5]

A second conference was held to develop consensus on a BICAMS validation protocol.

Results. Consensus Opinion: The BICAMS Validation Protocol

Table 1. Psychometric criteria of BICAMS tests in English speaking environments

	SDMT	CVLT2	BVMTR
Standardization	Smith 1982 Test 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. No or Little Data	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

Psychometric Standards

Standardization. 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 administration instructions are published in the test manuals.

Normalization of Raw Scores. Acquiring normative data can be an expensive endeavor. For example, the recent revised normative data for BVMTR has a sample size of 500. Manual based CVLT2 data are based on a sample of 500 healthy adults. These data were fairly recently acquired and are judged by the BICAMS committee to be current. The 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 MACFIMS battery, which includes the SDMT, in 2010,[6] using a healthy sample numbering 120. A potential problem is that normative data from one region may result in interpretive error when applied to raw test scores derived from a different culture, language, region or country.

Reliability. 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.[7] A more valid measure of test-retest reliability is the Pearson correlation coefficient.[8] For most purposes, the acceptable range of r values for test-retest correlation is > 0.80. [9] The SDMT has particularly high test-retest reliability. In one study of MS patients the test-retest r value was 0.97. In a 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 MS patients.[10]

Validity. 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. Clearly, if the test cannot discriminate MS patients from healthy individuals, the meaning of the test score is uncertain. All of the BICAMS tests discriminate very well between MS patients and healthy controls.[11] SDMT is most often the most sensitive measure in NP batteries.[11]

Table 2. Summary of national standardization requirements

- **Step 1, Standardization and Translation of Test Stimuli.** For abstract visual stimuli, determine if there are any semantic associations to stimuli in the culture or language under consideration. Or similarity to letters or other symbols in the target language. For CVLT2 must match new words on word frequency and meaning. It may be hard in other languages to match on freq, meaning and get a sensible standardization data set, for example particular animals. The match needs to be on structure of wordlist and recall performance of subjects.
- **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. Can we accept that in some centres the staff can speak English well enough for the test admin instructions to be used in English? Or do we as policy need a translation for every language we get normalization in, so that all professionals working in that language group (country) can use BICAMS?
- **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 from the same country, or matched to samples in other published descriptive MS studies.
- **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. Specify correlation type and value 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. We need to specify matching criteria. I suggest either no stat sig diffs between age, gender and years of education OR individually matched +/- 3 years of age, +/- 3 years of education, within 5% female proportion.

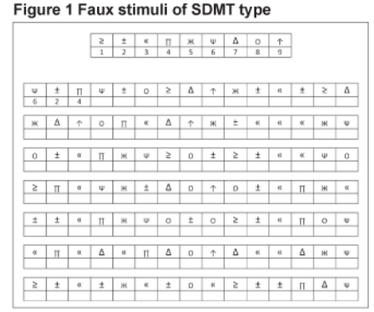
Conclusion

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. Much work needs to be done in order to validate BICAMS in non-English speaking settings. Here 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.

Brief International Cognitive Assessment for MS (BICAMS) Description

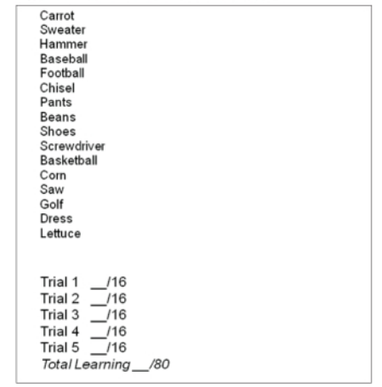
The SDMT3 presents a series of nine symbols, each paired with a single digit in a key at the top of an 8 ½ x 11 inch sheet. An adapted version of the test is presented in Figure 1. Patients are asked to voice the digit associated with each symbol as rapidly as possible for 90 sec. There is a single outcome measure – the number correct over the 90 sec time span.

Figure 1 Faux stimuli of SDMT type



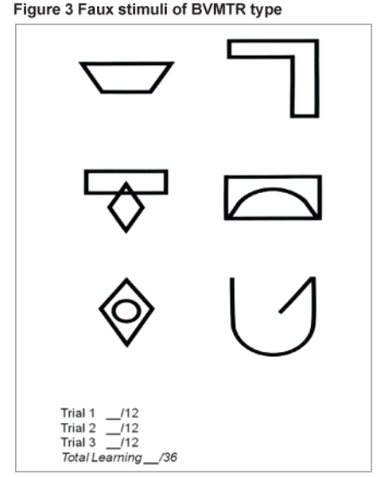
The auditory/verbal learning test is the CVLT2.[4] The test begins with the examiner reading a list of 16 words [Figure 2]. Patients listen to the list and report as many of the items as possible. There is no instruction as to the order in which items are recalled. After recall is recorded, the entire 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 validity with these measures, as the total number of recalled over the five learning trials is most sensitive.[10]

Figure 2 Faux stimuli of CVLT-II type



Visual/spatial memory is assessed in BICAMS using the BVMTR.[5] In this test, six abstract designs [Figure 3] are presented to patients 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. As in the case of the CVLT2, we find little evidence in the literature that the delayed recall and recognition tasks add incremental validity with MS samples, and therefore, only the initial learning trials are implemented in the BICAMS context.

Figure 3 Faux stimuli of BVMTR type



References: 1. Lezak M. Neuropsychological Assessment. 2nd ed. New York: Oxford University Press; 1995. 2. Rao SM. A Manual for the Brief, Repeatable Battery of Neuropsychological Tests in Multiple Sclerosis. National Multiple Sclerosis Society; 1991. 3. Smith A. Symbol digit modalities test: Manual. Los Angeles: Western Psychological Services; 1982. 4. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test Manual: Second Edition, Adult Version. San Antonio, TX: Psychological Corporation; 2000. 5. Benedict RHB. Brief Visuospatial Memory Test - Revised: Professional Manual. Odessa, Florida: Psychological Assessment Resources, Inc.; 1997. 6. Parmenter BA, Testa SM, Schretlen DJ, Weinstock-Guttman B, Benedict RHB. Journal of the International Neuropsychological Society 2010;16:6-16. 7. Horsfield MA, Rovaris M, Rocca MA, et al. Journal of the Neurological Sciences 2003;216:169-77. 8. Cohen J, Cohen P. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, Second Edition. Hillsdale, NJ: Lawrence Erlbaum; 1983. 9. Anastasi A. Psychological Testing. 6th ed. Macmillan Publishing Company; 1988. 10. Benedict RHB. Journal of the International Neuropsychological Society 2005;11:727-36. 11. Strober L, Englert J, Munschauer F, Weinstock-Guttman B, Rao S, Benedict RH. Mult Scler 2009;15:1077-84.

Professor Benedict has acted as a consultant or scientific advisory board member for Bayer, Biogen Idec, Merk Serono, EMD Serono, Pfizer, Novartis, royalties from Psychological Assessment Resources, Inc; financial support for research activities from Shire Pharmaceuticals. Biogen Idec. Professor Amato received personal compensation from Merck Serono, Biogen Dompé, Sanofi-Aventis, Bayer Schering for serving on scientific advisory board and for speaking. Prof. Amato received financial support for research activities from Merck Serono, Sanofi-Aventis, Biogen Dompé. Bayer Schering. Dr Boringa has consulted for Bayer Healthcare and served on speaker bureau for Exencia Pharma Academy. Dr. Brochet serves on scientific advisory boards for Bayer Schering Pharma, Novartis, and Merck Serono; has received funding for travel and speaker honoraria from Bayer Schering Pharma, Merck Serono, Biogen Idec, AB Sciences, and Teva Pharmaceutical Industries Ltd./Sanofi-Aventis, and serves as LEN editor for SEP et Neurosciences. Dr Foley has received honoraria for speaking at scientific and patient-education meetings, serving on scientific advisory boards and consulting activities from Bayer-Schering, Biogen, and Teva Neuroscience. Dr Fredrikson has received honoraria for lectures, educational activities and consultancy from Bayer, Biogen Idec and Sanofi-Aventis during the past year. Dr Hämäläinen received personal compensation from Bayer Healthcare and Novartis for serving on scientific advisory boards, consulting for Sanofi-Aventis, served on speaker bureau for Bayer Healthcare and Sanofi-Aventis. Prof Hartung received honoraria with approval by the Rector of Heinrich-Heine-University from Bayer Healthcare GmbH, Octapharma GmbH, Novartis, Teva Sanofi Aventis, Biogen Idec GmbH and Merck Serono GmBH for consulting and speaking at scientific symposia. Dr Krupp has served on speaker bureau, scientific advisory boards and/or been a consultant for Teva Neurosciences, BiogenIdec, EMD Serono, Multiple Sclerosis Association of America, Betaseron/Bayer Healthcare Pharmaceuticals, Pfizer, Sanofi-Aventis, Axon Advisors; she has received royalties from Genzyme, ER Squibb & Sons, NIMSS, Novartis, MedImmune. Abbott Laboratories, Johnson&Johnson, Roche, Health Professions Conferencing Corp. Dr Penner has received honoraria for speaking at scientific meetings, serving at scientific advisory boards and consulting activities from Bayer-Schering, Biogen, Merck Serono, Roche, Novartis. A.T.R has received compensation from many of the sources below and is an Advisory Board/ Consultant for: Abbott Laboratories, Immunoscience, Parsippany, NJ, American Medical Association, Chicago, IL, Astra Merck, Wayne, PA, Athena Neurosciences, South San Francisco, CA, *Aventis Pharma, Bridgewater, NJ, *Bayer HealthCare Pharmaceuticals, Montville, NJ, *Berlex Laboratories, Richmond, CA, *BioMS Medical Corp., Edmonton, Alberta, Canada, *Biogen and BiogenIdec, Cambridge, MA, Blue Cross, Blue Shield, Chicago, IL, Boehringer Ingelheim Pharmaceuticals Inc., Ridgefield, CT, *Caremark Rx, Northbrook, IL., Centocor, Inc, Malvern, PA, Cephalon, Inc, Delaware, MD, Connectics/Connective Therapeutics, Palo Alto, CA, CroMedica Global Inc. Victoria, BC, *Elan Pharmaceuticals, Inc., San Diego, CA, *Eli Lilly and Company, Indianapolis, IN, *Genentech, South San Francisco, CA, *Genzyme Corporation, San Antonio, Tx, GlaxoSmithKline, Research Triangle Park, NC, Hoechst Marion Roussel Canada Research, Inc., Laval, Quebec, Hoffman-LaRoche, Nutley, NJ, Idec, San Diego, CA, Immunex, Seattle, WA, Institute for Health Care Quality, Minneapolis, MN, Johnson & Johnson, Pharmaceutical Research & Development, LLC, Raritan, NJ, *MedImk/Neurobase electronic journal – editor and author, NARCOMS, Yale University, New Haven, CT & Barrow Neurological Institute, Phoenix, AZ, *National MS Society, National Multiple Sclerosis Society & Paralyzed Veterans of America, Pain Panel, NY, *Neurocrine Biosciences, San Diego, CA, *Novartis Corporation, NY, Parke-Davis, Morris Plains, NJ, *Pfizer Inc, NY, Pharmacia & Upjohn, Kalamazoo, MI, Protein Design Labs, Inc, Delaware, Quantum Biotechnologies, Inc., Laval, Quebec, Quintiles, Inc, San Diego, CA, *RENEW study (post-marketing study of Novantrone in MS), Serono, Sindoz (now Novartis) & Novartis, East Hanover, NJ, Sention, Inc, Providence, RI, *Schering, Berlin, Germany, *Serono, Norwell, MA, Smith Kline-Beecham, Philadelphia, PA, Specialized Therapeutics, a division of Berlipharm, Inc, Montville, NJ, Takeda Pharmaceuticals, Lincolnshire, IL, *Teva-Marion, Kansas City, MO. Anthony T Reder has also received research support from Bayer, Serono, Teva, MSociety. Dr Langdon has received funding for travel from Bayer Healthcare, Bayer Schering AG, serves on a Steering Committee for Bayer Healthcare; has served/serves as a consultant to Merck-Serono, Bayer Healthcare, Bayer Schering AG, serves on speaker bureau for Bayer Healthcare, Bayer Canada, Bayer Australia, Roche, Serono Symposia, TEVA Germany; receives research contract funding from Bayer Healthcare, receives educational grants from Bayer Healthcare and Bayer Schering AG.

The BICAMS committee meetings are supported by an educational grant from Bayer Healthcare Pharmaceuticals, Montville, NJ, USA