A new assessment tool for patients with multiple sclerosis from Spanish-speaking countries: validation of the Brief International Cognitive Assessment for MS (BICAMS) in Argentina

Sandra Vanotti, Audrey Smerbeck, Ralph H. B. Benedict & Fernando Caceres


To link to this article: http://dx.doi.org/10.1080/13854046.2016.1184317

Published online: 12 May 2016.
A new assessment tool for patients with multiple sclerosis from Spanish-speaking countries: validation of the Brief International Cognitive Assessment for MS (BICAMS) in Argentina

Sandra Vanotti a, Audrey Smerbeck b, Ralph H. B. Benedict c and Fernando Caceres a

aMultiple Sclerosis Clinic, INEBA – Neurosciences Institute of Buenos Aires, Buenos Aires, Argentina; bDepartment of Psychology, Rochester Institute of Technology, Rochester, NY, USA; cSUNY at Buffalo School of Medicine, Neurology, Buffalo, NY, USA

ABSTRACT

Background: The Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) is an international assessment tool for monitoring cognitive function in multiple sclerosis (MS) patients. BICAMS comprises the Symbol Digit Modalities Test (SDMT), the California Verbal Learning Test – Second Edition (CVLT II) and the Brief Visuospatial Memory Test – Revised (BVMT-R). Our objective was to validate and assess the reliability of BICAMS as applied in Argentina and to obtain normative data in Spanish for this population. Method: The sample composed of 50 MS patients and 100 healthy controls (HC). In order to test its reliability, BICAMS was re-administered in a subset of 25 patients. Results: The sample’s average age was 43.42 ± 10.17 years old, and average years of schooling were 14.86 ± 2.78. About 74% of the participants were women. The groups did not differ in age, years of schooling, or gender. The MS group performed significantly worse than the HC group across the three neuropsychological tests, yielding the following Cohen’s d values: SDMT: .85; CVLT I: .87; and BVMT-R: .40. The mean raw scores for Argentina normative data were as follows: SDMT: 56.71 ± 10.85; CVLT I: 60.88 ± 10.46; and BVMT-R: 23.44 ± 5.84. Finally, test–retest reliability coefficients for each test were as follows: SDMT: r = .95; CVLT I: r = .87; and BVMT-R: r = .82. Conclusion: This BICAMS version is reliable and useful as a monitoring tool for identifying MS patients with cognitive impairment.

Background

Multiple sclerosis (MS) is an autoimmune disease which presents as an acute inflammatory relapse or progressive disease of the white matter in the central nervous system (Pirko & Noseworthy, 2007). Variations in the prevalence of MS can be observed across continents, due to differences in geographical and environmental features (Aguirre-Cruz, Flores-Rivera, De La Cruz-Aguilera, Rangel-López, & Corona, 2011). The Argentinian population is characterized by a low prevalence and incidence in relation to reports from more developed countries. An annual incidence rate ranged from .15 to 1.9 cases per 100,000 person-years, and
prevalence ranged from .75 to 21.5 cases per 100,000 inhabitants (Cristiano, Patrucco, & Rojas, 2008). Regarding its clinical pattern, the disease generally displays a similar manifestation to the classic (or Western) forms (Aguirre-Cruz et al., 2011).

It is well known that cognitive impairment (CI) is common in patients with MS (Benedict & Zivadinov, 2011). Its prevalence is estimated between 40 and 70% (Chiaravalloti & DeLuca, 2008). In Latin America, unlike in North America and Western Europe, there are few studies on the subject. An epidemiological study carried out in Argentina reports that 43.2% of MS patients have CI (Cáceres, Vanotti, Rao, & RECONEM Workgroup, 2011), while another multicenter study of MS patients from Argentina, Chile, Colombia, Mexico, Uruguay, and Venezuela shows that there is a prevalence of 34.5% of CI in the early stages of the disease (Cáceres, Vanotti, Benedict, & RELACEM Workgroup, 2014).

Cognitive function impairment has adverse effects on different aspects of a person’s quality of life (Langdon, 2012), and an early and rapid detection helps mitigate the difficulties it causes. It not only affects a person’s capacity of safe driving, completion of household tasks, participation in social activities, physical independence, rehabilitation progress, and coping, but also adherence to treatment, mental health (Langdon et al., 2011), and performance at work (Benedict, Rodgers, Emmert, Kninger, & Weinstock-Guttman, 2014). For that reason, an expert committee of neurologists and neuropsychologists has developed the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS), an international assessment tool to monitor cognitive function in MS patients. The test requires an approximately 15-min-long administration. It was designed and optimized for small centers where there are few staff members, who may not have neuropsychological training. BICAMS was particularly designed for international usage, so as to facilitate comparison across settings (Benedict et al., 2012). It includes three tests of high reliability and good sensitivity: Rao’s (1991) adaptation of the Symbol Digit Modalities Test (SDMT) (Smith, 1982), the initial learning trials of the second edition of the California Verbal Learning Test (CVLT II) (Delis, Kramer, Kaplan, & Ober, 2000), and the total learning score of the Brief Visuospatial Memory Test – Revised (BVMT-R) (Benedict, 1997; Benedict et al., 2012).

BICAMS’ aim is to standardize assessment across languages and nationalities, and up to now, it had not been validated in Spanish. With that purpose in mind, we proceeded to validate and assess the reliability of BICAMS in Argentina and to obtain normative data in Spanish for this population.

Method

Participants

We studied 50 patients who had MS as diagnosed by standard criteria (Polman et al., 2011) and 100 healthy controls (HC). The groups did not differ with regard to age (MS group \(M = 43.42 \pm 10.17; HC = 42.37 \pm 10.07\) years, \(p = .55\)), years of schooling (MS group = 14.86 \(\pm 2.78\); HC = 14.94 \(\pm 2.49\), \(p = .86\)), or gender (MS patients: 74% women; HC: 75% women).

The majority of the first group had MS of the relapsing remitting type (78%, \(n = 39\); secondary progressive 18%, \(n = 9\); primary progressive 4%, \(n = 2\)). The average disease duration was 13.06 \(\pm 9.08\) years. The Expanded Disability Status Scale (EDSS) mean was 3.29 \(\pm 2.55\). MS Functional Composite (MSFC) z-score mean was .32 \(\pm 1.55\).
Participants were recruited from the general population of Buenos Aires city, Argentina; the process lasted from December 2013 to October 2014. Patients were contacted by research staff to ask whether they were interested in participating. As is customary in Argentina, they received no payment for participating. They all provided informed consent, and the study was approved by the Ethics Committee of the Buenos Aires Institute of Neurosciences. Inclusion criteria for the MS group were a diagnosis of clinically defined MS, 18–60 years old, fluency in Spanish, and ability to provide informed consent for all procedures. Exclusion criteria were current or previous neurological disorder other than MS, history of psychotic disorder, current psychiatric disorder related to mood, personality, or behavior changes following the onset of MS, a medical condition that might affect cognition, a history of developmental disorder, a history of substance or alcohol dependence, current substance abuse, motor or sensory disability that might interfere with cognitive test performance, and having experienced a relapse and/or corticosteroid pulse in the previous 4 weeks of assessment. HC were required to possess scores >26 on Folstein Mini-Mental State Examination (MMSE) (Butman et al., 2001) and <14 in the Beck Depression Inventory (BDI) (Beck, Steer y Brown, 2006) according to Argentinean adaptation (Brenlla & Rodriguez, 2006), and free from any history of neurological disease.

**Tests and procedures**

A Spanish adaptation of the BICAMS battery (Benedict et al., 2012) was administered by a staff of neuropsychologists.

The procedure began with the administration of the SDMT (Smith, 1982), in which a series of nine symbols were displayed on a standard sheet of paper, each paired with a single digit in a key at the top of the page. Over a period of 90 seconds, participants were asked to speak out the digit associated with each symbol as rapidly as possible. There was a single outcome measure: the number of correct answers obtained over the 90’s time span. Two equivalent versions were used for carrying out the test–retest procedure.

The official BICAMS battery would next include the CVLT II (Delis et al., 2000), but for our purposes, the Spanish version of the CVLT I was employed to provide a meaningfully equivalent assessment. This test began with the examiner reading a list of 16 words. Participants were asked to listen to the list and report as many of the items as possible, in any order. Afterward, the entire list was read again followed by a second attempt at recalling the words. There were five learning trials altogether. The outcome measure is the total number of recalled items over the five learning trials. The Spanish version of the CVLT I (Artiola, Fortuny, Hermosillo Romo, Heaton, & Pardee, 1998) was used. In the CVLT I validation (Artiola et al., 1998), the procedure for the selection of the stimuli words was as follows: Native Spanish speakers from 10 different countries were asked to generate as many words as possible that fell in each of the selected categories. Of all words generated for each category, the two words with the highest frequency were rejected. Any words and categories that did not share the same meaning in different regions were rejected (Artiola et al., 1998). The best remaining words were used to generate two lists, A and B. List A was used for the test procedure, while list B was for retest procedure.

Finally, the BVMT-R (Benedict, 1997) was administered, in which six abstract designs were displayed for ten seconds. The display was then removed from view and participants were asked to render the stimuli manually, with a pencil on a sheet of paper. Each drawing received
a score of 0–2 points, representing location and accuracy. Thus, overall trial scores range from 0 to 12. There were three learning trials, and the primary outcome measure is the total number of points earned over the three learning trials. Form 1 was used for the test procedure, and Form 4 was used for retesting. For the HC group, the MMSE was also administered, as adapted for Argentina (Butman et al., 2001).

In addition, the MSFC (Fischer, Rudick, Cutter, & Reingold, 1999) was administered, composed of a timed 25-foot walking task and a nine-hole peg placement task, executed with both hands. The PASAT 3 s interstimulus interval trial was combined with these motor tasks to yield a composite score in accordance with the published algorithm (Cutter et al., 1999). The EDSS (Kurtzke, 1983) was also completed by neurologists. Neuropsychologists also administered BDI in the adapted Spanish version (Brenlla & Rodríguez, 2006).

Twenty-five patients were examined twice at an interval of two weeks (test–retest group), using alternate forms to mitigate practice effects. For SDMT (Smith, 1982), Form 4 was used; for CVLT I (Artiola et al., 1998), list B, and for BVMT-R (Benedict et al., 2012), form 4.

**Statistical analysis**

A data analysis was carried out using SPSS Statistics, version 19.0. An analysis of variance was used to compare the MS group with the HC group. Effect sizes of the differences of means were calculated of each difference using Cohen’s $d$. Pearson’s correlation coefficient was used for retest study.

Regression-based norms were created following the procedure established by previous studies (Goretti et al., 2014; O’Connell, Langdon, Tubridy, Hutchinson, & McGuigan, 2015; Parmenter, Marc Testa, Schretlen, Weinstock-Guttman, & Benedict, 2010). The control group’s raw scores on each neuropsychological measure were converted to uncontrolled scaled scores ($M = 10$, $SD = 3$) using the cumulative frequency distribution of each measure. Then, a multiple regression model was generated to predict the HC participants’ actual raw scores incorporating age, age$^2$, sex (male = 1, female = 2), and education in years as predictor variables. The unstandardized $\beta$ weights for each predictor variable and the $SD$ of the residuals were retained for each model. These values were used to create an equation that compares an individual’s predicted performance with actual performance, yielding a $z$-score.

**Results**

**Psychometric characteristics**

**Evidence of criterion validity by contrasting the groups**

In order to assess criterion validity by contrasting the groups on the three tests, the mean raw scores obtained by each group along the three tests were compared. The MS group performed significantly worse than the HC group on the three neuropsychological tests (SDMT: $p < .001$; CVLT I: $p < .001$; BVMT-R: $p < .017$). Cohen’s $d$ was as follows for each test: SDMT: .85; CVLT I: .87, and BVMT-R: .40 (Table 1).
Reliability

Test–retest

Test–retest reliability coefficients for each test were as follows: SDMT: \( r = .95 \); CVLT I: \( r = .87 \); and BVMT-R:\( r = .82 \). Table 2 shows mean values and standard deviation values of the BICAMS battery for the group of 25 patients who received this test battery twice in an interval of 15 days.

Derivation and use of regression-based norms

The regression-based normative model for the BICAMS is provided in Tables 3 and 4. To calculate a patient’s z-score using the normative data provided, first calculate the predicted scaled score using the following equation, the constants provided in Table 3, and the patient’s demographic data: \( SS_{\text{predicted}} = \text{Constant} + (\text{Age} \times \beta_{\text{age}}) + (\text{Age}^2 \times \beta_{\text{age}^2}) + (\text{Sex} \times \beta_{\text{sex}}) + (\text{Education} \times \beta_{\text{Education}}) \). The patient’s sex should be coded 1 if male and 2 if female. Age and education should be provided in years. The value obtained from this equation is the patient’s predicted performance, given demographic information. Next, use Table 4 to convert the patient’s raw score into an unstandardized scaled score, \( SS_{\text{actual}} \). The difference between the actual and predicted scaled scores is standardized by dividing it by the standard deviation of the residuals, found...
in Table 3, using the following equation: \( Z = \frac{(SS_{\text{actual}} - SS_{\text{predicted}})}{SD_{\text{residual}}} \). The resulting \( z \)-score represents the patient’s performance adjusted for age, sex, and education.

**Discussion**

The BICAMS battery is a fast and reliable instrument that has been thoroughly validated. Its aim is to survey the cognitive difficulties caused by MS, and it is optimized for international usage and small centers that specialize in the disease (Langdon et al., 2011). Culture influences affect neuropsychological performance (Nisbett, Peng, Choi, & Norenzayan, 2001) not only during verbal tasks, but also during non-verbal tests (Rosselli & Ardila, 2003). Taking that into consideration, the following aims were established for this investigation: to validate BICAMS and assess its reliability for Argentina’s population, as well as to obtain normative data in Spanish. There exists normative data for populations of other countries, such as Czech Republic (Dusankova, Kalincik, Havrdova, & Benedict, 2012), Italy (Goretti et al., 2014), Hungary (Sandi et al., 2015), Ireland (O’Connell et al., 2015), and Brazil (Spedo et al., 2015), but up to now, there have not been any validations conducted in Spanish.

All three tests composing this validation of BICAMS (SDMT, CVLT I, and BVMT-R) show significant differences between MS patients and HCs, a finding that is in agreement with the other recently published validations (Dusankova et al., 2012; Goretti et al., 2014; O’Connell et al., 2015; Sandi et al., 2015; Spedo et al., 2015). These differences were more marked at the SDMT and CVLT I than the BVMT-R, and similar results were found by O’Connell et al. (2015) and Spedo et al. (2015). These results differ from findings by Dusankova et al. (2012), in which the tests that presented the biggest differences were SDMT and BVMT-R.

Test–retest reliability for raw scores was good to excellent for all the three tests in this validation, more than .80 in all measures, replicating prior findings (O’Connell et al., 2015) and somewhat exceeding the results found by Goretti and colleagues in control participants (Goretti et al., 2014).

As our HC sample was not fully demographically representative (e.g. with respect to sex), it was necessary to control for demographic variables when generating our norms.

---

**Table 4. Raw score to uncontrolled scaled score conversions.**

<table>
<thead>
<tr>
<th>Scaled score</th>
<th>SDMT</th>
<th>CVLT I</th>
<th>BVMT-R</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0–27</td>
<td>0–32</td>
<td>0–6</td>
</tr>
<tr>
<td>3</td>
<td>28–31</td>
<td>33–35</td>
<td>7–10</td>
</tr>
<tr>
<td>4</td>
<td>32–33</td>
<td>36–39</td>
<td>11</td>
</tr>
<tr>
<td>5</td>
<td>34–40</td>
<td>40–43</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>41–44</td>
<td>44–48</td>
<td>13–16</td>
</tr>
<tr>
<td>7</td>
<td>45–47</td>
<td>49–51</td>
<td>17–19</td>
</tr>
<tr>
<td>8</td>
<td>48–52</td>
<td>52–56</td>
<td>20–21</td>
</tr>
<tr>
<td>9</td>
<td>53–55</td>
<td>57–61</td>
<td>22</td>
</tr>
<tr>
<td>10</td>
<td>56–59</td>
<td>62–65</td>
<td>23–25</td>
</tr>
<tr>
<td>11</td>
<td>60–62</td>
<td>66–67</td>
<td>26</td>
</tr>
<tr>
<td>12</td>
<td>63–65</td>
<td>68–69</td>
<td>27–28</td>
</tr>
<tr>
<td>13</td>
<td>66–67</td>
<td>70–71</td>
<td>29</td>
</tr>
<tr>
<td>14</td>
<td>68–71</td>
<td>72–73</td>
<td>30–31</td>
</tr>
<tr>
<td>15</td>
<td>72–74</td>
<td>74–75</td>
<td>32</td>
</tr>
<tr>
<td>16</td>
<td>75–78</td>
<td>76</td>
<td>33</td>
</tr>
<tr>
<td>17</td>
<td>79–85</td>
<td>77</td>
<td>34–35</td>
</tr>
<tr>
<td>18</td>
<td>≥86</td>
<td>78–80</td>
<td>36</td>
</tr>
</tbody>
</table>

Notes: SDMT: Symbol Digit Modalities Test; CVLT I: California Verbal Learning Test; BVMT-R: Brief Visuospatial Memory Test Revised.
Normative values were obtained for the HC group using BICAMS as adapted for a Spanish-speaking Argentinian population, which proved to be a reliable and valid means of screening for and monitoring potential cognitive impairment in patients who have MS. As such, norm-referenced $z$-scores can be calculated for individual patients, comparing their present performance with their demographically predicted performance. For example, imagine a 27-year-old woman with 12 years of education who obtained a raw score of 36 on the SDMT. Using Table 4, this raw score can be converted to an uncontrolled scaled score of 5. (The uncontrolled scaled score compares the patient to the entire sample without regard for demographics and should not be directly interpreted.) The patient’s predicted scaled score can be calculated by combining her demographic information with the values in Table 3, as per the equation provided above: $SS_{\text{predicted}} = \text{Constant} + (\text{Age} \times \beta_{\text{age}}) + (\text{Age}^2 \times \beta_{\text{age}^2}) + (\text{Sex} \times \beta_{\text{sex}}) + (\text{Education} \times \beta_{\text{Education}})$. Thus, this patient’s predicted scaled score is equal to $10.9231 + (27 \times -0.2218) + (27^2 \times .0013) + (2 \times .2850) + (12 \times .3714)$, or 10.909. Her predicted and actual scaled scores are compared and the difference divided by the $SD_{\text{residual}}$ as found in Table 3. This yields a $z$-score, following the equation $Z = (SS_{\text{actual}} - SS_{\text{predicted}})/SD_{\text{residual}}$. For this patient, $Z = (5 - 10.909)/2.5842 = -2.2866$. Rounding, we find that this patient’s demographically adjusted SDMT $z$-score is $-2.29$.

**Conclusions**

It has been widely documented that CI has a substantial impact on the day-to-day function of patients with MS in the occupational, educational, and social realms. Early detection of cognitive difficulties is vital in order to design cognitive rehabilitation plans, provide vocational guidance, and quantify the disability caused by the disease. Unfortunately, in Argentina and other Latin American countries, not all of the health centers have neuropsychologists among their staff with clinical expertise in MS-related CI who can evaluate patients. The BICAMS battery has been designed with the aim to assess the cognitive performance of patients with MS and can be used in everyday practice by clinical neurologists or administered by local health care workers (Langdon, 2012). Thus, the adapted BICAMS will provide Latin American MS patients with the opportunity to receive regular and timely cognitive evaluation. Normative values have been obtained for Latin American countries through this validation.

**Abbreviations**

- **MS**: multiple sclerosis
- **BICAMS**: Brief International Cognitive Assessment for Multiple Sclerosis
- **SDMT**: Symbol Digit Modalities Test
- **CVLT**: California Verbal Learning Test
- **BVMT-R**: Brief Visuo-spatial Memory Test – Revised
- **CI**: cognitive impairment
- **HC**: healthy controls
- **MMSE**: Mini-Mental State Examination
- **MSFC**: MS Functional Composite
- **EDSS**: Expanded Disability Status Scale
Author contributions

SV contributed in drafting/revising the manuscript, study concept and design, analysis and interpretation of data, acquisition of data, statistical analysis, and coordination. ES contributed in revising the manuscript, analysis and interpretation of data, and statistical analysis. RHB contributed in study concept and design, revising the manuscript, analysis and interpretation of data, and statistical analysis. CF contributed in study concept and design, revising the manuscript, analysis and interpretation of data, and study supervision. All authors read and approved the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by Novartis Argentina.

References


