The Brief International Cognitive Assessment in Multiple Sclerosis (BICAMS): Validation in Arabic and Lebanese Normative Values

Hala Darwish[1],* , Pia Zeinoun[1] , Natali Farran[1], Husam Ghusn[2], Bassem Yamout[1] and Samia J. Khoury[1]

[1]American University of Beirut, P.O.Box 11-0236, Riad El-Solh / Beirut 1107 2020, Lebanon
[2]Ain Wazein Hospital, El Chouf, 5841, Lebanon

Abstract

Objective: Multiple sclerosis (MS) is often associated with cognitive deficits. Accurate evaluation of the MS patients' cognitive performance is essential for diagnosis and treatment recommendation. The Brief International Cognitive Assessment in Multiple Sclerosis (BICAMS), widely used cognitive testing battery, examines processing speed, verbal and visuospatial learning, and memory. Our study aims to examine the psychometric properties of an Arabic version of the BICAMS and to provide normative values in a Lebanese sample. **Method:** The BICAMS, comprised of the Symbol Digit Modalities Test (SDMT), Brief Visuospatial Memory Test-Revised (BVMT-R), and a newly developed verbal learning/memory test, the Verbal Memory Arabic Test (VMAT), were administered on healthy subjects and MS patients. The sample consisted of 180 healthy individuals, of whom 63 were retested after 2–3 weeks. Forty-three MS patients matched with 43 healthy subjects based on age, sex, and years of education were assessed. A sample of 10 MS patients was also examined on two occasions. Test–retest reliability and criterion-related validity were examined, and regression-based norms were derived. **Results:** The test–retest correlations showed good evidence of reliability with coefficients ranging between 0.64 and 0.73 in the healthy sample, and between 0.43 and 0.92 in the MS sample. The BICAMS was able to discriminate between MS patients and matched healthy participants on the SDMT and BVMT-R. Normative data were comparable to other studies. **Conclusions:** This new Arabic version of the BICAMS shows initial good psychometric properties. While good evidence of VMAT's reliability was shown in the healthy participants, less test–retest reliability in this tool was seen in the MS group, and partial criterion-related validity was evident. This renders further examination of the VMAT. We provide regression-based norms for a Lebanese sample and encourage the use of this battery in both research and clinical settings.

Keywords: BICAMS, Processing speed, Visuospatial memory, Verbal learning/memory, Psychometric properties, Lebanon

BACKGROUND

Cognitive impairment is increasingly recognized as a core feature of multiple sclerosis (MS) (Langdon, 2011), with prevalence rates ranging from 43 to 70% (Chiaravalloti & DeLuca, 2008). A characteristic pattern for these impaired cognitive domains has been identified in this population with some interpatient variability (Amato, Zipoli, & Portaccio, 2008). Deficits in processing speed and episodic memory, and to a lesser extent, in executive and visuospatial functions (Benedict et al., 2006; Chiaravalloti & DeLuca, 2008; Deloire et al., 2011; Rao, Leo, Bernardin, & Unverzagt, 1991; Strober et al., 2009; Sumowski et al., 2018; Trenova et al., 2016) are

the most common, while language abilities and comprehension are usually spared (Brassington & Marsh, 1998; Calabrese, 2006; Piras et al., 2003; Prosiegel & Michael, 1993; Rao, 1995; Rocca et al., 2015). Cognitive impairment is present in all types of MS as well as disease stages – independent from any physical disability. It influences several aspects of the patient's life negatively (Amato, Zipoli, & Portaccio, 2006), health-related quality of life (Mitchell, Benito-León, González, & Rivera-Navarro, 2005), activities of daily living (Goverover, Genova, Hillary, & DeLuca, 2007), social life (Bobholz & Rao, 2003), and employment activities (Clemens & Langdon, 2018). The latest longitudinal studies showed that cognitive impairment in MS predicts disability progression and cortical thinning (Moccia et al., 2016; Pitteri, Romualdi, Magliozzi, Monaco, & Calabrese, 2017). Subsequently, valid assessments of cognition in

*Correspondence and reprint requests to: Hala Darwish, American University of Beirut, P.O.Box 11-0236, Riad El-Solh / Beirut 1107 2020, Lebanon. E-mail: hd30@aub.edu.lb

this population could contribute to more accurate care plans, therefore, enhancing patients' functioning, productivity, and quality of life (Cheng et al., 2010; Kalb et al., 2018).

Several batteries for assessing cognitive function in MS have been developed and validated (Sumowski et al., 2018). The most commonly used ones are the Brief Repeatable Battery of Neuropsychological Tests (BRB-N) (Rao, 1990), Minimal Assessment of Cognitive Function in MS (MACFIMS) (Benedict et al., 2002), and the Brief International Cognitive Assessment for MS (BICAMS) (Langdon et al., 2012). The BICAMS has several advantages over other batteries, such as the BRB-N and MACFIMS – non-neuropsychology specialists can administer it requires a shorter time for administration, and international use was taken into consideration during development. Moreover, the battery was shown to have ecological validity; it predicted the real-life functional performance of MS patients (Goverover, Chiaravalloti, & DeLuca, 2016). The BICAMS examines the commonly affected cognitive functions in MS using three standardized tests with stable psychometric properties (Langdon et al., 2012); the Symbol Digit Modalities Test (SDMT) (Smith, 1982), California Verbal Learning Test-Second Edition (CVLT-II) (Delis, 2000), and the Brief Visuospatial Memory Test-Revised Edition (BVMT-R) (Benedict, 1997).

Although the BICAMS scores have shown evidence of their validity across 11 languages (Corfield & Langdon, 2018), such evidence is currently lacking among Arabic-speaking populations. An Arabic BICAMS with evidence of reliability is currently present (Kishk et al., 2017); however, the psychometric validation of the battery has not been completed and might not be appropriate for other Arab countries, as noted by Paul, Brown and Hughes (2019) in their recent systematic review (Paul et al., 2019). Results point to the need for additional translation, adaptation, and validation of standard MS cognitive measures for use in Arabic-speaking populations, although some limitations of BICAMS exist (El Ghoneimy, Hassan, Homos, Farghaly, & Dahshan, 2015; Hamdy et al., 2013; Kishk et al., 2017; Paul et al., 2019). In our study, we examined a new Arabic version of the battery following the BICAMS international standards for validation (aim 1) (Benedict et al., 2012). This version of the BICAMS includes a newly developed Verbal Memory Arabic Test (VMAT) that could be generalized to most Arab cultures (Zeinoun, Farran, Khoury, & Darwish, 2020). Additionally, given the importance of producing normative data relevant to local populations (Smerbeck et al., 2018), we provide normative values for the BICAMS in a Lebanese sample (aim 2).

Design and Methods

This cross-sectional observational study was approved by the American University of Beirut Institutional Review Board,

and all participants signed an informed consent. Human data were obtained in compliance with the Helsinki Declaration. The study was conducted in the period 2017–2019. The methods followed were primarily based on other studies that provided evidence of BICAMS validity (Filser et al., 2018; Goretti et al., 2014; Ozakbas et al., 2017; Polychroniadou et al., 2016; Sousa et al., 2018).

Sample

For the sample size of both healthy and MS patients, we followed Benedict et al. (2012) recommendations. The authors mentioned that 150 or more healthy persons are needed for data applicable to persons of all ages and diverse ethnicity. Benedict et al. (2012) likewise mentioned that additional 35 healthy participants could be recruited to round out the normalization sample. The authors also mentioned that ideally 65 healthy individuals should be matched with MS patients. Nevertheless, in our study, data from the MS patients were limited in terms of the number of participants.

Initially, we recruited patients using flyers posted at the hospital, clinics, and various social media platforms, then we relied on the participants' word of mouth and snowball techniques that we found to be more efficient. Around 75% of the final sample was recruited through word of mouth and snowballing. We attempted to purposefully target potential participants from different areas of Lebanon and age groups. There was no incentive for participating.

The final sample included 180 healthy participants recruited from the community. From this group, a subsample of 63 individuals was retested after 1–3 weeks. We subjected all participants to the same rigorous two phases of inclusion/exclusion screening process. During the first phase of enrollment, healthy participants, older than 16 years without a history of neurological disorders, traumatic brain injury, and psychiatric disorders, including alcohol and/or drug dependence, were included in the study. Men who consumed more than 15 drinks/week and women who consumed more than eight drinks/week were considered excessive alcohol consumers and excluded from the study (McGuire, 2011). During the 3 months prior to screening, individuals who were on antidepressants, mood stabilizers, or medications known to affect cognitive performance were excluded as well.

Excerpts from the BRFSS were used to complete the screening (Centers for Disease Control and Prevention, 2009). The BRFSS is a screening questionnaire that enquires about the participant's current health status, medical history, physical activity (weekly frequency, type, and duration of activity), smoking habits (current and previous smoking habits), age, years of education, and alcohol consumption (weekly and monthly). The following data were gathered using the BRFSS: age, years of education, marital status, educational attainment, current employment, annual household income, area of residence, primary language, diagnosis with any illness (if yes, indicate illness), current use of medications, current participation in volunteering activities,

frequency of physical activity (if any), presence of any difficulties that limit one's activities in addition to factors associated with cognitive function such as smoking including hubble bubble (if yes, amount smoked per day), alcohol intake in the past 30 days (frequency and amount), leisure activities, and cognitive performance in the last 12 months (presence of any difficulties, impact on daily activities.). Information about these variables was obtained through participant's self-report.

After completing the second phase of enrollment, we used the Arabic version of the Hopkins Symptoms Checklist-25 (HSCL-25) to screen for symptoms of depression (during the past week). This phase was administered post-consenting. Participants were also excluded if they scored 3.3 or more on the depression subscale (Fares, Dirani, & Darwish, 2019; Mahfoud et al., 2013; Winokur, Winokur, Rickels, & Cox, 1984).

The Montreal Cognitive Assessment (MoCA) was used post-consenting to screen for cognitive impairment (Nasreddine et al., 2005; Rahman & El Gaafary, 2009). A cutoff score of 26 for individuals below the age of 60 years and 24 for those 60 years or older was used (Carson, Leach, & Murphy, 2018). In other words, individuals below the age of 60 years were excluded if they scored less than 26, while those older than 60 years were excluded if they scored below 24.

MS patients were recruited from the American University of Beirut Medical Center MS center. MS patients were diagnosed according to the McDonald 2010 criteria by a neurologist (Polman et al., 2011). This information was checked in the patient's medical record. Only patients with a disease duration greater than 1 year were enrolled in this study. This was established through reviewing the patient's date of symptom appearance. Following the same eligibility criteria of the healthy subjects, 43 MS patients were matched to 43 healthy participants based on age, sex, and years of education on an individual level. (Table 1, Part A). The bands used were the following: age +/- 3 years, 1:1 ratio on sex, and +/- 2 years on education. We did not exclude MS patients with cognitive deficits or depression given its high prevalence in MS patients and importance for examining the discriminative abilities of the tests.

PROCEDURE

Data Collection

Training for data collectors

We trained two data collectors who were either in their senior year or holders of an undergraduate degree in psychology. Training consisted of a 3-hr practical workshop that included a live demonstration of flawed and flawless administrations. Then, the data collectors were observed in three mock administrations, until no errors were detected. Authors PZ and NF supervised the training and ongoing data collection; NF also collected data and was previously trained by PZ. All tests were scored by NF.

All tests were administered in a standardized manner, in a quiet room, using the Lebanese Arabic dialect. Tests were also performed in a fixed order, beginning with screening, including the BRFSS and HSCL-25, and next, the SDMT, and the VMAT learning trials and short delay recalls, and then by the BVMT-R. Other segments of the verbal memory test were performed last (25 min post-VMAT short delay recalls). Hence, the total duration of the BICAMS administration was around 55 min.

The oral version of the SDMT (Smith, 1982) was administered to all participants. Using a test form that contains a 9 symbol-digit paired (key) with a sequence of symbols (stimuli), the participant was required to respond by voicing the digit associated with each symbol as quickly as possible. A sequence of 10 symbols is first used for practice. Then the participant was given 90 s to complete as many items as possible present in the form after the practice items (Smith, 1982). The dependent variable was the number of correct responses in 90 s.

For the BVMT-R (Benedict, 1997), participants were asked to recall a matrix of six simple abstract designs after 10 s of visual exposure. The participants reproduced the designs using paper and pencil as accurately as possible in their correct positions. In total, the test was repeated three times for each participant. Each figure received a score of 0, 1, or 2, based on accuracy and location scoring criteria (Benedict, 1997). The dependent variable was the total score across the three trials.

We recently developed and validated a VMAT (Zeinoun et al., 2020), which substituted the CVLT-II in our study. The VMAT was developed indigenously in Arabic using quantitative and qualitative methods. Following a rigorous process, words that are more or less familiar to all Arab regions were selected during the development of the VMAT to facilitate use in other Arab countries. The instrument measures verbal learning, short-term memory, long-term memory, and recognition. Similar to other standardized verbal learning/memory tests, and in line with Benedict et al. (2012) recommendations, the examinee is presented with 15 words (List A) to be recalled freely, across 5 trials, and is then presented by another 15 words (List B) which serve as an interference trial. Following the recall of List B, the participant was required to recall List A with and without semantic cues. Following a 25-min delay, the test-taker was required to recall List A with and without cues and then recognize the words from List A from an array of 45 words that include List A, List B, and additional distractors. Several scores could be derived from the VMAT such as the number of words recalled per trial, the total number of words recalled on trials 1 to 5, in addition to a recognition discriminability index (i.e., the ability to endorse the 15 target items and reject all 30 distractors) (Zeinoun et al., 2020). However, the regression model and normative data included only the total number of words recalled in trials 1 to 5 as dependent variables. We chose this variable for the VMAT based on the methods of other studies

Table 1. Differences in BICAMS performance between MS patients matched with healthy individuals

BICAMS measure	Healthy group (n = 43)		MS group (n = 43)		Group comparisons		
	Median	Range†	Median	Range†	U	Cohen's d‡	p
SDMT	60	42	53	66	583.5	0.67	0.003
BVMT-R (T1 to T3)	25	25	21	30	650.5	0.53	0.018
VMAT							
Trial 1	7	8	6	9	851		0.521
Total trials 1 to 3	28	19	27	22	863.5		0.598
Total trials 1 to 5	54	29	52	38	861		0.583
Learning slope*	1	2	1	2	878		0.687
Short delay free	12	10	11	14	802		0.287
Short delay cued	12	10	11	10	792		0.249
Long delay free	12	9	11	10	822.5		0.375
Long delay cued	12	9	11	10	866		0.611
Recognition	44	9	44	12	831.5		0.409
Recognition discriminability	3	2.54	3	2.35	889.5		0.759

*Learning slope of trials 1 to 5.

† The difference between the lowest value and the highest value.

‡ Value transformed based on Cohen's 1998 and 2008 recommendations, to meet Mann–Whitney U test statistics requirements. Results of measures with significant differences between groups are reported.

(Filser et al., 2018; Ozakbas et al., 2017; Polychroniadou et al., 2016; Sousa et al., 2018).

Medical charts were also reviewed to collect the MS patients' disease type, disease duration, and Expanded Disability Status Scale score.

Normative Values

Regression-based norms were calculated following the previously described procedure applied for the MACFIMS (Parmenter, Testa, Schretlen, Weinstock-Guttman, & Benedict, 2010), which has recently begun to be utilized for the BICAMS (Goretti et al., 2014). To ensure the normal distribution of the raw test scores of healthy participants, we have first retrieved the cumulative frequency distribution of the SDMT, BVMT-R, and VMAT score of trials 1 to 5. The resulting distribution was converted into a standard scaled score with a mean (M) of 10 and a standard deviation (SD) of 3 (actual scaled score). Next, regression equations for the predicted scaled scores were modeled; stepwise regression analyses were performed including age, age^2, sex (1 = male, 2 = female), and years of education. A squared term of age was used to adjust for the nonlinear relationship between age and cognition (Goretti et al., 2014; Parmenter et al., 2010). We used stepwise regression based on Parmenter et al. (2010), MACFIMS, and Goretti et al. (2014). We performed a forced entry analysis to compare the results. The derived equations of the statistically significant models include unstandardized β-coefficients of the predictors and the constant. Both assumptions of homoscedasticity and normality of residuals were evaluated and met.

These regression models were used to generate normative values. First, the equations using specific demographic information were derived. Next, the predicted scaled score was subtracted from the actual scaled score. The difference was then divided by the residual SD of the healthy group tests. Finally, the derived value could be converted to other standardized scores, such as Z scores, to classify performance (Goretti et al., 2014).

Statistical Analysis

Descriptive statistics were calculated; mean (M), standard deviation (SD) and median for continuous data, and frequencies and percentages for categorical ones. For the test–retest reliability analysis, Pearson's correlations between test scores on both sessions were calculated (coefficient: r). When data violated the assumption of normality, the nonparametric alternative Spearman's Rho was used.

To evaluate whether the BICAMS score could differentiate between known groups membership, scores were compared between matched healthy subjects and MS patients using Mann–Whitney U-test, since the data were significantly skewed. Effect sizes were computed for variables that discriminated between the groups.

To derive normative values, a stepwise multiple regression analyses for each of the dependent variables, SDMT, BVMT-R, and VMAT (total score on trials 1 to 5), were conducted with age, age^2, sex, and years of education entered as predictors (see the previous section on "normative values" for more details). Regression analyses using forced entry method were also run to compare results – similar outcomes were obtained. All analyses were performed on SPSS version 25, a two-tailed test, and results with $p < 0.05$ were considered statistically significant.

RESULTS

Descriptives

Two hundred and thirty-four healthy participants from the community were screened for eligibility, 54 were excluded, and 2 from the MS group (psychiatric illnesses). Figure 1 is a flow chart of participants' recruitment. Table 2 summarizes demographic information of the full healthy sample, in addition to the MS group.

Healthy Individuals (Full Sample)

The average age of healthy individuals ($n = 180$) was 45.01 ± 19.36 years. The youngest participant was 16 and the oldest was 80 years old, and Table 3 reports the scores received on the different tests of BICAMS. The 63 individuals who were retested had a mean age of 33.15 ± 15.98 years, similar sex distribution (54% males), and high educational attainment (71.4% completed university).

MS Patients and the Matched Healthy Sample

Regarding the MS patients, mean MoCA scores were 25.15 ± 2.81 where 25 patients (61%) scored less than 26 on the MoCA, suggestive of cognitive impairment. Symptoms of depression scores were 1.91 ± 0.61; 27 patients showed low symptoms of depression (64.3%), 15 moderate levels (35.7%), and none with severe symptoms. Mean years of education were 14.63 ± 3.17 years.

Concurrently, the 43 healthy individuals matched with the 43 MS patients had a mean age of 36.7 ± 13.06 with 8 males and 35 females constituting the sample. Average years of education were 15.13 ± 3 years. Symptoms of depression scores were 1.52 ± 0.38, and MoCA scores were 28.19 ± 1.20.

There were no significant differences between the two matched samples on age ($t = 0.234$, $p = 0.816$) and education ($t = 0.751$, $p = 0.455$). There were also equal numbers of males and females. Nevertheless, the samples differed based on symptoms of depression scores ($t = -3.56$, $p = 0.001$) and MoCA ($t = 5.85$, $p < 0.001$).

MS Patients Retested

The mean age of the 10 MS patients who were subject to test–retest was 38.71 ± 12.87 years, and most were females ($n = 8$). Six individuals were diagnosed with RRMS, three with SPMS, and one with PPMS. Most of this subsample completed university education ($n = 7$), one completed high school, one obtained vocational education, and one had some high school education.

Reliability (Test–Retest)

Among the healthy participants, the test–retest agreement on the SDMT and BVMT-R were 73 and 60%, respectively ($p < 0.001$). SDMT scores of the subsample were 59.22 ± 12.27 on test, and 63.62 ± 12.45 on retest, while BVMT-R scores were 24.23 ± 6.66 on test and 29.87 ± 4.81 on retest. The test–retest reliability was good on the VMAT, total learning trials (1 to 5) $r = 0.69$ (test $= 54.10 \pm 8.71$, retest $= 61.9 \pm 8.26$). On the VMAT short delay segments, and the free and cued recall trials, the Spearman's Rho values were 0.64 (test $= 11 \pm 2.46$, retest $= 12.81 \pm 2.06$), and 0.73 (test $= 11.33 \pm 2.31$, retest $= 12.81 \pm 2.09$), respectively. Similar results were obtained on the long delay segments – free and cued recall Rho $= 0.60$ and 0.70, respectively, with free recall test scores equal to 11.51 ± 2.26, and retest scores equal to 13.11 ± 2.18, and cued recall test scores equal to 11.71 ± 2.43, and retest scores equal to 12.97 ± 2.27. The lowest test–retest reliability was on the VMAT recognition trial (Rho $= 0.39$, test $= 43.52 \pm 1.64$, retest $= 44.14 \pm 1.38$).

Among the MS patients, the test–retest reliability on both the SDMT and BVMT-R was excellent (Rho $= 0.92$, and $r = 0.83$, respectively), with SDMT test scores $= 47.2 \pm 17.98$, and retest $= 47.3 \pm 15.79$, as well as BVMT-R $t = 22 \pm 9.79$, while the $r = 26 \pm 8.79$. On the VMAT short delay free recall and cued recall, the test–retest agreement was fair, 59 and 43% respectively, as opposed to higher values on the total score of the learning trials 1 to 5 (Rho $= 0.64$), and the long delay segments (free $= 0.68$, cued $= 0.67$). During test period, MS patients scored 5 on the learning trials 56.9 ± 10.04, short delay free 11.1 ± 4.98, short delay cued 12.5 ± 3.21, long delay free 13.2 ± 2.3, and on long delay cued 13.1 ± 2.56. Alternatively, during retest, the MS patients scored on the learning trials 65.8 ± 8.7, short delay free 13.1 ± 2.28, short delay cued 13.9 ± 1.73, long delay free 13.5 ± 2.12, and long delay cued 13.7 ± 1.95. The highest test–retest reliability coefficient was on the recognition trial, Rho $= 0.92$ (test $= 43 \pm 3.3$, retest $= 43.6 \pm 2.27$).

Criterion-Related Validity (Group Differences)

Forty-three MS patients were matched with 43 healthy individuals. MS patients scored lower than healthy participants on SDMT, BVMT-R, VMAT trials 1 to 5, short delay-free and cued, as well as long delay-free and cued. The SDMT and BVMT-R discriminated the most between the groups with a larger effect size for the SDMT (Table 1).

Because the difference in scores between the groups on the VMAT did not reach statistical significance, for further validation, we examined the scores of the MS patients on the verbal memory section of the MoCA (total score 5), and we found a higher than average mean score of 3.31 ± 1.07.

Normative Values

Our data suggest that age and education were the strongest demographic contributors to test performance. Specifically, age or age^2 and education predicted the SDMT and BVMT-R scores. Age^2 mainly predicted the VMAT, total
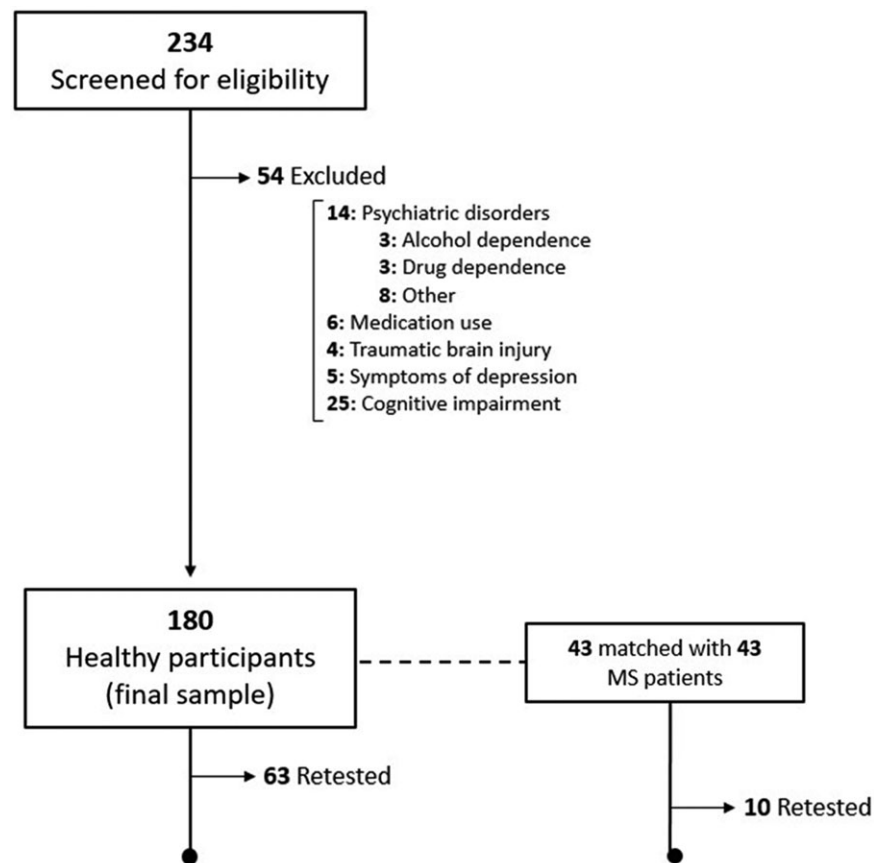
Fig. 1. Flow chart of participants' recruitment.

scores on trials 1 to 5. Table 4 reports the raw to scaled scores ($M = 10$, $SD = 3$) conversion using the BICAMS cumulative frequency distribution derived from our sample (Part A) and the results of the regression models, which were statistically significant (Part B). The derived equations used to compute predicted scaled scores are also listed in the table, alongside an example on how to apply them. To facilitate the adoption and usage of the proposed norms in the clinical setting, the reader can access an Excel file with built-in formulas through https://arabicbicams.wixsite.com/website to calculate the following parameters: actual scaled score, predicted scaled score, Z score, and t score.

DISCUSSION

This study contributes to the international utility of the BICAMS by providing partial evidence of validity and fair evidence of reliability for the Arabic version of the battery. The SDMT and BVMT-R tests mostly showed good psychometric properties. Here, we followed the recommendations and standards of the BICAMS consensus committee (Benedict et al., 2012). We also provide Lebanese normative data.

Although the BICAMS does not replace a more comprehensive evaluation of cognitive function in MS, it is valuable as a brief tool that can be integrated into broader assessments.

The availability of BICAMS increases accessibility for cognitive assessments in nonspecialized centers (Langdon et al., 2012). Translation and validation of the BICAMS were performed in many regions of the world and languages (Corfield & Langdon, 2018), and normative data were established for populations in several countries, such as Italy (Goretti et al., 2014) and Greece (Polychroniadou et al., 2016).

In this study, the BICAMS showed good evidence of test–retest reliability, which is essential for longitudinal assessments of cognition in MS. The SDMT, in particular, whereby different test forms and retest sessions were utilized, was higher than other measures. This result is in line with other studies (Benedict, 2005; Goretti et al., 2014; Sousa et al., 2018).

The VMAT showed fair reliability results (including the total learning trials 1 to 5 score). Regarding this main outcome variable, results are comparable to other BICAMS studies, which provided evidence of validity (Polychroniadou et al., 2016; Sousa et al., 2018). The highest values were on the cued recall trial. It should be noted that while good test–retest results were present for the healthy group, weaker results were found in the MS group in terms of the short delay free and cued recall trials. The use of one form of the VMAT in our study most likely elevated practice effects, thus hindering reliability analyses. We are currently developing alternative forms of the test and recruiting a larger and more diverse sample for additional psychometric validation.

Table 2. Demographic information of the full healthy sample and MS patients

	Healthy group (n = 180)		MS group (n = 43)	
	M	SD	M	SD
Age	45.01	19.36	36.06	12.37
Sex	F	%	F	%
Male	72	40	8	18.6
Female	108	60	35	81.4
Education				
Literate with no schooling	2	1.11	0	0
Elementary to intermediate	19	10.56	3	7
Some high school	11	6.11	3	7
Completed high school	28	15.56	10	23.3
Vocational education	11	6.11	4	9.3
University	109	60.56	23	53.3
Current employment status				
Employed	101	56.11	21	48.84
Unemployed	79	43.89	22	51.16
Marital status				
Married	77	42.78	21	48.84
In a relationship	14	7.78	1	2.33
Divorced, separated, or widowed	22	12	5	11.63
Single	67	37.22	16	37.21
Physical activity				
None	50	27.78	31	72.09
1–5 times per month	7	3.89	0	0
1–2 times per week	38	21.11	4	9.3
3–4 times per week	31	17.22	7	16.28
5 or more times per week	54	30	1	2.33
Cigarette smoking				
Smoker	51	28.33	11	25.58
Nonsmoker	129	71.67	32	74.42
MS clinical characteristics				
MS type				
Relapsing remitting			35	81.4
Secondary progressive			6	13.95
Primary progressive			2	4.65
			M	SD
Disease duration (years)			8.61	7.36
Expanded disability status scale			1.89	1.7

Table 3. BICAMS scores of healthy participants (n = 180)

Test	Min	Max	M	SD
SDMT (1)				
Correct responses	13	86	54.48	13.97
BVMT-R (4)				
Trial 1	0	11	4.62	2.45
Trial 2	1	12	7.69	2.80
Trial 3	1	12	9.46	2.62
Total	3	34	21.77	7.17
VMAT (13)				
Trial 1	1	12	6.57	2.00
Trial 2	4	15	9.54	2.32
Trial 3	4	15	10.81	2.28
Trial 4	4	15	11.58	2.27
Trial 5	4	15	11.88	2.25
Total	21	72	50.41	9.66
Interference	1	13	6.14	2.23
Short delay free	3	15	10.20	2.85
Short delay cued	5	15	10.62	2.62
Long delay free	4	15	10.64	2.68
Long delay cued	4	15	10.87	2.71
Recognition	31	45	42.56	2.70
Discriminability	0.73	4	3.14	0.75

well. In Zeinoun et al. (2020), results indicated that the VMAT could significantly discriminate between MS patients and healthy individuals on various subscores. In our current study, this significance in discrimination was absent – although MS patients scored lower than healthy individuals. We partly attribute this discrepancy in results to the different sampling methods and performance of the patients included in the studies. In the current study, 43 patients were matched with 43 healthy individuals, as opposed to the original VMAT study, in which 16 MS patients were matched with 32 healthy individuals. The scores of the healthy participants groups used in both studies were similar but differed for MS patients. In this study, the MS patients' scores on the VMAT were higher; for example, on the long delay free recall, MS patients in the original VMAT study scored 8.62 ± 3.07, while in the current study, they scored 10.83 ± 3.2. It could suggest that the examined MS sample had no or little impairment in verbal memory, which was confirmed when examining their performance on the verbal section of the MoCA. The majority scored above average. The better performance of the MS group in the current study when compared to the original VMAT study could have contributed to losing statistical significance when examining the discrimination of the test.

In essence, we encourage the use of the battery in Arab-speaking MS patients. Nevertheless, caution should be exercised in reference to the VMAT. In particular, while the full battery might be useful during an initial assessment of the patient, we advice to administer the battery again following a substantial period of time (such as 6 months). Alternatively, we are in the process of building and examining an alternative

The BVMT-R, on the other hand, showed slightly lower evidence of reliability in the healthy sample. This could be due to several participants reaching high ceiling scores on the first testing sessions, while others significantly increased their performance toward the second testing session. Nonetheless, the BVMT-R is known to be adequate for international use (Benedict et al., 2012) and showed good test–retest reliability measures among persons with MS.

The matched MS sample performed lower than the healthy participants on the Arabic BICAMS, but only SDMT and BVMT-R could significantly discriminate between the samples. We compared the current results on the VMAT with our prior work using the same tool in a smaller MS sample as

Table 4. Arabic BICAMS parameters for normative values

Measure	Part A. Raw to scaled score conversions																	Part B. Final regression models						
	Scaled scores (2–18)																							
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Predictor	β	SE B	t	Standardized B	R^2	Residual SD
SDMT	13	14–15	16–18	19–29	30–37	38–42	43–46	47–51	52–55	56–59	60–65	66–68	69–72	73–77	78–80	81–86	>86	(constant)	11.57	0.68	17.145		0.48	2.132
																		Age^2	−0.001	0	−10.13	−0.62		
																		Education	0.082	0.04	2.181	0.134		
BVMTR		4	5–6	7–8	9–10	11–14	15–17	18–21	22	23–25	26–27	28–29	30–31	32	33–34		>34	(constant)	12.12	0.91	13.286		0.276	2.458
																		Age	−0.067	0.01	−6.33	−0.45		
																		Education	0.085	0.04	2	0.142		
VMAT (T1 to T5)	21	22	23–28	29–33	34–37	38–41	42–43	44–48	49–51	52–54	55–57	58–60	61–63	64–65	66–68	69–70	>70	(constant)	12.28	0.32	38.841		0.249	2.575
																		Age^2	−0.001	0	−7.66	−0.499		

Equations derived from Part B to convert MS raw scores to regression-based *t* scores or any other type of standardized scores:

SDMT scaled score$_{predicted}$ = 11.569 − 0.001 (age^2) + 0.082 (education).

BVMT-R scaled score$_{predicted}$ = 12.121 − 0.067 (age) + 0.085 (education).

VMAT (trials 1 to 5) scaled score$_{predicted}$ = 12.282 − 0.001 (age^2).

An example on how to apply the normative values:

Consider a 40-year-old female MS patient with 16 years of education. This patient scored 42 on the SDMT. From Part A, we know that her SDMT raw score corresponds to an actual scaled score of 7, and the predicted scaled score is 11.281 based on the formula provided above [11.569 − 0.001 (402) + 0.082 (16)]. Next, to deduce the Z score, we subtract the predicted scaled score from the patient's actual scaled score and then divide the difference by the residual *SD* of the healthy participants; residual *SD* is provided in Part B. Here, we have (7 − 11.281) ÷ 2.132, which equals to − 2 (i.e., *t* score: 30) and suggests that her SDMT performance can be classified as impaired (Z ≤ -2 *SD*). To facilitate the adoption and usage of the proposed norms in the clinical setting, the corresponding author, upon request, can provide a spreadsheet table with a built-in formula that calculates and yields the above parameters.

version of the VMAT which might yield more reliable results in the clinical or research settings.

In this study, we also provide regression-based norms for a Lebanese sample. This approach has several advantages over conventional norming methods, such as linearly adjusting covariates. Also, it applies to smaller samples (Oosterhuis, van der Ark, & Sijtsma, 2016). Although the scores and the impact of a few demographic variables on test results were similar to other studies (Goretti et al., 2014; Vanotti, Smerbeck, Benedict, & Caceres, 2016), not all variables contributed equally. This is not surprising in light of cultural differences in testing. The current results, age, and education as main predictors of cognitive performance are similar to findings from another study in the same country where we validated and normed a visual memory test (Rey Complex Figure Test) (Darwish, Zeinoun, Farran, & Fares, 2018). Furthermore, the Portuguese BICAMS validation study and regression-based norms reached similar conclusions (Sousa et al., 2018). This highlights the importance of national validation of cognitive tests across cultures as subtle differences in test performance could be present.

Several limitations need to be taken into account when interpreting our findings. The majority of the MS sample were diagnosed with RRMS, and the sex distribution was not balanced. Further evidence on the validity of the new Arabic version of the BICAMS with a more diverse profile and a larger sample will be pursued in future studies. Along a similar vein, examining external validity can potentiate a wider use of the battery, especially in other Arab countries.

Lastly, the frequency of individuals who attained a university degree in this study was higher in younger individuals when compared to older participants. This is in line with the age distribution, literacy, and education rates in Lebanon. The literacy rate is 99.24% for those between the ages of 15 and 24 years, and 60.15% for the 65 years and older. Sixteen percent of the population is between the ages 15 and 24 years, the majority, 45.27%, are between the ages 25 and 54, 8.3 and 7% are 55–64 and 65 years or older, respectively, with around 11–12 years of school life expectancy (CIA, 2020). Also, in 2017, it was reported that 93% of the population finished primary education, 63–70% secondary education, and 45–49% tertiary education (UNESCO, 2017).

The current study shows encouraging psychometric properties of a new Arabic version of the BICAMS and provides regression-based norms for a Lebanese sample. The findings and data presented can enhance MS-related clinical practice in this region, and we encourage the use of this battery in both research and clinical settings.

ACKNOWLEDGMENT

CONFLICT OF INTEREST

The authors have nothing to disclose.

REFERENCES

Amato, M.P., Zipoli, V., & Portaccio, E. (2006). Multiple sclerosis-related cognitive changes: A review of cross-sectional and longitudinal studies. *Journal of the Neurological Sciences*, *245*(1–2), 41–46.

Amato, M.P., Zipoli, V., & Portaccio, E. (2008). Cognitive changes in multiple sclerosis. *Expert Review of Neurotherapeutics*, *8*(10), 1585–1596.

Benedict, R. (1997). *Brief Visual Memory Test-Revised: Professional Manual*. Odessa, FL: Psychological Assessment Resources.

Benedict, R.H. (2005). Effects of using same-versus alternate-form memory tests during short-interval repeated assessments in multiple sclerosis. *Journal of the International Neuropsychological Society: JINS*, *11*(6), 727.

Benedict, R.H., Amato, M.P., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., . . . Penner, I. (2012). Brief International Cognitive Assessment for MS (BICAMS): International standards for validation. *BMC Neurology*, *12*(1), 55.

Benedict, R.H., Cookfair, D., Gavett, R., Gunther, M., Munschauer, F., Garg, N., & Weinstock-Guttman, B. (2006). Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *Journal of the International Neuropsychological Society: JINS*, *12*(4), 549.

Benedict, R.H., Fischer, J.S., Archibald, C.J., Arnett, P.A., Beatty, W.W., Bobholz, J., . . . Caruso, L. (2002). Minimal neuropsychological assessment of MS patients: A consensus approach. *The Clinical Neuropsychologist*, *16*(3), 381–397.

Bobholz, J.A. & Rao, S.M. (2003). Cognitive dysfunction in multiple sclerosis: A review of recent developments. *Current Opinion in Neurology*, *16*(3), 283–288.

Brassington, J.C. & Marsh, N.V. (1998). Neuropsychological aspects of multiple sclerosis. *Neuropsychology Review*, *8*(2), 43–77.

Calabrese, P. (2006). Neuropsychology of multiple sclerosis. *Journal of Neurology*, *253*(1), i10–i15.

Carson, N., Leach, L., & Murphy, K.J. (2018). A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. *International Journal of Geriatric Psychiatry*, *33*(2), 379–388.

Centers for Disease Control and Prevention. (2009). *Behavioral Risk Factor Surveillance System Survey Questionnaire*. Atlanta, Georgia: US Department of Health and Human Services.

Cheng, E.M., Crandall, C.J., Bever, C.T., Jr., Giesser, B., Haselkorn, J.K., Hays, R.D., . . . Vickrey, B.G. (2010). Quality indicators for multiple sclerosis. *Multiple Sclerosis Journal*, *16*(8), 970–980.

Chiaravalloti, N.D. & DeLuca, J. (2008). Cognitive impairment in multiple sclerosis. *The Lancet Neurology*, *7*(12), 1139–1151.

CIA. (2020). *The World Fact Book: Lebanon*. Washington, DC: Central Intelligence Agency. https://www.cia.gov/the-world-factbook/; https://www.cia.gov/Library/publications/the-world-factbook/geos/le.html

Clemens, L. & Langdon, D. (2018). How does cognition relate to employment in multiple sclerosis? A systematic review. *Multiple Sclerosis and Related Disorders*, *26*, 183–191.

Corfield, F. & Langdon, D. (2018). A systematic review and meta-analysis of the brief cognitive assessment for multiple sclerosis (BICAMS). *Neurology and Therapy*, *7*(2), 287–306.

Darwish, H., Zeinoun, P., Farran, N., & Fares, S. (2018). Rey Figure Test with recognition trial: normative data for Lebanese adults. *The Clinical Neuropsychologist*, *32*(Suppl 1), 102–113.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (2000). *California Verbal Learning Test – second edition. Adult Version Manual*. San Antonio, TX: Psychological Corporation.

Deloire, M.S., Ruet, A., Hamel, D., Bonnet, M., Dousset, V., & Brochet, B. (2011). MRI predictors of cognitive outcome in early multiple sclerosis. *Neurology*, *76*(13), 1161–1167.

El Ghoneimy, A.T., Hassan, A., Homos, M., Farghaly, M., & Dahshan, A. (2015). Thalamic involvement and its impact on disability and cognition in multiple sclerosis: A clinical and diffusion tensor imaging study. *Egyptian Journal of Neurology Psychiatry and Neurosurgery*, *52*(2), 139–145.

Fares, S., Dirani, J., & Darwish, H. (2019). Arabic validation of the hopkins symptom checklist-25 (HSCL) in a Lebanese sample of adults and older adults. *Current Psychology*, 1–8.

Filser, M., Schreiber, H., Pöttgen, J., Ullrich, S., Lang, M., & Penner, I.-K. (2018). The brief international cognitive assessment in multiple sclerosis (BICAMS): Results from the German validation study. *Journal of Neurology*, *265*(11), 2587–2593.

Goretti, B., Niccolai, C., Hakiki, B., Sturchio, A., Falautano, M., Minacapelli, E., . . . Murgia, M. (2014). The brief international cognitive assessment for multiple sclerosis (BICAMS): Normative values with gender, age and education corrections in the Italian population. *BMC Neurology*, *14*(1), 171.

Goverover, Y., Chiaravalloti, N., & DeLuca, J. (2016). Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) and performance of everyday life tasks: Actual reality. *Multiple Sclerosis Journal*, *22*(4), 544–550.

Goverover, Y., Genova, H. M., Hillary, F. G., & DeLuca, J. (2007). The relationship between neuropsychological measures and the Timed Instrumental Activities of Daily Living task in multiple sclerosis. *Multiple Sclerosis Journal*, *13*(5), 636–644.

Hamdy, S., Shaheen, H.A., Aboumousa, A.M., Farghaly, M., Ezzat, L.M., & Daker, L. (2013). Does the disease course or treatment type have impact on executive functions and cognition in multiple sclerosis patients? A clinical and 3 tesla MRI study. *Egyptian Journal of Neurology, Psychiatry & Neurosurgery*, *50*(4), 389–396.

Kalb, R., Beier, M., Benedict, R.H., Charvet, L., Costello, K., Feinstein, A., . . . Harris, C. (2018). Recommendations for cognitive screening and management in multiple sclerosis care. *Multiple Sclerosis Journal*, *24*(13), 1665–1680.

Kishk, N., Shalaby, N., Shehata, H., Hassan, A., Hegazy, M., Elmazny, A., . . . Farghaly, M. (2017). Reliability of BICAMS (Arabic version) in Egyptian multiple sclerosis patients. *Multiple Sclerosis Journal*, *23*, 155–156.

Langdon, D., Amato, M., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., . . . Penner, I. (2012). Recommendations for a brief international cognitive assessment for multiple sclerosis (BICAMS). *Multiple Sclerosis Journal*, *18*(6), 891–898.

Langdon, D.W. (2011). Cognition in multiple sclerosis. *Current Opinion in Neurology*, *24*(3), 244–249.

Mahfoud, Z., Kobeissi, L., Peters, T., Araya, R., Ghantous, Z., & Khoury, B. (2013). The Arabic validation of Hopkins Symptoms Checklist-25 against MINI in a disadvantaged suburb of Beirut, Lebanon. *The International Journal of Educational and Psychological Assessment (TIJEPA)*, *13*(1), 17–33.

McGuire, S. (2011). *US Department of Agriculture and US Department of Health and Human Services, Dietary Guidelines*

for Americans, 2010. Washington, DC: US government printing office.

Mitchell, A.J., Benito-León, J., González, J.-M.M., & Rivera-Navarro, J. (2005). Quality of life and its assessment in multiple sclerosis: Integrating physical and psychological components of wellbeing. *The Lancet Neurology*, *4*(9), 556–566.

Moccia, M., Lanzillo, R., Palladino, R., Chang, K.C.-M., Costabile, T., Russo, C., … Maniscalco, G.T. (2016). Cognitive impairment at diagnosis predicts 10-year multiple sclerosis progression. *Multiple Sclerosis Journal*, *22*(5), 659–667.

Nasreddine, Z.S., Phillips, N.A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699.

Oosterhuis, H.E., van der Ark, L.A., & Sijtsma, K. (2016). Sample size requirements for traditional and regression-based norms. *Assessment*, *23*(2), 191–202.

Ozakbas, S., Yigit, P., Cinar, B.P., Limoncu, H., Kahraman, T., & Kösehasanoğulları, G. (2017). The Turkish validation of the brief international cognitive assessment for multiple sclerosis (BICAMS) battery. *BMC Neurology*, *17*(1), 1–6.

Parmenter, B.A., Testa, S.M., Schretlen, D.J., Weinstock-Guttman, B., & Benedict, R.H. (2010). The utility of regression-based norms in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *Journal of the International Neuropsychological Society: JINS*, *16*(1), 6.

Paul, S., Brown, A., & Hughes, A.J. (2019). Cognitive assessment of individuals with multiple sclerosis in the Arab world: A systematic review. *Neuropsychology Review*, *29*(3), 259–269.

Piras, M.R., Magnano, I., Canu, E.D.G., Paulus, K.S., Satta, W.M., Soddu, A., … Aiello, I. (2003). Longitudinal study of cognitive dysfunction in multiple sclerosis: Neuropsychological, neuroradiological, and neurophysiological findings. *Journal of Neurology, Neurosurgery & Psychiatry*, *74*(7), 878–885.

Pitteri, M., Romualdi, C., Magliozzi, R., Monaco, S., & Calabrese, M. (2017). Cognitive impairment predicts disability progression and cortical thinning in MS: An 8-year study. *Multiple Sclerosis Journal*, *23*(6), 848–854.

Polman, C.H., Reingold, S.C., Banwell, B., Clanet, M., Cohen, J.A., Filippi, M., … Kappos, L. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology*, *69*(2), 292–302.

Polychroniadou, E., Bakirtzis, C., Langdon, D., Lagoudaki, R., Kesidou, E., Theotokis, P., … Boziki, M. (2016). Validation of the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) in Greek population with multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *9*, 68–72.

Prosiegel, M. & Michael, C. (1993). Neuropsychology and multiple sclerosis: diagnostic and rehabilitative approaches. *Journal of the Neurological Sciences*, *115*, S51–S54.

Rahman, T.T.A. & El Gaafary, M.M. (2009). Montreal Cognitive Assessment Arabic version: Reliability and validity prevalence of mild cognitive impairment among elderly attending geriatric clubs in Cairo. *Geriatrics & Gerontology International*, *9*(1), 54–61.

Rao, M. and The Cognitive Function Study Group of the National Multiple Sclerosis Society. (1990). A manual for the Brief Repeatable Battery of Neuropsychological Tests in multiple sclerosis Medical College of Wisconsin: Milwaukee, WI.

Rao, S.M. (1995). Neuropsychology of multiple sclerosis. *Current Opinion in Neurology*, *8*(3), 216–220.

Rao, S.M., Leo, G.J., Bernardin, L., & Unverzagt, F. (1991). Cognitive dysfunction in multiple sclerosis: I. Frequency, patterns, and prediction. *Neurology*, *41*(5), 685–691.

Rocca, M.A., Amato, M.P., De Stefano, N., Enzinger, C., Geurts, J.J., Penner, I.-K., … Filippi, M. (2015). Clinical and imaging assessment of cognitive dysfunction in multiple sclerosis. *The Lancet Neurology*, *14*(3), 302–317.

Smerbeck, A., Benedict, R.H., Eshaghi, A., Vanotti, S., Spedo, C., Blahova Dusankova, J., … Langdon, D. (2018). Influence of nationality on the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). *The Clinical Neuropsychologist*, *32*(1), 54–62.

Smith, A. (1982). *Manual for the Symbol Digit Modalities Test*. Los Angeles: Western Psychological Services.

Sousa, C., Rigueiro-Neves, M., Miranda, T., Alegria, P., Vale, J., Passos, A.M., … Sá, M.J. (2018). Validation of the brief international cognitive assessment for multiple sclerosis (BICAMS) in the Portuguese population with multiple sclerosis. *BMC Neurology*, *18*(1), 1–7.

Strober, L., Englert, J., Munschauer, F., Weinstock-Guttman, B., Rao, S., & Benedict, R. (2009). Sensitivity of conventional memory tests in multiple sclerosis: Comparing the Rao brief repeatable neuropsychological battery and the minimal assessment of cognitive function in MS. *Multiple Sclerosis Journal*, *15*(9), 1077–1084.

Sumowski, J.F., Benedict, R., Enzinger, C., Filippi, M., Geurts, J.J., Hamalainen, P., … Rocca, M.A. (2018). Cognition in multiple sclerosis: State of the field and priorities for the future. *Neurology*, *90*(6), 278–288.

Trenova, A.G., Slavov, G.S., Manova, M.G., Aksentieva, J.B., Miteva, L.D., & Stanilova, S.A. (2016). Cognitive impairment in multiple sclerosis. *Folia Medica*, *58*(3), 157–163.

UNESCO. (2017). *Education and Literacy: Lebanon*. Montreal, Canada: UNESCO Institute for Statistics, Sustainable Development Goals. http://uis.unesco.org/en/country/lb

Vanotti, S., Smerbeck, A., Benedict, R.H., & Caceres, F. (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. *The Clinical Neuropsychologist*, *30*(7), 1023–1031.

Winokur, A., Winokur, D.F., Rickels, K., & Cox, D.S. (1984). Symptoms of emotional distress in a family planning service: Stability over a 4-week period. *The British Journal of Psychiatry*, *144*(4), 395–399.

Zeinoun, P., Farran, N., Khoury, S.J., & Darwish, H. (2020). Development, psychometric properties, and pilot norms of the first Arabic indigenous memory test: The Verbal Memory Arabic Test (VMAT). *Journal of Clinical and Experimental Neuropsychology*, *42*(5), 505–515.