*© 2022. This manuscript version is made available under the CC-BY-NC-ND 4.0 license http://creativecommons.org/licenses/by-nc-nd/4.0/*

**Validation of the Norma Latina neuropsychological assessment battery in individuals with multiple sclerosis in Mexico**

**Diego Rivera, Ph.D**

Department of Health Sciences, Public University of Navarre, Pamplona, Spain

ORCID: 0000-0001-7477-1893

**Daniela Ramos Usuga**

BioCruces Bizkaia Health Research Institute, 48903 Barakaldo, Spain.

Biomedical Research Doctorate Program, University of the Basque Country, Leioa, Spain

**Eliana María Fuentes Mendoza**

Maestría en Psicometría y Evaluación Educativa, Universidad Nacional Autónoma de Honduras, Tegucigalpa, Honduras.

**Adriana Aguayo Arelis. Ph.D**

Departamento de neurociencias, centro universitario de ciencias de la salud. Universidad de Guadalajara. Fundación Mexicana Para la Esclerosis Múltiple A.C.

**Brenda Viridiana Rabago Barajas. Ph.D**

Departamento de neurociencias, centro universitario de ciencias de la salud. Universidad de Guadalajara. Fundación Mexicana Para la Esclerosis Múltiple A.C.

**Miguel Ángel Macías Islas. M.D**

Departamento de neurociencias, centro universitario de ciencias de la salud. Universidad de Guadalajara. Fundación Mexicana Para la Esclerosis Múltiple A.C.

**Denise Krch., Ph.D**

Center for Traumatic Brain Injury Research, Kessler Foundation, East Hanover, NJ, USA;

Department of Physical Medicine and Rehabilitation, Rutgers – New Jersey Medical School, Newark, NJ, USA

**Anthony H. Lequerica, Ph.D**

Center for Traumatic Brain Injury Research, Kessler Foundation, East Hanover, NJ, USA;

Department of Physical Medicine and Rehabilitation, Rutgers – New Jersey Medical School, Newark, NJ, USA

**Juan Carlos Arango-Lasprilla., Ph.D**

BioCruces Bizkaia Health Research Institute, 48903 Barakaldo, Spain.

IKERBASQUE, Basque Foundation for Science, Bilbao, Spain.

Department of Cell Biology and Histology, University of the Basque Country, Leioa, Spain.

**Corresponding author:**

Juan Carlos Arango Lasprilla, Ph.D

Biocruces Bizkaia Health Research Institute.

IKERBASQUE. Basque Foundation for Science.

Plaza de Cruces s/n. 48903

Barakaldo. Bizkaia, Spain.

Phone: (34) 946006000 (Ext. 7963)

E mail: jcalasprilla@gmail.com

Declarations of interest: none

**Abstract**

**Background.** Between 50-60% of Multiple Sclerosis (MS) patients have cognitive alterations. There are several batteries to assess cognitive impairments in MS, however, few exist for Latin Americans. The objective of this study is to evaluate the neuropsychological profile of Mexican people with MS (PwMS) and assess the utility of Norma Latina, a new battery for cognitive assessment in Latin America, in differentiating cognitive test performance between PwMS and healthy controls (HC). **Methods.** 100 PwMS and 100 HC from Mexico were evaluated with the Norma Latina battery. The following analyses were conducted: 1) low-percentiles of each participant were calculated, 2) Area Under the Curve was used to determine whether the battery discriminated between PwMS and HC, 3) four composite scores were calculated, and student’s t-test was used to compare groups according to these domains. **Results.** PwMS obtained a greater number of impaired scores compared to HCs, principally in executive function. The battery successfully discriminated between PwMS and HCs, with the strongest capacity to discriminate in the executive functions, and the weakest in memory. **Conclusion.** Establishing validation of a neuropsychological battery for Mexican PwMS will help to more accurately detect cognitive alterations, which will guide the decisions of professionals in terms of cognitive rehabilitation.

**Keywords**: Multiple Sclerosis; Assessment; cognitive problems; Latin America; Low scores.

* 1. **Introduction**

Multiple sclerosis (MS) is an autoimmune and degenerative disease of the central nervous system (CNS) caused by a progressive demyelination of the axons. MS is more common in women and is typically diagnosed between the ages of 20 and 40 (1). It is estimated that around 2 million individuals suffer from MS world-wide, whereas in Mexico approximately 20,000 inhabitants suffer from MS (2).

Cognitive alterations are one of the main consequences of MS, present in 50-60% of patients (3). The numerous investigations done with this population have allowed to establish a neuropsychological profile characterized by the presence mainly of alterations in processing speed, attention, memory, and executive functions (4). This characterization has been made possible by the existence of numerous neuropsychological batteries developed for the evaluation of these patients.

One of the most widely used batteries is the Brief International Cognitive Assessment for MS (BICAMS), which was created through an international initiative aimed at designing a brief and universal assessment instrument for people with MS (PwMS). This instrument mainly assesses processing speed and memory, and has been validated in many countries (5–11). Within Spanish-speaking populations, validation has only been carried out in Argentina (12) and Colombia (13).

The Minimal Assessment of Cognitive Function in MS (MACFIMS) is also widely used for the evaluation of these patients. It is composed of seven tests that assess speed of processing, learning and memory, executive function, visual-spatial processing, and word retrieval (14), and has been validated in Italy (15), France (16), the Czech Republic (17), the United States (18), and for the Persian language (19).

Regarding instruments for the Spanish-speaking population, Duque et al. (20) designed and validated the brief neuropsychological battery (BNB) for individuals in Spain, made up of 5 neuropsychological tests. This tool is based on the Brief Repeatable Battery of Neuropsychological Tests (BRB-N), widely used in U.S., and validated in countries such as Germany (21), the Netherlands (22), and Brazil (23).

Since most of these batteries are composed of classic neuropsychological tests such as the Symbol Digit Modalities Test (SDMT), the Paced Auditory Serial Attention Test (PASAT), California Verbal Learning Test-II (CVLT-II) or Brief Visuospatial Memory Test- Revised (BVMR-R), among others, some authors have chosen to use their own battery of tests in patients with MS. For example, Higginson, Arnett, & Voss (24) evaluated a group of 31 patients with the Rivermead Behavioral Memory Test (RBMT), Test of Everyday Attention (TEA), CVLT-II, 7/24 Spatial Recall Test, SDMT and PASAT with the aim of determining the validity of the RBMT and TEA to detect cognitive problems in patients with MS. Along these same lines, Matías-Guiu and collaborators (25) recently validated the NEURONORMA Project battery in individuals with MS in Spain. Other authors have carried out studies with the objective of validating the use of the Montreal Cognitive Assessment (MoCa) as a screening instrument in patients with MS (26).

In general, the vast majority of the batteries and neuropsychological tests described have been developed in Anglo-Saxon countries and validated in a non-Spanish-speaking population, with the exception of the BNB and the NEURONORMA projects in Spain. In fact, in Latin America (LA), the BICAMS validation study has only been carried out in Argentina (12) and Colombia (13).

The scarcity of validated and standardized assessment tools for the LA population is one of the greatest limitations that exist for the neuropsychologists in these countries to be able to properly diagnose and make recommendations for their patients (27). For this reason, large-scale multi-site, multi-country study, “Norma Latina” was recently carried out to generate healthy normative data for a battery of neuropsychological tests in adult and children from various LA countries, including Mexico (28-31), Spain (31), and Portugal (32, 33). The recent creation of these norms now affords the possibility of evaluating the cognitive status of Spanish-speaking individuals with MS, and to monitor the progression of the disease over time and the effects of treatment. The objectives of this study were to evaluate the neuropsychological profile of Mexican individuals with MS as well as to assess the utility of the Norma Latina battery in differentiating cognitive test performance between PwMS and healthy groups.

* 1. **Material and methods**

***1.2.1 Participants***

The sample consisted in 200 Mexican participants divided in two groups. The PwMS group was composed of 100 adults, mainly woman (69.0%) with an average age of 30.9 (SD=11.1) and mean of 13.7 (SD=4.2) years of education. The Healthy Control (HC) group was composed of 100 adults matched for age, sex, and education (see Table 1), and selected from the Norma Latina database from Mexico (28).

**Insert Table 1**

The inclusion/exclusion criteria for PwMS were: a) diagnosis of MS according to McDonald 2010 criteria, b) have signed the informed consent, c) ≥18 years old, d) no history of neurological or psychiatric diagnosis (e.g., schizophrenia, schizophreniform, schizoaffective, or delusional disorders), e) no history of alcohol/drugs abuse, f) no severe visual and/or hearing deficit, f) not have had a relapse occurring within the last 30 days, and g) no treatment with corticosteroids within the last 30 days. HCs met the following criteria: a) age between 18-90 years, b) were born and currently live in Mexico, c) spoke Spanish as their native language, d) had completed at least one year of formal education, and e) were able to read and write, f) no history of neurological or psychiatric conditions, g) no history of alcohol/drug abuse, h) no history of chronic disease (e.g., diabetes mellitus), i) no regular use of any medication that may impact cognitive performance, j) no severe visual and/or hearing deficit, or k) no history of learning disabilities or developmental disorders.

***1.2.2 Measures***

The neuropsychological battery was comprised of the following tests most commonly used by clinicians and researchers from LA countries (27): Rey–Osterrieth Complex Figure (ROCF), Stroop Color and Word Test, Modified Wisconsin Card Sorting Test (M-WCST), Trail Making Test (TMT), Verbal Fluency Test (VFT; in this study, animals, fruits, professions, and letters F/A/S/M), Boston Naming Test (BNT), SDMT, Brief Test of Attention (BTA) and Hopkins Verbal Learning Test–Revised (HVLT-R).

***1.2.3 Procedure***

The University Health Sciences Center of the University of Guadalajara (Guadalajara, Mexico) Research and Ethics Committee approved this study and was conducted in compliance with the declaration of Helsinki. The PwMS participants were recruited from the Mexican Foundation for Multiple Sclerosis. Those participants who met the inclusion/exclusion criteria completed a neuropsychological evaluation of approximately 2 hours’ duration. Data collection started in April 2018 and finished in January 2020. The HC data was extracted from the Norma Latina database for Mexico (28). All the participants in both groups were volunteers and did not receive financial compensation for participation. For further information regarding study procedure, see Guàrdia-Olmos et al. (28).

***1.2.4 Data analysis***

The Kolmogorov-Smirnov and Levene tests were applied to evaluate normal distribution and homoscedasticity in quantitative variables (demographic and test-scores) in both groups. Since the vast majority of scores did not meet assumptions of normality and homoscedasticity, the Mann-Whitney U test was used to verify that the samples did not differ in age and number of years of formal education. Sex distribution was evaluated by the chi-square test.

Raw scores were converted to both Z-scores and percentiles using the normative data for HVLT-R (34), M-WCST (35), SDMT (36), TMT (37), BNT (38), VFT (39, 40), BTA (41), ROCF (42), and Stroop test (43) for Mexican population. Percentiles were utilized for comparing groups on individual test scores and on the number of low-scores at various percentile cutoffs, whereas Z-scores were utilized to create cognitive domain composites for group comparison.

For each participant, the number of test scores at or below each of the percentiles was calculated. The following thresholds of low performance were analyzed in this study: (a) below the 25th percentile, (b) below the 16th percentile, (c) below the 10th percentile, (d) below the 5th percentile, and (e) below the 2nd percentile. These cut-off points were chosen to represent various levels of low performance. The Mann-Whitney U test was used to evaluate the difference between PwMS and HC on each of the 24 tests and the number of low-scores.

Receiver Operating Characteristics (ROC) or ROC curve analysis was conducted in order to determine whether the neuropsychological battery discriminated between PwMS and HC using low percentiles. The curve was created by plotting the true positive rate (also known as sensitivity) against the false positive rate (also known as specificity) for various cutoff points. The Area Under the Curve (AUC) was calculated as a measure of accuracy or precision, wherein high accuracy was deemed at a value equal to .9 or higher, moderate accuracy at a value from .7 to .9, and low accuracy at a value from .5 to .7. Additionally, Youden Index and Index of Union were calculated to determinate the optimal cutoff point for the number of low-scores below the 10th percentile and below the 5th percentile to discriminate PwMS or HC.

Finally, composite scores were created from the 24 Z-scores for the following domains: learning and memory [5 scores (ROCF Immediate Recall, ROCF Delayed Recall, HVLT-R Total Recall, HVLT-R Delayed Recall, and HVLT-R Recognition) ], language [8 scores (Letters F, A, S, M, Animals, Fruits, Professions, and BNT Total)], attention and processing speed [6 scores (TMT A & B, Stroop Color, Stroop Word, SDMT and BTA)], and executive functions [5 scores (M-WCST Categories, M-WCST Perseverations, and M-WCST Total Errors; Stroop Total Word-Color, and Stroop Interference)] following the classification suggested by Rivera et al. (44, 45) and Olabarrieta-Landa et al. (46). Each composite score was created using Stouffer's Z method (47). Student’s t-test was used to compare groups according to these four domains. For all comparisons, effect size (*r*) was calculated to determine significant group differences. Effect sizes were interpreted as small when *r* ≥.1, medium when *r* ≥.3, and large when *r* ≥.5. The data was analyzed using SPSS version 25 for Windows.

* 1. **Results**

Mann-Whitney U tests showed statistically significantly poorer performance by PwMS relative to HCs on 22 of the 24 neuropsychological test scores (see Table 2). The largest difference was observed on the M-WCST (perseveration: *p*<.001; *r*=.62; total errors: *p*<.001; *r*=.61; categories: *p*<.001; *r*=.56) and the smallest difference on Stroop Interference (*p*=.02; *r*=.16). Almost half (45.8%) of effect sizes were medium. There were no significant differences between the PwMS and HC groups on Stroop Word (*p*=.07, *r*=.12) and the ROCF copy (*p*=.37, *r*=.06).

**Insert Table 2**

Normative data for the Mexican population were used to convert raw scores into demographically-adjusted percentiles for each participant and determine base rates of low-scores for PwMS and HC groups. Mann-Whitney U tests showed significant differences between low-score distributions between MS and HC groups (*p’s*<.001; see Table 3), where the PwMS group presented with a greater number of low-scores at each percentile cutoff (25th, 16th, 10th, 5th, and 2nd). For example, at the <10th percentile cutoff (a threshold commonly used to reflect borderline impaired performance), the PwMS sample obtained a median of seven low-scores, while the HC group obtained a median of two low-scores. All comparisons showed large effect sizes (*r*>.59). The AUCs indicated a moderate degree of accuracy in discriminating between individuals with MS and HCs at every percentile cutoff (25th [AUC=.863; CI=.812, .915], 16th [AUC=.872; CI=.822, .922], 10th [AUC=.867; CI=.816, .918], 5th [AUC=.840; CI=.784, .896], and 2nd [AUC=.841; IC=.786, .896]).

**Insert Table 3**

To determine the optimal cut of point to discriminate PwMS or HC based on the number of low-scores, the Youden index and index of union (IU) were used. The cutoff points with maximum Youden index and minimal IU showed that the optimal cutoff point was ≥4 (Sensitivity=.810 and Specificity=.813) for the number of low-scores in the battery below the 10th percentile. Regarding the number of low-scores falling below the 5th percentile, the maximum Youden index showed ≥3 and ≥4 low scores as potential cutoffs; however, IU, which defines the optimal cut-off as the value that minimizes the difference between sensitivity and specificity, as well as minimizing each of their differences from the area under the ROC curve, showed that the optimal cutoff point was ≥3 (Sensitivity= .730 and Specificity = .844). See Table 4 for more details.

**Insert table 4**

The results presented in Table 5 showed significant differences between the groups in all cognitive domains (learning and memory [*t*=7.399; *p*<.001; *r*=.47]); (language [*t*=8.366; *p*<.001; *r*=.51]); (attention and speed processing [*t*=8.275; *p*<.001; *r*=.51]); (executive functions [*t*=12.048; *p*<.001; *r*=.65]) where PwMS presented worse mean Z-scores in each domain compared to HC participants (see Figure 1). The AUCs for each specific cognitive domain suggest high accuracy (AUC’s >.773, 95% CI=.706 - .839) in discriminating between PwMS and HC on composite Z-scores (see Figure 2).

**Insert Table 5 and Figures 1 & 2**

* 1. **Discussion**

The study was designed to evaluate the neuropsychological profile of Mexican PwMS and assess the utility of the Norma Latina battery to differentiate cognitive test performance between PwMS and HCs. Considering the individual tests of the battery, PwMS performed worse than HCs on 22 of the 24 tests. Group comparisons showed that PwMS obtained a greater number of impaired scores at each impairment threshold relative to HCs, principally in executive function, attention and processing speed, language, and learning and memory.

A closer look at individual test performance revealed that the largest effects were observed on all variables of the M-WCST and TMT-B. Medium effects were seen on SDMT, TMT-A, ROCF Recall, HVLT-R Total recall and Delayed recall, BNT, and several fluency tasks (animals and the letters M, A, and F). Small effect sizes were observed on BTA, all Stroop variables, the remainder of the fluency tasks (fruits, occupations, and the letter S), and HVLT-R Recognition. No statistically significant group differences were observed on Stroop Word and ROCF copy. These results may be used to inform clinicians in their prioritization of tests to select in a diagnostic setting, particularly when assessment time and access to test resources are limited.

In a co-normed battery, base rates of the prevalence of low-scores obtained by HCs at any given cutoff can be used to ascertain whether the number of low-scores obtained by PwMS is broadly normal or below expectation. In the current study, the HC group obtained a median of one low-score below the most impaired cutoff level (< 2nd percentile), versus PwMS who obtained a median of three low-scores. At the most liberal cutoff (<25th percentile), HCs obtained a median of 5 low-scores below the cutoff point, while their MS patient counterparts obtained a median of 12. Overall, PwMS obtained 2.4 to 4 times as many low-scores at any given threshold relative to HCs. These findings revealed that the co-normed battery is a significant classifier of PwMS versus HCs and the consideration of these base rates may be useful in making clinical inferences and reducing the likelihood of misdiagnosing cognitive impairments.

The battery evaluated in this study covers the main cognitive domains affected within MS, with results revealing that performance in the executive function, attention/processing speed, language, and learning and memory domains discriminated groups with medium to large effect sizes. The mean PwMS group performance fell in the Impaired range for executive function and attention/processing speed and the Borderline Impaired range for language, whereas HC performance fell in the Average range in these domains. Although the PwMS group performed within the Impaired range in the learning and memory domain, HC performance was below expectation (Low Average range), likely explaining why this effect size fell just below the threshold for a large effect size. Nonetheless, these results indicate that all domains in the Norma Latina battery were sensitive to the effects of MS.

In consideration of individual test scores, base rates at various cut points and performance by domain, the neuropsychological battery demonstrated a strong ability to discriminate between PwMS and HCs. For the battery as a whole, having ≥4 scores falling below the 10th percentile or ≥3 scores falling below the 5th percentile had optimal sensitivity and specificity for discriminating between PwMS and HCs. These findings robustly support the use of this battery for clinical assessment of PwMS in Mexico and suggest its potential applicability to PwMS in other countries. This study makes an important contribution in expanding the ability to diagnose cognitive impairments associated with MS in LA, adding to the one existing study conducted in Argentina (12) and the existing tools available for Spain populations (e.g., BNB and the NEURONORMA).

**1.4.1 Strengths, Limitations, and Future Directions**

The PwMS and HC groups were well matched to allow for comparison, and the fact that the tests were co-normed in a sample of HCs also from Mexico facilitates the use of low-score base rates among HCs to assist in interpretation of the cognitive profiles among PwMS (48). Another strength of the study was the fact that the sample contained a mixture of MS subtypes, extending its representativeness. The ROCF was included in the current study by virtue of its being part of the Norma Latina battery. However, it is appreciated that there is a significant motor component required for intact performance on this test. In the current sample, performance on the copy trial for the MS group was not significantly different than that of HCs, indicating that impairment on the recall trial for the MS group was not likely due to motor limitations. Nonetheless, motor impairments are not uncommon in MS and should impaired performance be observed on the copy trial in practice settings, clinical judgment should guide the use or interpretation of the recall trial.

Secondly, the HC group was recruited from a prior study [Norma Latina (28)], the associated exclusion criteria were rigorous to ensure the HC sample was truly “healthy.” For the MS sample, we wanted to be inclusive of the typical population commonly seen in clinics within this region. Albeit the MS sample did not exclude for other illnesses or conditions that may impact cognition, we felt it was important that this sample was representative of the broad population of individuals with MS in Mexico that may be encountered in clinical practice. We took steps to match the groups on other sociodemographic characteristics without limiting the generalizability within the MS group.

Future studies may wish to examine individuals with relapsing-remitting or primary progressive subtypes separately if more specific characterization is desired within each subtype. In addition, future studies may consider adding other objective measures of disease burden or functional status to better characterize the PwMS sample. Nevertheless, the finding that this battery is sensitive to the cognitive deficits associated with MS opens the possibility for future studies examining other aspects of cognition such as changes over time or responsiveness to interventions to establish efficacy for the promotion of evidence-based cognitive rehabilitation practices in Mexico.

**1.4.2 Clinical Implications**

This battery successfully discriminated between PwMS and a well-matched HC group. This study provides compelling support for the utility of the battery in assessment of PwMS. Beyond identifying cognitive sequelae of the neural degeneration, assessment of cognitive strengths and weaknesses can provide valuable information about domains to target for cognitive rehabilitation in order maximize functional independence and community participation for individuals living with MS. Additionally, the brevity of this battery and the accessibility of the measures comprising the battery make it an exceedingly valuable diagnostic tool within clinical settings in Mexico and LA. Follow-up research is needed to evaluate the differentiability of this battery in other patient populations and in other countries.

**Funding**

D.R.U. was supported by a predoctoral fellowship from the Basque Government (PRE\_2019\_1\_0164).

**References**

1. Porcel J, Olivares T. Rehabilitación neuropsicológica. Intervención y práctica clínica. En: Esclerosis múltiple. España: Elsevier; 2011. p. 189-206.
2. Cuevas C, Velázquez M, Núñez L, Skromne E, Árcega R, Barroso N, et al. Consenso mexicano para la esclerosis múltiple. Guía diagnóstica y terapéutica. Revista Mexicana de Neurociencia. 2007;8(2):155-62.
3. Meca-Lallana V, Gascón-Giménez F, Ginestal-López R. C, Higueras Y, Téllez-Lara N, Carreres-Polo J, et al. Cognitive impairment in multiple sclerosis: diagnosis and monitoring. Neurol Sci, 2021;1-11.
4. Arango-Lasprilla JC, DeLuca J, Chiaravalloti N. [Neuropsychological profile of multiple sclerosis]. Psicothema. febrero de 2007;19(1):1-6.
5. Betscher E, Guenter W, Langdon DW, Bonek R. Polish validation of the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS battery): correlation of cognitive impairment with mood disorders and fatigue. Neurol Neurochir Pol. 2021;55(1):59-66.
6. Sandi D, Rudisch T, Füvesi J, Fricska-Nagy Z, Huszka H, Biernacki T, et al. The Hungarian validation of the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) battery and the correlation of cognitive impairment with fatigue and quality of life. Multiple Sclerosis and Related Disorders. noviembre de 2015;4(6):499-504.
7. Costers L, Gielen J, Eelen PL, Schependom JV, Laton J, Remoortel AV, et al. Does including the full CVLT-II and BVMT-R improve BICAMS? Evidence from a Belgian (Dutch) validation study. Multiple Sclerosis and Related Disorders. noviembre de 2017;18:33-40.
8. Polychroniadou E, Bakirtzis C, Langdon D, Lagoudaki R, Kesidou E, Theotokis P, et al. Validation of the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) in Greek population with multiple sclerosis. Multiple Sclerosis and Related Disorders. septiembre de 2016;9:68-72.
9. Giedraitienė N, Kizlaitienė R, Kaubrys G. The BICAMS Battery for Assessment of Lithuanian-Speaking Multiple Sclerosis Patients: Relationship with Age, Education, Disease Disability, and Duration. Med Sci Monit. 10 de diciembre de 2015;21:3853-9.
10. Farghaly M, Langdon DW, Shalaby NM, Shehata HS, Abokrysha NT, Hassan A, et al. Reliability and validity of Arabic version of the brief international cognitive assessment for multiple sclerosis: Egyptian dialect. Egypt J Neurol Psychiatry Neurosurg. diciembre de 2021;57(1):51.
11. Hämäläinen P, Leo V, Therman S, Ruutiainen J. Validation of the Finnish version of the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) and evaluation of the applicability of the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ) and the Fatigue Scale for Motor and Cognitive Functions (FSMC). Brain Behav. junio de 2021;11(6):e02087.
12. Vanotti S, Benedict R, Caceres F. Validation Of The Brief International Cognitive Assessment For Multiple Sclerosis (BICAMS) in Argentina (P5. 201). Neurology. 2015;84.
13. Alarcón AN, Ayala OD, García JR, Montañés P. Validation of the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) in a Colombian Population. Mult Scler Relat Disord. julio de 2020;42:102072.
14. Benedict RHB, Fischer JS, Archibald CJ, Arnett PA, Beatty WW, Bobholz J, et al. Minimal neuropsychological assessment of MS patients: a consensus approach. Clin Neuropsychol. agosto de 2002;16(3):381-97.
15. Migliore S, Ghazaryan A, Simonelli I, Pasqualetti P, Landi D, Palmieri MG, et al. Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS) in the Italian population. Neurol Sci. agosto de 2016;37(8):1261-70.
16. Maubeuge N, Deloire MSA, Brochet B, Erhlé N, Charré-Morin J, Saubusse A, et al. French validation of the Brief International Cognitive Assessment for Multiple Sclerosis. Rev Neurol (Paris). febrero de 2021;177(1-2):73-9.
17. Dusankova JB, Kalincik T, Havrdova E, Benedict RHB. Cross cultural validation of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS) and the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). Clin Neuropsychol. 2012;26(7):1186-200.
18. Gromisch ES, Zemon V, Holtzer R, Chiaravalloti ND, DeLuca J, Beier M, et al. Assessing the criterion validity of four highly abbreviated measures from the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS). Clin Neuropsychol. octubre de 2016;30(7):1032-49.
19. Eshaghi A, Riyahi-Alam S, Roostaei T, Haeri G, Aghsaei A, Aidi MR, et al. Validity and reliability of a Persian translation of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS). Clin Neuropsychol. 2012;26(6):975-84.
20. Duque P, Ibanez J, Del Barco A, Sepulcre J, de Ramon E, Fernandez-Fernandez O, et al. [Normalisation and validation of the Brief Neuropsychological Battery as the reference neuropsychological test in multiple sclerosis]. Rev Neurol. 1 de marzo de 2012;54(5):263-70.
21. Scherer P, Baum K, Bauer H, Göhler H, Miltenburger C. [Normalization of the Brief Repeatable Battery of Neuropsychological tests (BRB-N) for German-speaking regions. Application in relapsing-remitting and secondary progressive multiple sclerosis patients]. Nervenarzt. octubre de 2004;75(10):984-90.
22. Boringa JB, Lazeron RH, Reuling IE, Adèr HJ, Pfennings L, Lindeboom J, et al. The brief repeatable battery of neuropsychological tests: normative values allow application in multiple sclerosis clinical practice. Mult Scler. agosto de 2001;7(4):263-7.
23. Damasceno A, Amaral JMSDS, Barreira AA, Becker J, Callegaro D, Campanholo KR, et al. Normative values of the Brief Repeatable Battery of Neuropsychological Tests in a Brazilian population sample: discrete and regression-based norms. Arq Neuropsiquiatr. marzo de 2018;76(3):163-9.
24. Higginson CI, Arnett PA, Voss WD. The ecological validity of clinical tests of memory and attention in multiple sclerosis. Arch Clin Neuropsychol. abril de 2000;15(3):185-204.
25. Matías-Guiu JA, Sánchez-Benavides G, Rivera-Àvila N, Cortés-Martínez A, Delgado-Alonso C, Delgado-Álvarez A, et al. Validation of the Neuronorma battery for neuropsychological assessment in multiple sclerosis. Mult Scler Relat Disord. julio de 2020;42:102070.
26. Gómez-Moreno SM, Cuadrado ML, Cruz-Orduña I, Martínez-Acebes EM, Gordo-Mañas R, Fernández-Pérez C, et al. Validation of the Spanish-language version of the Montreal Cognitive Assessment as a screening test for cognitive impairment in multiple sclerosis. Neurologia (Engl Ed). 23 de enero de 2020;S0213-4853(19)30149-5.
27. Arango-Lasprilla JC, Stevens L, Morlett Paredes A, Ardila A, Rivera D. Profession of neuropsychology in Latin America. Appl Neuropsychol Adult. agosto de 2017;24(4):318-30.
28. Guàrdia-Olmos J, Peró-Cebollero M, Rivera D, Arango-Lasprilla JC. Methodology for the development of normative data for ten Spanish-language neuropsychological tests in eleven Latin American countries. NeuroRehabilitation. 2015;37(4):493-9.
29. Rodríguez-Lorenzana A, Núñez-Fernández S, Adana-Díaz L, Mascialino G, Ponce TY, Rivera D, et al. Normative data for test of learning and memory in an ecuadorian adult population. The Clinical Neuropsychologist. 1 de diciembre de 2020;34(sup1):54-69.
30. Rodríguez-Lorenzana A, Benito-Sánchez I, Adana-Díaz L, Paz CP, Yacelga Ponce T, Rivera D, et al. Normative Data for Test of Verbal Fluency and Naming on Ecuadorian Adult Population. Front Psychol. 27 de mayo de 2020;11:830.
31. Rivera D, Arango-Lasprilla JC. Methodology for the development of normative data for Spanish-speaking pediatric populations. NeuroRehabilitation. 2017;41:581-92.
32. Vicente SG, Rivera D, Barbosa F, Gaspar N, Dores AR, Mascialino G, et al. Normative data for tests of attention and executive functions in a sample of European Portuguese adult population. Aging, Neuropsychology, and Cognition. 4 de mayo de 2021;28(3):418-37.
33. Vicente SG, Ramos-Usuga D, Barbosa F, Gaspar N, Dores AR, Rivera D, et al. Regression-Based Norms for the Hopkins Verbal Learning Test-Revised and the Rey–Osterrieth Complex Figure in a Portuguese Adult Population. Archives of Clinical Neuropsychology. 21 de mayo de 2021;36(4):587-96.
34. Arango-Lasprilla JC, Rivera D, Garza MT, Saracho CP, Rodríguez W, Rodríguez-Agudelo Y, et al. Hopkins Verbal Learning Test- Revised: Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):699-718.
35. Arango-Lasprilla JC, Rivera D, Longoni M, Saracho CP, Garza MT, Aliaga A, et al. Modified Wisconsin Card Sorting Test (M-WCST): Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):563-90.
36. Arango-Lasprilla JC, Rivera D, Rodríguez G, Garza MT, Galarza-Del-Angel J, Rodríguez W, et al. Symbol Digit Modalities Test: Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):625-38.
37. Arango-Lasprilla JC, Rivera D, Aguayo A, Rodríguez W, Garza MT, Saracho CP, et al. Trail Making Test: Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):639-61.
38. Olabarrieta-Landa L, Rivera D, Morlett-Paredes A, Jaimes-Bautista A, Garza MT, Galarza-del-Angel J, et al. Standard form of the Boston Naming Test: Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):501-13.
39. Olabarrieta-Landa L, Rivera D, Galarza-Del-Angel J, Garza MT, Saracho CP, Rodríguez W, et al. Verbal fluency tests: Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):515-61.
40. Rivera D, Olabarrieta-Landa L, Van der Elst W, Gonzalez I, Rodríguez-Agudelo Y, Aguayo Arelis A, et al. Normative data for verbal fluency in healthy Latin American adults: Letter M, and fruits and occupations categories. Neuropsychology. marzo de 2019;33(3):287-300.
41. Rivera D, Perrin PB, Aliaga A, Garza MT, Saracho CP, Rodrŕguez W, et al. Brief Test of Attention: Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):663-76.
42. Rivera D, Perrin PB, Morlett-Paredes A, Galarza-Del-Angel J, Martínez C, Garza MT, et al. Rey-Osterrieth Complex Figure - copy and immediate recall: Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):677-98.
43. Rivera D, Perrin PB, Stevens LF, Garza MT, Weil C, Saracho CP, et al. Stroop Color-Word Interference Test: Normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation. 2015;37(4):591-624.
44. Rivera D, Mascialino G, Brooks BL, Olabarrieta-Landa L, Longoni M, Galarza-Del-Angel J, et al. Multivariate Base Rates of Low Scores on Tests of Executive Functions in a Multi-Country Latin American Sample. Developmental Neuropsychology. 2 de enero de 2021;46(1):1-15.
45. Rivera D, Olabarrieta-Landa L, Brooks BL, Ertl MM, Benito-Sánchez I, Quijano MC, et al. Multivariate Base Rates of Low Scores on Tests of Learning and Memory Among Latino Adult Populations. J Int Neuropsychol Soc. septiembre de 2019;25(08):834-44.
46. Olabarrieta-Landa L, Ramos Usuga D, Rivera D, Leal G, Bailey KC, Calderón Chagualá A, et al. Prevalence of low scores on language tests as a potential factor in misdiagnosis of cognitive impairment in a Spanish-speaking adult population. Applied Neuropsychology: Adult. 2 de enero de 2022;29(1):41-52.
47. Whitlock MC. Combining probability from independent tests: the weighted Z-method is superior to Fisher’s approach: Combining probabilities from many tests. Journal of Evolutionary Biology. 25 de agosto de 2005;18(5):1368-73.
48. Brooks BL, Iverson GL. Comparing Actual to Estimated Base Rates of «Abnormal» Scores on Neuropsychological Test Batteries: Implications for Interpretation. Archives of Clinical Neuropsychology. 1 de febrero de 2010;25(1):14-21.
49. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). Neurology. 1 de noviembre de 1983;33(11):1444-1444.

**Table 1.** *Comparison of the samples by age, education and gender.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Group** | **Statistic** | **Sig.** | **Effect size (*r*)** |
|  | **PwMS (n= 100)** | **HC (n= 100)** |
|  | **Median** | **Min** | **Max** | **Median** | **Min** | **Max** |
| Age | 38.0 | 19.0 | 69.0 | 38.0 | 19.0 | 69.0 | U=4837.0 | .690 | 0.028 |
| Education | 15.0 | 1.0 | 22.0 | 12.0 | 1.0 | 20.0 | U=4270.0 | .072 | 0.127 |
| Disease duration | 5.0 | 0.0 | 28.0 | -- | -- | -- | -- | -- | -- |
| EDSS | 3.7 | 0.0 | 6.5 | -- | -- | -- | -- | -- | -- |
| Sex | Female | 69 | 69.0% | Female | 69 | 69.0% | $$X^{2}= 0.0$$ | 1.00 | 0.00 |
| Male | 31 | 31.0% | Male | 31 | 31.0% |
| Subtypes | Relapsing–remitting | 93 | 93.0% | -- | -- | -- | -- | -- | -- |
| Progressive | 5 | 5.0% | -- | -- | -- | -- | -- | -- |
| Primary progressive | 2 | 2.0% | -- | -- | -- | -- | -- | -- |

*Note:* HC = Healthy Control; PwMS = People with Multiple Sclerosis; EDSS (49) = Expanded Disability Status Scale

**Table 2.** *Comparison between groups on neuropsychological test scores.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Test-Score** | **Group** | **Median** | **Min** | **Max** | **Mann-Whitney U** | **Sig.** | **Effect size (*r*)** |
| ROCF Copy | HC | 34.0 | 15.5 | 36.0 | 4313.000 | .375 | .06 |
| MS | 33.5 | 4.5 | 36.0 |
| ROCF Recall | HC | 21.0 | 5.0 | 34.0 | 3003.500 | <.001 | .30†† |
| MS | 16.0 | 0.0 | 32.0 |
| Stroop Word | HC | 97.0 | 10.0 | 140.0 | 3539.000 | .080 | .12† |
| MS | 87.0 | 38.0 | 160.0 |
| Stroop Color | HC | 68.0 | 38.0 | 113.0 | 2972.000 | .001 | .24† |
| MS | 61.0 | 13.0 | 101.0 |
| Stroop Word-Color | HC | 41.0 | 21.0 | 98.0 | 2734.000 | <.001 | .28† |
| MS | 37.0 | 3.0 | 67.0 |
| Stroop Interference | HC | 3.5 | -16.0 | 46.9 | 3340.000 | .021 | .16† |
| MS | 1.5 | -36.3 | 18.0 |
| M-WCST Categories | HC | 6.0 | 1.0 | 6.0 | 1652.000 | <.001 | .56††† |
| MS | 3.0 | 1.0 | 6.0 |
| M-WCST Perseveration errors | HC | 1.0 | 0.0 | 42.0 | 1215.000 | <.001 | .62††† |
| MS | 11.0 | 0.0 | 39.0 |
| M-WCST Total errors | HC | 4.0 | 0.0 | 42.0 | 1241.000 | <.001 | .61††† |
| MS | 24.0 | 2.0 | 42.0 |
| TMT-A | HC | 40.0 | 19.0 | 100.0 | 2491.500 | <.001 | .40†† |
| MS | 61.5 | 14.0 | 100.0 |
| TMT-B | HC | 64.0 | 28.0 | 212.0 | 1939.500 | <.001 | .50††† |
| MS | 118.5 | 26.0 | 300.0 |
| BTA | HC | 16.0 | 7.0 | 20.0 | 3559.000 | <.001 | .25† |
| MS | 15.0 | 3.0 | 20.0 |
| VFT Letter F | HC | 13.0 | 3.0 | 25.0 | 3291.000 | <.001 | .30†† |
| MS | 10.0 | 2.0 | 21.0 |
| VFT Letter A | HC | 13.0 | 5.0 | 27.0 | 3055.500 | <.001 | .34†† |
| MS | 10.0 | 2.0 | 23.0 |
| VFT Letter S | HC | 13.0 | 4.0 | 23.0 | 3345.000 | <.001 | .29† |
| MS | 10.0 | 1.0 | 22.0 |
| VFT Letter M | HC | 14.0 | 4.0 | 30.0 | 3072.000 | <.001 | .33†† |
| MS | 11.0 | 3.0 | 19.0 |
| VFT Animals | HC | 20.0 | 7.0 | 33.0 | 3268.000 | <.001 | .30†† |
| MS | 17.0 | 6.0 | 28.0 |
| VFT Fruits | HC | 15.0 | 6.0 | 28.0 | 3964.500 | .011 | .18† |
| MS | 13.0 | 0.0 | 22.0 |
| VFT Occupations | HC | 13.0 | 5.0 | 26.0 | 3750.500 | .002 | .22† |
| MS | 12.0 | 0.0 | 22.0 |
| BNT | HC | 52.0 | 33.0 | 67.0 | 3193.000 | <.001 | .31†† |
| MS | 49.0 | 0.0 | 60.0 |
| SDMT | HC | 48.0 | 18.0 | 87.0 | 2216.500 | <.001 | .47†† |
| MS | 34.5 | 5.0 | 80.0 |
| HVLT-R Total learning | HC | 23.0 | 12.0 | 36.0 | 2965.500 | <.001 | .35†† |
| MS | 20.0 | 6.0 | 32.0 |
| HVLT-R Delayed recall | HC | 9.0 | 3.0 | 12.0 | 2961.500 | <.001 | .36†† |
| MS | 7.0 | 0.0 | 12.0 |
| HVLT-R Recognition | HC | 12.0 | 0.0 | 12.0 | 3692.500 | <.001 | .25† |
| MS | 11.0 | 4.0 | 12.0 |

*Note:* BTA = Brief Test of Attention; HC = Healthy Control; HVLT-R = Hopkins Verbal Learning Test – Revised; MS = Multiple Sclerosis; Min = Minimum; Max = Maximum; M-WCST = Modified Wisconsin Cart Sorting Test; ROCF = Rey-Osterrieth Complex Figure; SDMT = Symbol Digit Modalities Test; TMT = Trail Making Test; VFT = Verbal Fluency Test. † = Small effect, †† = Medium effect; ††† = Large effect

**Table 3.** *Comparison between groups on the number of test scores falling below specified percentile cutoffs and associated ROC Curve characteristics.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cutoff | Group | Median | Min. | Max. | Mann- Whitney U | Sig. | Effect size | ROC |
| AUC | Lower bound | Upper bound |
| <25th percentile | HC | 5 | 0 | 14 | 1310.500 | <.001 | .62††† | .863 | .812 | .915 |
| MS | 12 | 1 | 23 |
| <16th percentile | HC | 3 | 0 | 12 | 1231.500 | <.001 | .64††† | .872 | .822 | .922 |
| MS | 9 | 1 | 22 |
| <10th percentile | HC | 2 | 0 | 12 | 1277.500 | <.001 | .63††† | .867 | .816 | .918 |
| MS | 7 | 1 | 21 |
| <5th percentile | HC | 1 | 0 | 7 | 1538.500 | <.001 | .59††† | .840 | .784 | .896 |
| MS | 4 | 0 | 20 |
| <2nd percentile | HC | 1 | 0 | 5 | 1528.000 | <.001 | .60††† | .841 | .786 | .896 |
| MS | 3 | 0 | 19 |

*Note:* HC = Healthy Control; MS = Multiple Sclerosis; Min = Minimum; Max = Maximum; ††† = Large effect; Lower bound and upper bound refer to the 95% confidence intervals of the AUC.

**Table 4.** *Cut-points and associated sensitivity and specificity values.*

|  |  |  |
| --- | --- | --- |
| **Cut-point** | **<5th percentile** | **<10th percentile** |
| **Se** | **Sp** | ***J*** | **IU** | **Se** | **Sp** | ***J*** | **IU** |
| ≥2 | .830 | .635 | .465 | .214 | .910 | .479 | .389 | .431 |
| ≥3 | .730 | .844 | .574 | .114 | .850 | .635 | .485 | .248 |
| ≥4 | .650 | .927 | .577 | .277 | .810 | .813 | .623 | .111 |
| ≥5 | .470 | .938 | .408 | .468 | .720 | .896 | .616 | .176 |
| ≥6 | .420 | .948 | .368 | .528 | .590 | .948 | .538 | .358 |
| ≥7 | .350 | .958 | .308 | .608 | .510 | .958 | .468 | .448 |

Se = Sensitivity; Sp = Specificity; *J* =Youden index; IU = Index of Union

**Table 5.** *Average performance between groups by cognitive domain and associated ROC Curve characteristic.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Domain** | **Group** | **Mean Z-score** | **SD** | ***t*** | ***df*** | ***Sig.*** | **Effect size** | **ROC** |
| **AUC** | **Lower bound** | **Upper bound** |
| Executive Function | HC | 0.17 | 1.29 | 12.048 | 198 | < .001 | .65††† | .891 | .847 | .934 |
| MS | -2.47 | 1.77 |
| Attention & Processing Speed | HC | 0.29 | 1.43 | 8.275 | 198 | < .001 | .51††† | .812 | .751 | .872 |
| MS | -2.30 | 2.79 |
| Language | HC | 0.27 | 1.52 | 8.366 | 198 | < .001 | .51††† | .795 | .734 | .856 |
| MS | -1.96 | 2.19 |
| Learning and Memory | HC | -0.79 | 1.40 | 7.399 | 198 | < .001 | .47†† | .773 | .706 | .839 |
| MS | -2.81 | 2.34 |

*Note:* df = degrees of freedom; HC = Healthy Control; MS = Multiple Sclerosis; SD = Standard Deviation; †† = Medium effect; ††† = Large effect

**Figure 1. *Performance between groups by cognitive domain.***

**Figure 2. *ROC curves representing different cognitive domains.*****