OXFORD UNIVERSITY PRESS

Archives of Clinical Neuropsychology 00 (2022) 1–15

Archives of CLINICAL NEUROPSYCHOLOGY

Multiple Sclerosis and Depression: Translation and Adaptation of the Spanish Version of the Chicago Multiscale Depression Inventory and the Study of Factors Associated with Depressive Symptoms

Carla Sanchis-Segura[1], Álvaro Javier Cruz-Gómez[2], Sónia Félix Esbrí[1], Alba Sebastián Tirado[1], Peter A. Arnett[3], Cristina Forn[1,*]

[1]*Departament de Psicología Bàsica, Clínica i Psicobiología, Universitat Jaume I, Castelló de la Plana, Spain*
[2]*Grupo de Neuroimagen y Psicofisiología, Instituto de Investigación e Innovación Biomédica de Cádiz (INiBICA), Castelló, Spain*
[3]*Department of Psychology, The Pennsylvania State University, University Park, PA, USA*

*Corresponding author at: Departament de Psicología Bàsica, Clínica i Psicobilogia, Universitat Jaume I, Universitat de Ciències de la Salut. Avda. Sos Baynat s/n (Edifici d'investigació II), 12071 Castelló de la Plana, Spain.
E-mail address: forn@uji.es (C. Forn).

Accepted 2 November 2022

Abstract

Objective: Depressive disorder occurs in up to 50% of persons with Multiple Sclerosis (PwMS). Accurate assessment of depression in MS is essential in clinical settings because depressive symptomatology can affect the clinical course of the disease.
Methods: We translated, adapted, and tested the Spanish version of the Chicago Multiscale Depression Inventory (CMDI), a specific test to assess depression in neurological disorders. We compare our results with those obtained with previous versions of the questionnaire (English and Italian). Finally, we also analyze the relationship between the results obtained on the CMDI and demographic, clinical, and cognitive variables.
Results: The results obtained with the Spanish version of the CMDI were similar to those observed in previous published versions. We also observed higher depression scores in PwMS (especially in progressive forms) compared with healthy controls. Moreover, depression symptomatology was related to higher disability and fatigue and worse cognitive performance in PwMS.
Conclusions: The results support the validity of the CDMI in the Spanish population, as well as the association between depression and other characteristic symptoms of MS. These findings also emphasize the importance of good assessment and multidisciplinary treatment of depression in PwMS.

Keywords: Multiple sclerosis; Depression; Assessment; Fatigue; Cognition

Introduction

Multiple sclerosis (MS) is a chronic disease characterized by neurological symptoms that include motor and sensory deficits, fatigue, pain, cognitive impairment, and clinical depression. The lifetime prevalence of risk of depression in this clinical population has been estimated at around 50% (Arnett, Barwick, & Beeney, 2008), and major depression is more frequently observed in persons with MS (PwMS) than in other disabling neurological disorders (Solaro, Gamberini, & Fabio, 2018). Moreover, depression directly affects quality of life and the clinical course of MS (Arnett et al., 2008; Hoffmeister et al., 2021) because it is associated with worse cognitive performance (Arnett et al., 2008, 2021; Siegert & Abernethy, 2005), decreased adherence to pharmacological treatment, increased fatigue, and physical deficits (Brown et al., 2009; Bruce, Hancock, Arnett, & Lynch, 2010; Siegert & Abernethy, 2005).

https://doi.org/10.1093/arclin/acac096

Regarding the latter, there is a significant overlap between somatic symptoms of depression and the specific clinical symptomatology of MS, principally fatigue, pain, or poor sleep (Hind et al., 2016). This overlap leads to the specific diagnosis of depression in MS and may result in overestimating or underestimating depression in this clinical population (Nyenhuis et al., 1995; Siegert & Abernethy, 2005). Therefore, to ensure the proper assessment of depressive symptomatology in MS patients, psychometric instruments are needed that can differentiate between the somatic symptoms of depression and MS.

In this regard, the Chicago Multiscale Depression Inventory (CMDI) is a self-report measure that was specifically developed to assess depression in neurological patients (Nyenhuis et al., 1995; Nyenhuis et al., 1998). This questionnaire contains three adjective-based subscales that can be used and interpreted separately to examine unique features of depression, and it has demonstrated adequate validity and reliability to measure depressive symptomatology in PwMS (Chang et al., 2003; Hind et al., 2016; Solari et al., 2003). Unfortunately, a Spanish version of this inventory does not yet exist.

Taking these precedents into account, the overarching goal of this study was to adapt the CMDI to Spanish and compare its results in healthy control (HC) and PwMS with results from studies that developed the two currently available versions of this test (English: Nyenhuis et al., 1995; Italian: Solari et al., 2003). Moreover, to enhance the current understanding of the association between different kinds of depressive symptomatology and features of MS, we also explored the relationships between the CMDI scores and a large number of demographic, clinical, and cognitive variables in PwMS.

Method

Participants

One hundred eighty-nine PwMS and 76 HCs were tested. One hundred fifty-seven PwMS had a clinically defined MS diagnosis, according to the McDonald criteria, with different clinical courses (relapsing–remitting -RR-, secondary progressive -SP- and primary progressive -PP-), and 32 were diagnosed with clinically isolated syndrome (CIS). All the PwMS were free of exacerbation or steroid treatment in the 3 months prior to enrolment and had no history of alcohol or drug abuse or central nervous system diseases other than MS. Thirty-nine percent of PwMS reported using antidepressants, and 80% reported using MS disease-modifying pharmacotherapy. HC were recruited from acquaintances and relatives of PwMS. The only exclusion criteria applied to HC were the self-report of: any current or previous neurological or psychiatric diagnosis; a history of alcohol or drug abuse (excluding tobacco). After providing informed consent, all participants answered the CMDI questionnaire and were assessed neuropsychologically.

Instruments

CMDI questionnaire. The CMDI is a 42-item, self-report questionnaire that includes three subscales (14 items in each subscale) that assess: mood (dysphoria), vegetative symptoms (physical malfunctioning), and evaluative symptoms (negative self-concept and self-criticism). Examinees are asked to rate the extent to which each word or phrase describes them during the past week, including today, on a Likert scale from 1 to 5, where 1 is "not at all" and 5 is "extremely."

Procedure

To adapt the CMDI questionnaire to Spanish, two professional English to Spanish translators with previous experience in clinical psychology terminology carried out two independent translations of the CMDI items. In a second step, the two translators compared and discussed their respective proposals, resolving any discrepancies until reaching a first consensual CMDI version. This document was presented in a joint meeting of the two translators and the authors of the present study, who back-translated the items to English and worked together with the translators until reaching a pre-definitive version of the inventory. This pre-definitive version of the CMDI was administered to a sample of 10 healthy volunteers and 12 PwMS with differing educational levels, who assessed the items' clarity and comprehensibility. None of the items were identified as problematic, and so no further modifications were made. The resulting Spanish-adapted CMDI was used in subsequent stages of this study, and it is available upon request from the authors.

Neuropsychological assessment. Data were collected by two expert neuropsychologists in the facilities of the Associació d'Esclerosi Múltiple de Castelló (AEMC) and of the Hospital General de Castelló (Spain).

All the participants were assessed first with the Spanish-adapted CMDI and then with the Spanish version of the Brief Repeatable Battery-Neuropsychology Test (BRB-N; Sepulcre et al., 2006), which included the following tests: SRT: Selective

Table 1. Demographic, clinical, and neuropsychological data of all participants. Data are reported as means and standard deviations (between brackets) except in the case of, Expanded Disability Status Scale which is reported as median and inter-quantile range (Q25, Q75)

	HC ($n = 76$; 48.7% F)	**All PwMS** ($n = 189$; 68.3% F)	**CIS** ($n = 32$; 59.4% F)	**RR** ($n = 113$; 64.6% F)	**SP** ($n = 31$; 83.9% F)	**PP** ($n = 13$; 84.6 F)
Age	33.96 (8.56) $t = 5.17$, $p < .001$, $d = 0.66$	40.5 (10.9)	34.62 (8.91)	38.57 (9.89)	49.26 (9.23)	50.62 (9.47)
Disease years	—	8.24 (7.56)	2.28 (2.89)	7.06 (5.65)	17.48 (8.97)	11.08 (6.29)
Education years	13.57 (2.71) $t = -6.29$, $p < .001$, $d = -0.83$	11.19 (3.14)	12.03 (3.15)	11.18 (3.08)	10.65 (2.98)	9.77 (3.59)
Expanded Disability Status Scale	—	2 (2, 3.5)	1.09 (0, 2)	2.41 (2, 3)	4.76 (3.75, 6)	3.88 (3,5)
Fatigue Severity Scale	3.21 (1.26) $t = 4.09$, $p < .001$, $d = 0.52$	3.99 (1.74)	2.97 (1.48)	4.11 (1.79)	4.26 (1.57)	4.86 (1.29)
Matrix subtest	19.58 (3.32) $t = -6.61$, $p < .001$, $d = -0.84$	16.3 (4.40)	17.38 (4.37)	16.48 (4.16)	15.26 (5.04)	14.38 (4.41)
Selective Reminding Test Long-Term Storage	47.3 (12.12) $t = -3.70$, $p < .001$, $d = -0.49$	41.1 (13.0)	49.59 (9.16)	40.73 (12.57)	34.9 (13.78)	37.69 (13.79)
Selective Reminding Test Long-Term Retrieval	38.91 (10.61) $t = -4.71$, $p < .001$, $d = -0.57$	31.4 (14.0)	42.25 (11.44)	31 (12.88)	24.13 (14.85)	26 (11.43)
Selective Reminding Test Delayed Recall	9.21 (2.01) $t = -4.37$, $p < .001$, $d = -0.57$	7.97 (2.26)	9.09 (1.65)	7.92 (2.14)	7.35 (2.54)	7.15 (2.97)
Spatial Recall test Long-Term Storage	22.72 (4.01) $t = -8.73$, $p < .001$, $d = -1.05$	19.4 (5.51)	19.94 (4.77)	19.98 (5.13)	17.32 (5.46)	18.31 (5.57)
Spatial Recall Test Delayed-Recall	8.21 (1.77) $t = -12.50$, $p < .001$, $d = -1.53$	6.79 (2.22)	7.09 (2.01)	7.1 (2.16)	5.52 (2.35)	6.46 (1.98)
Symbol Digit Modalities Test	59.96 (10.72) $t = -8.64$, $p < .001$, $d = -1.05$	46.1 (14.0)	51.69 (11.57)	48.7 (12.6)	34.16 (14.92)	38.92 (11.57)
Paced Auditory Serial Addition test 3 sec	48.30 (9.56) $t = -10.89$, $p < .001$, $d = -1.15$	33.86 (18.6)	44.06 (11.65)	34.47 (18.7)	24.58 (19.21)	25.62 (17.01)
Phonetic fluency	13.95 (4.2) $t = -4.22$, $p < .001$, $d = -0.52$	11.3 (5.27)	13.62 (6.63)	11.3 (5.15)	10.32 (4.07)	8.62 (2.93)
Semantic fluency	22.37 (4.64) $t = -3.37$, $p < .001$, $d = -0.43$	20.1 (5.55)	22.59 (4.91)	20.38 (5.44)	17.9 (5.34)	16.92 (5.77)

Reminding Test (SRT); SPART: Spatial Recall Test (SPART); SDMT: Symbol Digit Modalities Test (SDMT); PASAT: Paced Auditory Serial Addition Test (PASAT); and phonetic and semantic fluency. We also used the Matrix Reasoning subtest (WAIS-III) to assess intelligence quotient and the Fatigue Severity Scale (FSS; Krupp, Larocca, Muir Nash, & Steinberg, 1989) to assess fatigue (see also Table 1, Table S1).

Data analysis

All the statistical analyses were conducted with freely available packages for R (R Core Team, 2020). These analyses can be grouped in four major domains.

Psychometric characteristics of the Spanish-adapted CMDI. A confirmatory factor analysis (CFA) was conducted to determine to which extent the Spanish-adapted CMDI reproduces the same factorial structure as that described for the original CMDI (Chang et al., 2003; Nyenhuis et al., 1998). Of note, although the CMDI is composed of 3 subscales and 42 scoring items, the

original CMDI has a five-factor structure that is calculated from only 24 items (see Nyenhuis et al., 1998 and Chang et al., 2003 for details). Thus, a CFA aiming to validate this structure was conducted using the *cfa* function from the *lavaan* package using the robust MLM estimator. This analysis was based on *a priori* specification that 5 factors characterized the data and each item was constrained to load on one and only one factor, then allowing for empirically validating the factor structure described for the original CMDI. Model fit was first evaluated with the same two indexes employed by Chang et al. (2003) when validating the original CMDI: the comparative fit index (CFI) and the non-normed fit index (NNFI). For these indexes, values higher than 0.9 were considered acceptable, and values equal to or higher than 0.95 as desirable (Carlson & Mulaik, 1993; Hooper, Coughlan, & Mullen, 2008). Moreover, following the recommendations of the so-called Two-Index Presentation Strategy (Hu & Bentler, 1999), the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR) were also calculated. According to the same recommendations, these indices were judged as adequate when their values were below 0.09 and 0. 06, respectively.

Measurement invariance of the CMDI for PwMS and HC was evaluated following the four-step method proposed by Widaman and Reise (1997), hence successively testing configural, metric (also known as weak factorial), scalar (also known as strong factorial), and residual (also known as strict or invariant uniqueness) measurement invariance. For this assessment, the *compareFit* function of the *semTools* package and the same fit indices (CFI, NNFI, RMSEA, and SRMR) and cut-off values previously described were employed. The observed changes (delta) in these indices at each successive model comparison are also reported. To our knowledge, no previous study has established the metric, scalar, or residual invariance of neither the English or the Italian versions of the CMDI. Therefore, the obtained results could not be compared with preceding findings and were solely judged in reference to previously published cutoffs.

The overall reliability of the Spanish-adapted CMDI (the 42 items) was evaluated by calculating the overall omega (ω_u; Flora, 2020) reliability coefficient and its 95% confidence interval (CI) with the *ci.reliability* function of the *MBESS* package. Moreover, the omega-higher-order (ω_{ho}) and omega estimates for each factor of the previous CFA (24 items) were calculated with the *reliabilityL2* and *reliability* functions of the *semTools* package (Flora, 2020). However, because omega indexes have not been reported for either the English or the Italian versions of the CMDI, we also calculated the reliability measures provided by Nyenhuis et al. (1995) and Solari et al. (2003) for these two previous CMDI versions. Thus, Cronbach's alpha coefficient (and corresponding CI) for the 14 items of each subscale and for the total scores on this inventory (42 items). As in these two studies, item-subscale correlations were also calculated.

Comparison with of the Spanish-adapted CMDI scores with those of other CMDI versions. The results obtained with the Spanish-adapted CMDI were compared with those obtained with two other currently available versions. More specifically, we corroborated that the means on each subscale of the Spanish-adapted CMDI fell within the 95% prediction intervals (PIs; Spence & Stanley, 2016) for the means observed in the studies with the English and Italian versions.

The proportions of "depressed" HC and PwMS observed in the present study were also compared with their reported equivalents in the studies by Nyenhuis et al. (1995) and Solari et al. (2003). Because there is no validated method to calculate the 95% PI for proportions, these comparisons were conducted using χ^2-based tests. Note that, although we used the same criterion (1.5 standard deviations above the control group mean) as in previous studies to separate "depressed"/"non-depressed" individuals, we use the terms "probably-depressed" (P-depressed) and "non-depressed" (N-depressed) to make it clear that these labels refer to individuals who had high scores on one or more CMDI subscale but had not been diagnosed with any DSM-V depression-related disorder at the time of the study.

Descriptive statistics and between-group comparisons. Although their assumptions could not be always satisfied, between-group comparisons involving HC versus MS, MS subtypes, or P- versus N-depressed PwMS were conducted using the same parametric approach as in the Nyenhuis et al. (1995) and Solari et al. (2003) reports (i.e., mean comparisons through Student's *t* tests for independent samples or ANOVAs). In all cases, *p*-values were FDR-adjusted for multiple testing (Benjamini & Hochberg, 1995), and effect sizes were estimated using Cohen's *d* and its 95% CI. Nevertheless, the same between-group comparisons were also conducted with equivalent, robust, and non-parametric tests of the *WRS2* package (Mair & Wilcox, 2020). These analyses yielded very similar results and identical conclusions as the parametric comparisons and, therefore, they are solely reported as supplementary material.

In addition, proportion-based between-group comparisons were performed using χ^2 tests, and, when appropriate, the magnitude of these effects was estimated in terms of Cohen's H or risk ratios (and their 95% CIs).

Relationship between CMDI scores and clinical, demographic, or cognitive variables. The Boruta algorithm (Kursa, Jankowski, & Rudnicki, 2010; Kursa & Rudnicki, 2010) was used to select and rank the relative importance of all the variables with predictive value in each CMDI subscale or total scores.

The Boruta algorithm has the same foundations as the better-known Random Forest algorithm (Kursa et al., 2010; Kursa & Rudnicki, 2010). However, conversely to the Random Forest, the Boruta algorithm does not only rank features' importance, but it does actually allow distinguishing which of them should be considered as relevant, tentatively relevant, or irrelevant as predictors of the outcome of interest. To do so, the Boruta algorithm: compares, at each iteration, the predictive importance of each potential predictor to that of the so-called "shadow features" (features artificially created by shuffling the values of each original variable across instances or subjects), so designating a feature as relevant when its importance is significantly higher than that of the best-performing shadow feature (Kursa & Rudnicki, 2010); uses the outcomes of all iterations (in the present study, 2000) to calculate the binomial distribution of "successes" (the number of times that a particular variable is identified as relevant) and "failures" (the number of times that it is not) for each variable, so determining whether the observed frequency of successes falls into the rejection, irresolution, or acceptance area according to a predefined p-value (in the present study, the default value of $p < .01$ was employed).

Independent Boruta-based analyses were performed using as the outcome variable the scores of each CMDI subscale or the CMDI total scores and clinical-demographic or cognitive variables as predictors. In a second step, we estimated the Spearman's correlations between the predictors that were not rejected by the Boruta algorithm and the CMDI scores. When appropriate, coefficients of overlapping correlations were compared by means of Hittner's test (Hittner, May, & Silver, 2003).

Results

Demographic, Clinical, and Neuropsychological Data

Table 1 shows demographic characteristics of the study sample, including the different phenotypes of the PwMS. Compared with HC, PwMS were older and had fewer years of education, higher scores on the FSS, and lower scores on all the neuropsychological tests. Regarding the differences between the MS phenotypes, CIS patients were very similar to HC on all variables, whereas progressive MS forms (SP and PP) presented worse clinical and cognitive scores than the RR phenotype.

The Spanish-adapted CMDI: Psychometric characteristics

A CFA revealed that the Spanish-adapted CMDI has a similar factorial structure to the original (English) version of the CMDI described in Nyenhuis et al. (1998) and in Chang et al. (2003). Thus, the observed CFI (0.92) and NNFI (0.91) values were acceptable according to proposed cut-offs (Carlson & Mulaik, 1993) and very similar to those reported for the original CMDI (0.93 and 0.92, respectively; see Chang et al., 2003). Further support for the fitted model was provided by two other indices. Thus, the obtained SRMR (0.06) and RMSEA (0.046, [0.039, 0.054]) values were below their recommended thresholds (0.09 and 0. 06, respectively; Hu & Bentler, 1999). Moreover, all items loaded high (range: 0.59–0.92) and exclusively in their respective factors (see Supplementary material online, Table S2) except the item "*Easily awakened*" that, as in the original CMDI, had relatively low factor loading (0.31 in the present study, 0.26 and 0.47 in the original CMDI, see Nyenhuis et al., 1998 and Chang et al., 2003). Taken together, these observations confirm that the model provides an adequate approximation to the data, which is very similar to that of the original CMDI, and also that items behaved as expected.

The Spanish-adapted CMDI also seems to exhibit adequate configural, metric, scalar, and residual invariance across HC and PwMS. More specifically, as shown in Table 2, all CFI and NNFI values were above the 0.9 cut-off, whereas all RMSEA values were smaller than 0.05. Moreover, all the SRMR values were below the 0.09 cut-off value except in the residual invariance model. Taken together, these results confirm that the Spanish-adapted CMDI provides a similar measurement of the latent constructs in these two groups of participants.

Finally, the Spanish-adapted CMDI showed adequate reliability. More specifically, the ω_u□value was 0.95 and its 95% CI was narrow [0.93, 0.97]. Similarly, the omega-higher-order was high (0.86) as also were those of each factor of the previously conducted CFA (range: 0.75–0.92). However, because the two previously existing versions of the CMDI (English and Italian) did not report omega indices but Cronbach's alpha coefficients for each CMDI subscale and total score, we also calculate these reliability indices. The obtained Cronbach's alpha coefficients were high, and their 95% CIs were narrow (Mood = 0.92 [0.90, 0.94]; Evaluative: 0.87 [0.82, 0.90]; Vegetative: 0.84 [0.81, 0.87]; Total score = 0.92 [0.90, 0.93]), reflecting an adequate level of homogeneity of the item pool within the subscales and for the total CMDI score. Moreover, these Cronbach's alpha values were similar to those observed for the two currently available versions of the CMDI (English and Italian), which ranged between 0.91–0.98 (Nyenhuis et al., 1995) and 0.86–0.95 (Solari et al., 2003), respectively. In a similar vein, item-subscale correlations were adequate (Mood: median = 0.74, range = 0.48–0.83; Evaluative: median = 0.47, range = 0.33–0.63; Vegetative: median = 0.46, range = 0.32–0.55) and similar to those reported for the Italian-adapted CMDI (Solari et al., 2003).

Table 2. Metric invariance between HC and PwMS

	M1: Configural invariance	M2: Metric Invariance	M3: Scalar invariance	M4: Residual invariance
CFI	0.965	0.938	0.936	0.930
NNFI	0.959	0.931	0.932	0.929
RMSEA [90% CI RMSEA]	0.026 [0.000, 0.041]	0.034 [0.017, 0.046]	0.034 [0.017, 0.046]	0.035 [0.019, 0.047]
SRMR	0.07	0.083	0.084	0.093
Compared to model	—	M1	M2	M3
Delta CFI	—	−0.027	−0.002	−0.006
Delta NNFI	—	−0.028	0.001	−0.003
Delta RMSEA	—	0.008	0	0.001
Deltat SRMR	—	0.013	0.001	0.009

The Spanish-adapted CMDI: Scores and comparison with the English- and Italian-adapted CMDI

The scores (mean ± 95% CI) of the HC (light blue) and MS (orange) groups on each subscale of the CMDI are depicted in Fig. 1A. A series of *t*-tests revealed that, as expected, PwMS had higher depression scores than HC on all the CMDI subscales (Mood subscale: 26.17 vs. 21.93, $t_{263} = 3.44$, $p < .001$, $d = 0.5$; Evaluative subscale: 21.17 vs. 18.14, $t_{263} = 2.90$, $p < .005$, $d = 0.43$; Vegetative subscale: 33.58 vs. 28.64, $t_{263} = 3.77$, $p < .001$, $d = 0.53$). Consequently, PwMS also exhibited higher total CMDI scores than HC (81.34 vs. 68.72, $t_{263} = 3.80$, $p < .001$, $d = 0.56$; Fig. 1B).

Figure 1C depicts the HC and MS means reported by Nyenhuis et al. (1995) (triangles) and Solari et al. (2003) (squares) for each CMDI subscale and the total depression scores. In this Figure, the 95% PIs (vertical lines) of these means are also depicted, illustrating the range within which these results can be considered successfully replicated (Spence & Stanley, 2016). As the figure shows, all the HC means in the current study (circles) fell within the 95% PI of the corresponding means in the Nyenhuis et al. (1995) and Solari et al. (2003) studies, thus confirming that scores obtained with the original and Spanish-adapted CMDI are fully comparable with those obtained with the English and Italian versions of the CMDI. Although CMDI scores of PwMS are expected to exhibit more variation due to between-sample differences in clinical variables, the MS averages obtained with the Spanish-adapted CMDI were also very similar to those found with other CMDI versions. Thus, all the MS average scores in the current study were included within the 95% PI calculated for the MS averages reported in the study by Nyenhuis et al. (1995). Similarly, the MS average scores in the current study were also within the 95% PI calculated for the MS averages reported by Solari et al. (2003) on the Mood and Evaluative subscales of the Italian version of the CMDI.

Because the Spanish-adapted CMDI average scores were very similar to those found in other CMDI versions, the criterion used in the original CMDI (1.5 standard deviations above the control group mean) could be also reasonably adopted when aiming to employ this new psychometric instrument to classify individuals as "probably-depressed" (P-depressed) and "non-depressed" (N-depressed). Thus, in the present study, individuals were categorized as P-depressed according to the Mood, Evaluative, or Vegetative subscale when their scores were equal to or higher than 33, 27, or 41 in these subscales, respectively. When attending to the total scores, the cut-off to be classified as P-depressed was set at 96 points. All these values are very similar to those derived from the original CMDI (35, 26, 38, and 93, respectively).

P-depressed rates in PwMS doubled those observed in HC (Mood subscale: 24.9% vs. 11.9%, $\chi^2 = 5.51$, $p = .019$, Risk Ratio = 2.09 [1.08, 4.06], $p = .025$; Evaluative subscale: 22.2% vs. 10.5%, $\chi^2 = 4.84$, $p = .028$, Risk Ratio = 2.11 [1.04, 4.28], $p = .028$; Vegetative subscale: 22.2% vs. 9.21%, $\chi^2 = 6.09$, $p = .014$, Risk Ratio = 2.41 [1.13, 5.13], $p = .025$; Total CMDI scores = 26.5% vs. 10.53%, $\chi^2 = 8.04$, $p = .005$, Risk Ratio = 2.51 [1.25, 5.04], $p = .018$). Moreover, as Table 3 shows, the rates of P-depressed participants in the present study were not statistically different from those reported by Nyenhuis et al. (1995) and Solari et al. (2003), thus confirming the comparability between the outcomes obtained with the Spanish-adapted CMDI and those obtained with the other two currently available versions of this inventory.

Relationship Between CMDI Scores and Clinical/Demographic Variables of PwMS

A series of ANOVAs revealed that the MS subtypes differed in their average scores on all the CMDI subscales (Mood: $F_{3,185} = 3.18$, $p = .025$; Evaluative: $F_{3,185} = 3.46$, $p = .017$; Vegetative: $F_{3,185} = 7.11$, $p < .001$) and, consequently, on the total CMDI scores ($F_{3,185} = 5.12$, $p = .002$). Similarly, differences between MS subtypes in the proportion of P-depressed patients were statistically significant on the Mood ($\chi^2_{(3)} = 8.07$, $p = .044$) and Evaluative: ($\chi^2_{(3)} = 8.23$, $p = .041$) subscales, as well as on the Total CMDI score ($\chi^2_{(3)} = 8.74$, $p = .033$), and they approached statistical significance on the Vegetative subscale

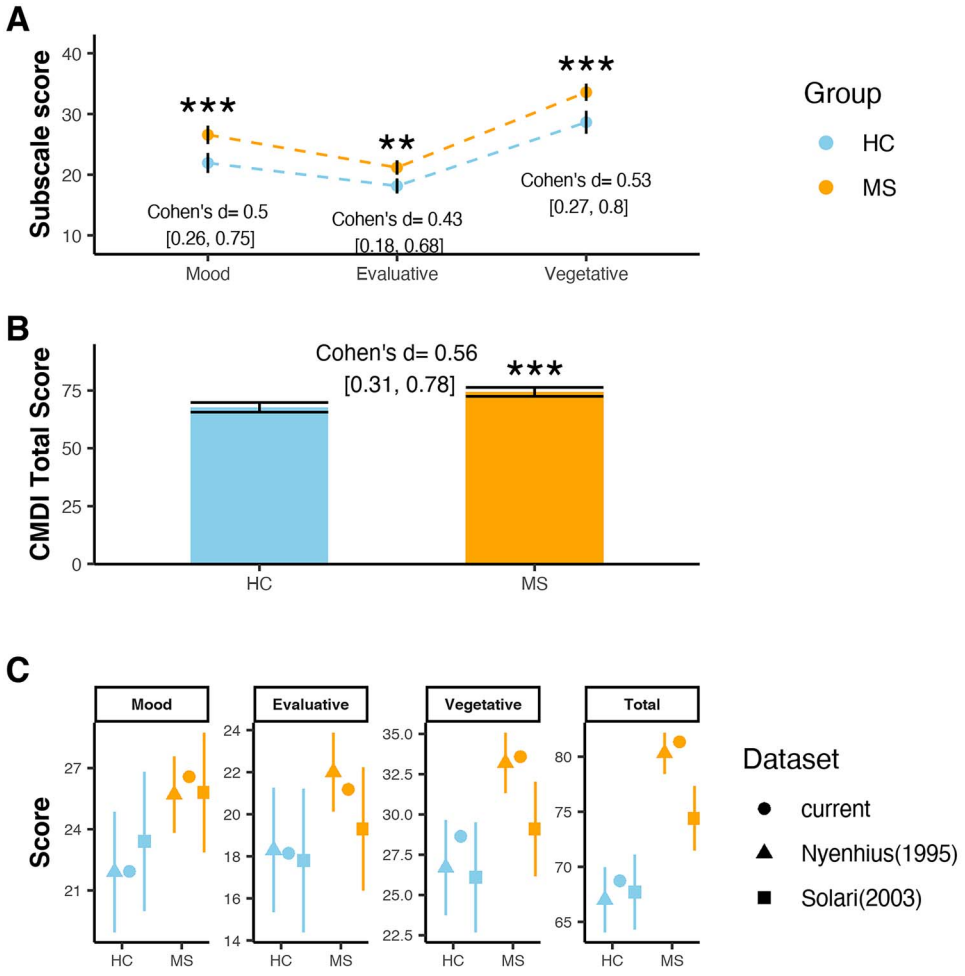

Fig. 1. Participants (HC and MS) scores in the Spanish-adapted CMDI. Panel A displays mean + 95% confidence interval (CI) of HC (light blue) and MS (orange) scores in each subscale of this test. Panel B illustrates CMDI total scores of the same groups of participants. For completeness, statistical significance (***$p < .001$, **$p < .005$) and effect size (Cohen's d [95% CI]) of the corresponding between group differences are also included. Finally, panel C depicts the 95% prediction intervals (PIs) (vertical lines) of the means exhibited by HC and MS patients in each CMDI subscale and in the overall depression scores of the studies of Nyenhuis et al., 1995 (triangles) and Solari et al., 2003 (squares). As it can be readily observed, all the HC means observed in the current study (circles) fell within the 95% PIs of their corresponding means of the Nyenhuis and colleagues and Solari and colleagues, studies, then confirming that scores obtained with the original and the Spanish-adapted CMDI are fully comparable.

Table 3. Proportion of "probably depressed" participants (HC and MS) according to the different CMDI scores

	Mood subscale	Evaluative subscale	Vegetative subscale	Total score
MS current study vs. MS Nyenhuis et al. (1995)	24.9% vs. 17.7%, $\chi^2_{(1)} = 1.25$, $p = .52$	22.2% vs. 22.2%, $\chi^2_{(1)} = 8.26e{-}31$, $p = 1$	22.2% vs. 34.6%, $\chi^2_{(1)} = 3.95$, $p = .187$	26.5% vs. 26.6%, $\chi^2_{(1)} = 1.14e{-}30$, $p = 1$
MS current study vs. MS Solari et al. (2003)	24.9% vs. 18.8% $\chi^2_{(1)} = 1.62$, $p = .27$	22.2% vs. 16.0% $\chi^2_{(1)} = 1.94$, $p = .27$	22.2% vs. 17.8% $\chi^2_{(1)} = 0.8$, $p = .37$	26.5% vs. 19.9% $\chi^2_{(1)} = 1.71$, $p = .27$
HC current study vs. HC Solari et al. (2003)	11.9% vs. 7.3% $\chi^2_{(1)} = 1.04$, $p = .60$	10.5% vs. 8.3% $\chi^2_{(1)} = 0.241$, $p = .62$	9.2% vs. 3.1% $\chi^2_{(1)} = 2.86$, $p = .36$	10.5% vs. 7.3% $\chi^2_{(1)} = 0.56$, $p = .61$

This table reports the rates of P-depressed individuals observed in the HC and MS of the current study. The table also includes the results of χ^2-based comparisons of these rates to those reported in the studies of Nyenhuis et al. (1995) and Solari et al. (2003). Note that Nyenhuis et al. (1995) did only report P-depressed rates for MS patients. Displayed p-values are FDR corrected for the multiple comparisons contained in each row.

($\chi^2_{(3)} = 6.91$, $p = .074$). However, because MS subtypes integrate differences in several clinical/demographic variables, analyzing CMDI outcomes according to the MS subtype makes it difficult to identify which of these variables are actually related to CMDI

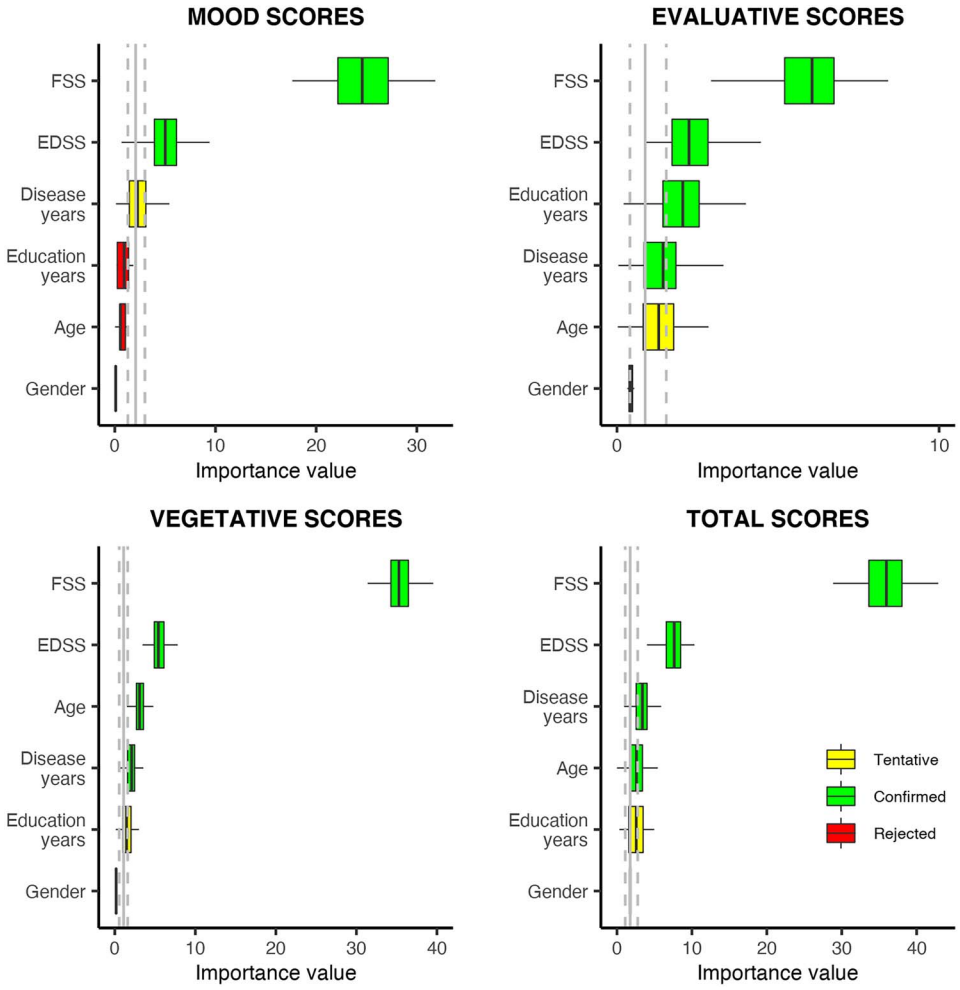

Fig. 2. Predictive importance of demographic and clinical variables in CMDI scores. Each panel depicts the importance profile of each CMDI subscale (mood, evaluative, vegetative) or CMDI total scores (2,000 runs, $p < .001$). Within each panel, green/yellow/red boxplots illustrate the distribution of the estimated importance value of demographic and clinical variables confirmed as relevant/tentative and non-relevant predictors of CMDI scores, respectively. The vertical line depicts the median, whereas the dotted lines correspond to the values of the first and third quartile of the maximum shadow attribute.

scores. Therefore, we decided to investigate the relationships between CMDI scores and clinical and demographic variables in PwMS without introducing any subtype categorization.

First, the Boruta algorithm was used to identify all the demographic and clinical variables that could have predictive values in the CMDI scores. As Fig. 2 shows, fatigue (FSS) scores were the most relevant predictor of CMDI scores, although the strength of this relationship varied in the different CMDI subscales (Vegetative > Mood> Evaluative). A substantive, albeit weaker, relationship between physical disability (EDSS) and depression scores was also confirmed for all the CMDI subscales. Other variables, such as age, disease evolution in years, or educational level, only had predictive value in the scores on specific CMDI subscales.

In a second step, Spearman's correlations between the CMDI scores and the clinical/demographic variables that were not rejected in the previous Boruta-based analysis were calculated (Fig. 3). These correlation values corroborated and refined the results obtained with the Boruta algorithm, leading to three main conclusions. First, fatigue was confirmed as the clinical-demographic variable most clearly associated with the CMDI scores, and, as revealed in a series of Hittner's tests, the strength of this relationship was greater for the Vegetative subscale (0.54) and total CMDI scores (0.49) than for the Mood (0.41) and Evaluative (0.36) subscales ($\bar{Z}^{*}_{\text{FSS-Veg vs. FSS-Total}} = 1.71$, $p = .105$; $\bar{Z}^{*}_{\text{FSS-Veg vs. FSS-Eval}} = 3.25$, $p = .003$; $\bar{Z}^{*}_{\text{FSS-Veg vs. FSS-Mood}} = 2.65$, $p = .012$; $\bar{Z}^{*}_{\text{FSS-Total vs. FSS-Eval}} = 3.53$, $p = .002$; $\bar{Z}^{*}_{\text{FSS-Total vs. FSS-Mood}} = 3.02$, $p = .005$; $\bar{Z}_{\text{FSS-Mood vs. FSS-Eval}} = 1.08$, $p = .276$). Second, the EDSS had a weaker but more uniform relationship with all the CMDI subscales (correlations range: 0.25–0.32;

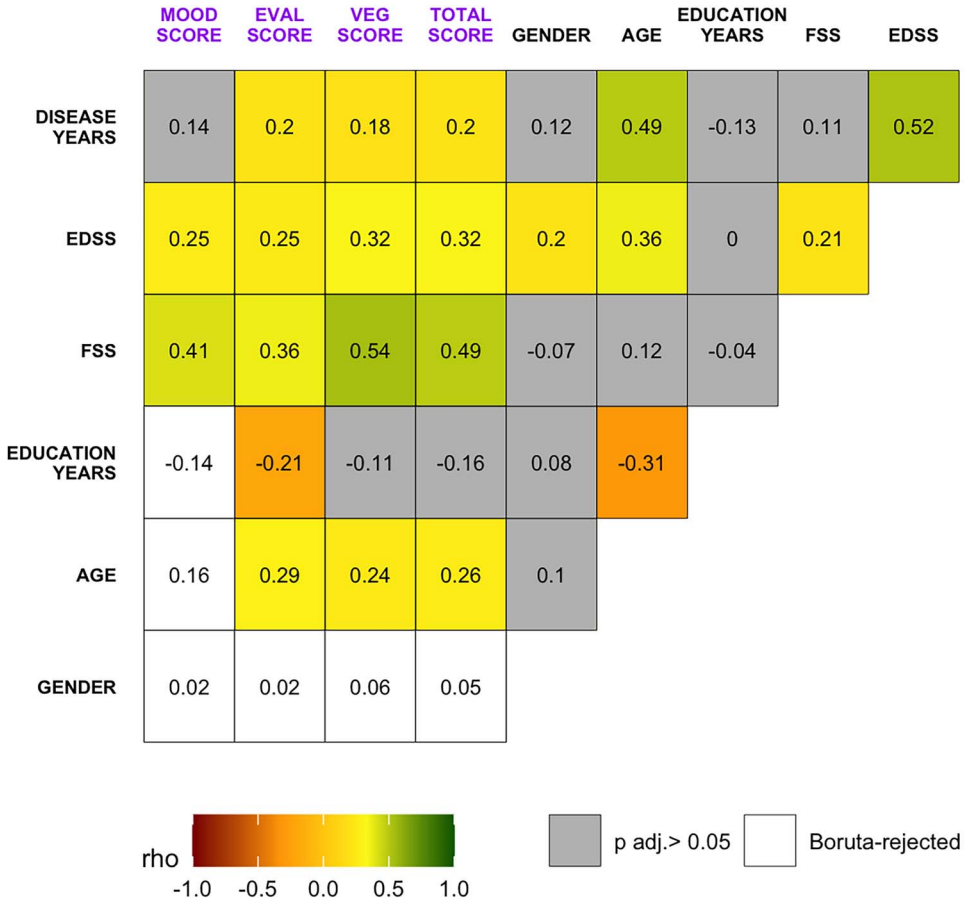

Fig. 3. Correlations between demographic/clinical variables and CMDI scores. The figure summarizes the values of the ordinal correlations (Spearman's rho) between demographic/clinical variables and the CMDI scores. The correlations that remain significant after multiple testing corrections are depicted using a gradient scale, whereas those that did not or that had been rejected as possible predictors by the Boruta algorithm (Fig. 2) are depicted in gray and white colored cells, respectively.

Hittner's tests: $p > .05$ in all cases). Third, all the other clinical/demographic variables (except gender) were selectively associated with specific CMDI subscales.

To offer a complementary perspective, the means of P-depressed and N-depressed PwMS (as defined by each CMDI subscale or total score) on all the clinical/demographic variables considered in the present study were compared (Table 4). More specifically, this table displays the descriptive statistics (means and, in brackets, standard deviations) for each group and the outcomes of the tests used to assess their differences (Student's t or χ^2 statistics, FDR-corrected p-values, Cohen's d or H and its 95% CI; see methods section for details). Similar to the results of our previous correlational assessment, the largest and most reliable differences between P-depressed and N-depressed PwMS were observed in the FSS and EDSS scores, especially if these groups were defined based on their scores on the Vegetative subscale.

Relationship Between CMDI Scores and Cognitive Variables of PwMS

As Fig. 4 shows, the Boruta algorithm identified the PASAT3 as the cognitive test with the highest predictive value in the CMDI scores. Although with lower relative importance, the SDMT and semantic fluency were also related to all the CMDI scores. Conversely, phonetic fluency, verbal memory (SRT), and visual memory (SPART) measures exhibited weaker and/or subscale-dependent associations with the CMDI scores.

Spearman's correlation values (Fig. 5) corroborated and refined these results, leading to three major conclusions: the PASAT3 was the cognitive variable most clearly related to the CMDI scores, and, as confirmed in a series of Hittner's tests, the strength of this association was greater for the Evaluative subscale (−0.45) and the total CMDI score (−0.39) than for the Mood

Table 4. Differences between P-depressed and N-depressed PwMS in clinical/demographic variables

	Mood subscale		Evaluative subscale		Vegetative Subscale		Total scores	
	P-dep	N-dep	P-dep	N-dep	P-dep	N-dep	P-dep	N-dep
Age	41.7 (11.4)	40.1 (10.7)	42.9 (10.9)	39.8 (10.8)	43.4 (12.3)	39.6 (10.3)	43.5 (11.7)	39.4 (10.4)
	$t = 0.89$, $p = .41$		$t = 1.65$, $p = .10$		$t = 1.83$, $p = .09$		**$t = 2.15$, $p = .04$**	
	$d = 0.15$ [−0.18, 0.49]		$d = 0.29$ [−0.06, 0.64]		**$d = 0.35$ [0.01, 0.7]**		**$d = 0.37$ [0.05, 0.7]**	
Education years	10.5(3.2)	11.4(3.1)	10.1 (3.0)	11.4 (3.1)	9.9 (2.9)	11.5(3.1)	10 (3.1)	11.6 (3.1)
	$t = -1.68$, $p = .16$		**$t = -2.45$, $p = .03$**		**$t = -3.06$, $p = .007$**		**$t = -3.1$, $p = .007$**	
	$d = -0.28$ [0.48, −0.18]		**$d = -0.42$ [−0.07, −0.77]**		**$d = -0.51$ [−0.17, −0.86]**		**$d = -0.51$ [−0.18, −0.84]**	
Disease years	9.1 (8.6)	7.9 (7.2)	10.5 (9.3)	7.6 (6.8)	9.4 (8.2)	7.9 (7.3)	9.8 (9.4)	7.7 (6.7)
	$t = 0.83$, $p = .41$		$t = 1.83$, $p = .08$		$t = 1.08$, $p = .28$		$t = 1.48$, $p = .14$	
	$d = 0.15$ [−0.62,0.05]		**$d = 0.37$ [0.03, 0.73]**		$d = 0.20$ [−0.14, 0.55]		$d = 0.28$ [−0.04, 0.61]	
Expanded Disability Status Scale	3.2 (1.7)	2.5 (1.5)	3.2 (1.8)	2.5 (1.5)	3.2 (1.5)	2.5 (1.6)	3.3 (1.8)	2.5 (1.5)
	$t = 2.64$, $p = .02$		**$t = 2.32$, $p = .04$**		**$t = 2.67$, $p = .02$**		**$t = 2.80$, $p = .01$**	
	$d = 0.47$ [0.13, 0.81]		**$d = 0.45$ [0.1, 0.79]**		**$d = 0.44$ [0.1, 0.79]**		**$d = 0.49$ [0.17, 0.82]**	
Fatigue Severity Scale	5.0 (1.4)	3.6 (1.7)	4.8 (1.6)	3.8 (1.7)	5.2 (1.4)	3.7 (1.7)	5.2 (1.4)	3.6 (1.7)
	$t = 5.65$, $p < .001$		**$t = 3.71$, $p = .002$**		**$t = 5.76$, $p < .001$**		**$t = 6.83$, $p < .001$**	
	$d = 0.85$ [0.51, 1.2]		**$d = 0.62$ [0.27, 0.97]**		**$d = 0.93$ [0.57, 1.29]**		**$d = 1.03$ [0.69, 1.37]**	
Gender (% females)	74.5%	66.2%	73.8%	66.7%	73.8%	66.7%	76%	65.5%
	$\chi^2_{(1)} = 0.76$, $p = .38$		$\chi^2_{(1)} = 0.48$, $p = .5$		$\chi^2_{(1)} = 0.47$, $p = .49$		$\chi^2_{(1)} = 1.42$, $p = .23$	
	$H = 0.18$ [−0.15, 0.51]		$H = 0.16$ [−0.19, 0.5]		$H = 0.16$ [−0.19, 0.5]		$H = 0.23$ [−0.09, 0.56]	

This table displays the means and standard deviation (between brackets) of the values observed in P-depressed patients and N-depressed patients (as defined by each CMDI subscale or total scores) in all the clinical/demographic variables considered in the present study. For all variables except gender, differences between groups were tested by means of Student's *t* tests and their size was estimated in terms of Cohen's d and its 95% CI. The *p*-values associated to these *t*-tests are FDR-corrected for all *t*-test based comparisons included within each column, whereas the 95% CI for Cohen's *d* that do not continue the zero value allow identifying between-means differences that were significant at uncorrected $p < .05$. Gender-associated differences were tested by comparing the proportion of females included in each group by means of a chi-squared test of independence and the size of the differences in these proportions were estimated using Cohen's *H*. All statistically significant differences are highlighted in bold.

(−0.31) and Vegetative (−0.29) subscales ($Z^*_{\text{PASAT-Eval vs. PASAT-Total}} = 1.88$, $p = .071$; $Z^*_{\text{PASAT-Eval vs. PASAT-Mood}} = 3.14$, $p = .007$; $Z^*_{\text{PASAT-Eval vs. PASAT-Veg}} = 2.90$, $p = .007$; $Z^*_{\text{PASAT-Total vs. PASAT-Mood}} = 2.82$, $p = .007$; $Z^*_{\text{PASAT-Total vs. PASAT-Veg}} = 3.03$, $p = .007$; $Z_{\text{PASAT-Mood vs. PASAT-Veg}} = 0.43$, $p = .667$). The SDMT and fluency measures exhibited statistically significant correlations with all the CMDI subscales and the total score. Once again, the values of these correlations (and especially those involving the SDMT) were higher for the Evaluative subscale. Verbal and visual memory measures were solely and weakly correlated with the Evaluative subscale scores.

These conclusions received convergent support from mean-based comparisons of P-depressed and N-depressed PwMS on the cognitive variables. Table 5 displays the means and standard deviations (in brackets) of the values observed in P-depressed and N-depressed PwMS (as defined by each CMDI subscale or total score) on all the tests included in the BRB-N, as well as the Matrix Reasoning subtest of the WAIS-III battery. More specifically, these comparisons confirmed that differences between P-depressed and N-depressed PwMS were larger on the PASAT3, the SDMT, and the semantic fluency measure, and these differences were larger when these groups were created according to their scores on the Evaluative subscale. Moreover, differences in verbal and visual memory were only observed when defining P-depressed/ N-depressed PwMS based on the Evaluative subscale.

Discussion

The first goal of our study was to adapt the CMDI to Spanish. Our findings indicate that this Spanish CMDI adaptation yields very similar scores to those of the existing English (Nyenhuis et al., 1995) and Italian (Solari et al., 2003) CMDI versions, not only the total scores but also on the three different subscales of this inventory. The same conclusion is reached when comparing the percentage of "probably depressed" (P-depressed) participants observed in our study, which did not differ significantly from the percentages reported by Nyenhuis et al. (1995) and Solari et al. (2003). All these results, together with those that reveal adequate psychometric characteristics, support the validity of the Spanish version and its usefulness as a good measure to diagnose depression in Spanish-speaking PwMS.

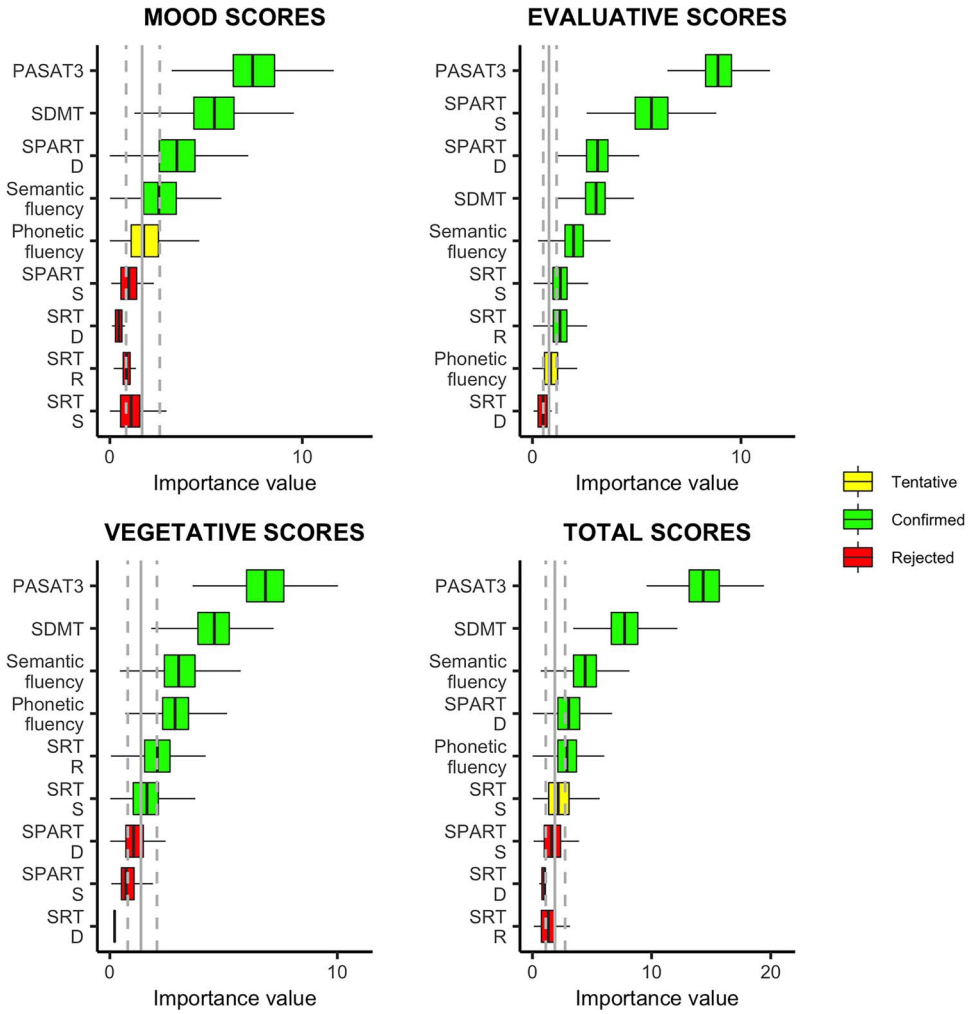

Fig. 4. Predictive importance of cognitive variables in CMDI scores. Each panel depicts the importance profile of each CMDI subscale (mood, evaluative, vegetative) or CMDI total scores (2,000 runs, $p < .001$). Within each panel, green/yellow/red boxplots illustrate the distribution of the estimated importance value of cognitive variables confirmed as relevant/tentative and non-relevant predictors of CMDI scores, respectively. The vertical line depicts the median, whereas the dotted lines correspond to the values of the first and third quartile of the maximum shadow attribute.

As in previous studies using the CMDI (e.g., Arnett et al., 2008; Nyenhuis et al., 1995; Solari et al., 2003), our results also confirmed that PwMS have higher CMDI scores than HCs. These differences were observed on all the CMDI subscales (mood, evaluative, and vegetative), on the total CMDI scores, and when comparing the percentage of P-depressed individuals. The size of these differences can be considered "moderate-to-large" (Cohen, 1969), and they confirm the association between MS and enhanced depressive symptomatology (Diamond, Johnson, Kaufman, & Graves, 2008; Solaro et al., 2018).

In addition, we provide the first evidence indicating that the severity of depressive symptomatology in PwMS differs across MS subtypes. More specifically, we found that the MS progressive phenotypes were associated with higher CMDI scores (on all subscales) than those observed in RR PwMS. In contrast, CIS individuals exhibited lower CMDI scores than RR PwMS, and their scores on all the CMDI subscales were very similar to those of HC. These results are not surprising because progressive forms of MS are associated with more advanced age, more years of disease evolution, larger motor and sensory deficits, a higher degree of cognitive impairment, and greater fatigue (Dekker et al., 2019; Lorefice et al., 2015; Planche et al., 2016; Siegert & Abernethy, 2005).

Moreover, we corroborated that depressive symptomatology is closely associated with fatigue and, to a lesser extent, with physical disability (EDSS scores). The association between depression, EDSS, and fatigue scores has been described in previous studies (Lorefice et al., 2015; Patten, Marrie, & Carta, 2017; Pittion-Vouyovitch et al., 2006; Siegert & Abernethy, 2005; Solaro et al., 2016). Thus, as we also observed, other studies showed that fatigue is correlated with disability

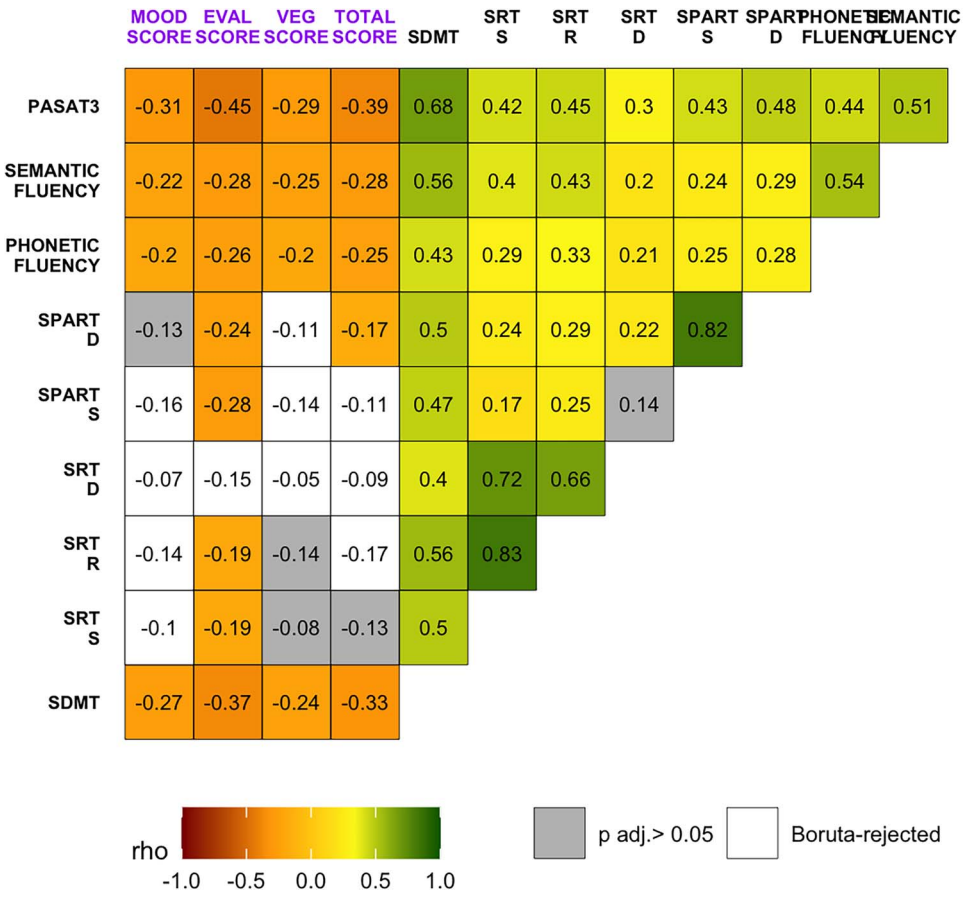

Fig. 5. Correlations between cognitive variables and CMDI scores. The figure summarizes the values of the ordinal correlations (Spearman's rho) between cognitive variables and the CMDI scores. The correlations that remain significant after multiple testing corrections are depicted using a gradient scale, whereas those that did not or that had been rejected as possible predictors by the Boruta algorithm (Fig. 4) are depicted in gray and white colored cells, respectively.

(Jones, Motl, & Sandroff, 2021) and that both disability and fatigue are correlated with depression measures (Arnett et al., 2008, 2021; Jones et al., 2021; Krupp, Serafin, & Christodoulou, 2010). In addition, and highlighting the importance of using psychometric instruments that allow us to independently evaluate different dimensions of depression, we observed that fatigue and physical disability were more clearly, but not exclusively, associated with vegetative depressive symptomatology. This observation confirms that depression is related to fatigue and physical disability, even when somatic symptomatology is excluded from rating scales (Brenner & Piehl, 2016). Therefore, the increased prevalence of depression in PwMS cannot only be attributed to a possible confound between the vegetative symptoms associated with both disorders.

Our study also shows that higher CMDI scores are associated with worse cognitive performance in PwMS. The relationship between depressive symptomatology and cognition has been well-established (for a review, see Arnett et al., 2008; Siegert & Abernethy, 2005), and it is more noteworthy in "cognitively effortful" tasks that require information processing speed, attention, working memory, and executive functions (Arnett, Higginson, Voss, Bender, et al., 1999; Arnett, Higginson, Voss, Wright, et al., 1999; Arnett et al., 2008; Arnett, Higginson, & Randolph, 2001; Diamond et al., 2008; Hoffmeister et al., 2021; Takeda et al., 2021). In agreement with these previous studies, we observed that the inverse correlations between CMDI scores and performance on the PASAT and SDMT were larger than those observed for other cognitive tests. Of note, and again highlighting the importance of using psychometric instruments to independently evaluate different dimensions of depression, we also observed that the size of these associations was larger when considering the scores on the Evaluative subscale of the CMDI than when considering the scores of the same participants on the Mood or Vegetative subscales.

In conclusion, the present study provides a valid Spanish-adapted version of the CMDI that yields similar results to those obtained with other previous versions of this inventory (Nyenhuis et al., 1995; Solari et al., 2003). Using this new psychometric instrument, we found that the prevalence of depression in PwMS is higher than in HCs but also that different dimensions of depressive symptomatology (which are separately evaluated by the three CMDI subscales) are differentially associated with

Table 5. Differences between P-depressed and N-depressed PwMS in cognitive variables

	Mood subscale		Evaluative subscale		Vegetative Subscale		Total scores	
	P-dep	N-dep	P-dep	N-dep	P-dep	N-dep	P-dep	N-dep
Paced Auditory Serial Addition test 3 s	25.8 (18.9)	36.5 ((17.8)	21.2 (17.6)	37.5 (17.2)	26.9 (18.5)	35.8 (18.2)	24 (17.8)	37.4 (17.6)
	$t = -3.4$, $p = .011$		**$t = -5.28$, $p < .001$**		**$t = -2.76$, $p = .025$**		**$t = -4.55$, $p < .001$**	
	$d = -0.59$ [−0.93, −0.25]		**$d = -0.94$ [−1.29, −0.58]**		**$d = -0.49$ [−0.84, −0.14]**		**$d = -0.75$ [−1.09, −0.42]**	
Symbol Digit Modalities Test	41.8 (13.9)	47.6 (13.8)	39.2 (15.4)	48.1 (13)	42.1 (14.6)	47.3 (13.7)	40.8 (13.9)	48.1 (13.6)
	$t = -2.46$, $p = .039$		**$t = -3.42$, $p = .005$**		$t = -2.05$, $p = .112$		**$t = -3.18$, $p = .005$**	
	$d = -0.42$ [−0.75, −0.08]		**$d = -0.66$ [−1.01, −0.31]**		**$d = -0.37$ [−0.72, 0.002]**		**$d = -0.53$ [−0.86, −0.2]**	
Semantic fluency	18.1 (4.9)	20.8 (5.6)	17.6 (5.1)	29.8 (5.5)	17.8 (4.9)	20.8 (5.6)	17.3 (5.1)	21.1 (5.4)
	$t = -3.04$, $p = .015$		**$t = -3.55$, $p = .004$**		**$t = -3.36$, $p = .006$**		**$t = -4.44$, $p < .001$**	
	$d = -0.48$ [−0.81, −0.14]		**$d = -0.59$ [−0.94, −0.24]**		**$d = -0.55$ [−0.90, −0.21]**		**$d = -0.71$ [−1.04, −0.38]**	
Phonetic fluency	9.9 (4)	11.8 (5.5)	9.7 (4.4)	11.8 (5.4)	9.5 (3.4)	11.9 (5.6)	9.4 (4.2)	12.1 (5.5)
	$t = -2.65$, $p = .030$		**$t = -2.59$, $p = .016$**		**$t = -3.47$, $p = .006$**		**$t = -3.61$, $p = .006$**	
	$d = -0.38$ [−0.72, −0.05]		**$d = -0.40$ [−0.75, −0.06]**		**$d = -0.47$ [−0.82, 0.12]**		**$d = -0.53$, [−0.86, −0.2]**	
Spatial Recall Test Long-Term Storage	19.3 (5.3)	19.4 (2.3)	17.3 (5.2)	20 (5.1)	18.9 (5.4)	19.6 (5.2)	18.4 (5.1)	19.8 (5.2)
	$t = -0.09$, $p = .980$		**$t = -3.01$, $p = .009$**		$t = -0.54$, $p = .667$		$t = -1.58$, $p = .169$	
	$d = -0.02$ [−0.35, 0.32]		**$d = -0.53$ [−0.88, −0.18]**		$d = -0.13$ [−0.47, 0.22]		$d = -0.26$ [−0.58, 0.07]	
Spatial Recall Test Delayed-Recall	6.8 (2)	6.8 (2.3)	6.2 (2)	7 (2.2)	6.6 (2)	6.8 (2.3)	6.5 (1.9)	6.9 (2.3)
	$t = -0.02$, $p = .981$		**$t = -2.14$, $p = .044$**		$t = -0.71$, $p = .667$		$t = -1.03$, $p = .336$	
	$d = 0.0$ [−0.34, 0.33]		**$d = -0.35$ [−0.7, −0.01]**		$d = -0.09$ [−0.43, 0.-26]		$d = -0.16$ [−0.48, 0.17]	
Selective Reminding Test Long-Term Storage	39.9 (12.2)	41.4 (13.3)	37.7 (15.4)	42 (12.9)	40.4 (11.6)	41.3 (13.4)	39 (12)	41.8 (13.3)
	$t = -0.73$, $p = .67$		$t = -1.91$, $p = .066$		$t = -0.40$, $p = .689$		$t = -1.37$, $p = .215$	
	$d = -0.12$ [−0.45, −0.22]		$d = -0.34$ [−0.68, 0.01]		$d = -0.06$ [−0.41, 0.28]		$d = -0.22$ [−0.54, 0.11]	
Selective Reminding Test Long-Term Retrieval	28.4 (13.3)	32.4 (14.1)	26.8 (12.6)	32.8 (14)	28.9 (11.8)	32.2 (14.5)	28 (12.6)	32.7 (14.2)
	$t = -1.78$, $p = .129$		**$t = -2.64$, $p = .02$**		$t = -1.49$, $p = .231$		$t = -2.16$, $p = .055$	
	$d = -0.29$ [−0.63, −0.04]		**$d = -0.43$ [−0.78 -0.09]**		$d = -0.23$[−0.58, 0.11]		**$d = -0.34$ [−0.66, −0.01]**	
Selective Reminding Test Delayed Recall	8.1 (2.1)	7.9 (2.3)	7.8 (2.1)	8 (2.3)	7.8 (2.3)	8 (2.5)	7.9 (2.1)	8 (2.3)
	$t = -0.33$, $p = .924$		$t = -0.65$, $p = .521$		$t = -0.53$, $p = .667$		$t = -0.43$, $p = .669$	
	$d = 0.05$ [−0.29, 0.39]		$d = -0.11$ [−0.45, 0.24]		$d = -0.09$ [−0.44, 0.25]		$d = -0.07$ [−0.39, 0.26]	
Matrix subtest	14.9 (5.1)	16.7 (4.1)	14.5 (4.8)	16.8 (4.2)	15.2 (4.6)	16.6 (4.3)	14.8 (4.8)	16.8 (4.2)
	$t = -2.16$, $p = .067$		**$t = -2.78$, $p = .014$**		$t = -1.80$, $p = .153$		**$t = -2.58$, $p = .023$**	
	$d = -0.41$ [−0.74, −0.07]		**$d = -0.53$ [−0.87, −0.18]**		$d = -0.33$ [−0.68, 0.02]		**$d = -0.45$ [−0.78, −0.13]**	

This table displays the means and standard deviation (between brackets) of the values observed in P-depressed PwMS and N-depressed PwMS (as defined by each CMDI subscale or total scores) in all the tests included in the BRB-N as well as the Matrix Reasoning subtest of the WAIS-III battery. For all variables, differences between groups were tested by means of Student's *t* tests and their size was estimated in terms of Cohen's *d* and its 95% CI. The *p*-values associated to these *t*-tests are FDR-corrected for all comparisons included within each column, whereas the 95% CI for Cohen's *d* that do not continue the zero value allow identifying between-means differences that were significant at uncorrected $p < .05$. All statistically significant differences are highlighted in bold.

specific handicaps and sequelae characteristic of this clinical population, such as physical disability, fatigue, and cognitive impairment. These findings are relevant for improving the diagnosis of depression in PwMS because they help to clarify its relationship with physical and cognitive deficits in MS. Nevertheless, the present study is not devoid of limitations and, therefore, some caution should be applied to our results and conclusions. In particular, it should be noted that, although it cannot be considered "small," our sample could be suboptimal for some of the performed statistical analyses (e.g., CFA) and it did not allow performing separate analyses for different MS subtypes.

Acknowledgements

We appreciate the collaboration of the Associació Esclerosis Múltiple de Castelló (AEMC).

Supplementary material

Supplementary material is available at *Archives of Clinical Neuropsychology* online.

Funding

MCI/AEI (UJI B2020-02 to CF and CS-S); UJI (PREDOC/2020/22 to SFA). These funding sources did not play any role in designing the study or in the collection, analysis, and interpretation of the data.

Conflict of interest

The authors declare no conflict of interest.

References

Arnett, P. A., Barwick, F. H., & Beeney, J. E. (2008). Depression in multiple sclerosis: Review and theoretical proposal. *Journal of the International Neuropsychological Society*, *14*(*5*), 691–724. https://doi.org/10.1017/S1355617708081174.

Arnett, P. A., Cadden, M., Roman, C. A. F., Guty, E. T., Riegler, K. E., & Thomas, G. A. (2021). Sensory-motor and affective-fatigue factors are associated with symbol digit performance in multiple sclerosis. *Journal of the International Neuropsychological Society*, *28*(*4*), 362–370. https://doi.org/10.1017/S1355617721000540.

Arnett, P. A., Higginson, C. I., & Randolph, J. J. (2001). Depression in multiple sclerosis: Relationship to planning ability. *Journal of the International Neuropsychological Society*, *7*(*6*), 665–674. https://doi.org/10.1017/s1355617701766027.

Arnett, P. A., Higginson, C. I., Voss, W. D., Bender, W. I., Wurst, J. M., & Tippin, J. M. (1999). Depression in multiple sclerosis: Relationship to working memory capacity. *Neuropsychology*, *13*(*4*), 546–556. https://doi.org/10.1037/0894-4105.13.4.546.

Arnett, P. A., Higginson, C. I., Voss, W. D., Wright, B., Bender, W. I., Wurst, J. M. et al. (1999). Depressed mood in multiple sclerosis: Relationship to capacity-demanding memory and attentional functioning. *Neuropsychology*, *13*(*3*), 434–446. https://doi.org/10.1037/0894-4105.13.3.434.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, *57*(*1*), 289–300. https://doi.org/10.1111/J.2517-6161.1995.TB02031.X.

Brenner, P., & Piehl, F. (2016). Fatigue and depression in multiple sclerosis: Pharmacological and non-pharmacological interventions. *Acta Neurologica Scandinavica*, *134*, 47–54. https://doi.org/10.1111/ANE.12648.

Brown, R. F., Valpiani, E. M., Tennant, C. C., Dunn, S. M., Sharrock, M., Hodgkinson, S. et al. (2009). Longitudinal assessment of anxiety, depression, and fatigue in people with multiple sclerosis. *Psychology and Psychotherapy: Theory, Research and Practice*, *82*, 41–56. https://doi.org/10.1348/147608308X345614.

Bruce, J. M., Hancock, L. M., Arnett, P., & Lynch, S. (2010). Treatment adherence in multiple sclerosis: Association with emotional status, personality, and cognition. *Journal of Behavioral Medicine*, *33*(*3*), 219–227. https://doi.org/10.1007/s10865-010-9247-y.

Carlson, M., & Mulaik, S. A. (1993). Trait ratings from descriptions of behavior as mediated by components of meaning. *Multivariate Behavioral Research*, *28*(*1*), 111–159. https://doi.org/10.1207/s15327906mbr2801_7.

Chang, C. H., Nyenhuis, D. L., Cella, D., Luchetta, T., Dineen, K., & Reder, A. T. (2003). Psychometric evaluation of the Chicago multiscale depression inventory in multiple sclerosis patients. *Multiple Sclerosis Journal*, *9*(2), 160–170. https://doi.org/10.1191/1352458503ms885oa.

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. Academic Press. https://www.utstat.toronto.edu/~brunner/oldclass/378f16/readings/CohenPower.pdf

Dekker, I., Leurs, C. E., Hagens, M. H. J., van Kempen, Z. L. E., Kleerekooper, I., Lissenberg-Witte, B. I. et al. (2019). Long-term disease activity and disability progression in relapsing-remitting multiple sclerosis patients on natalizumab. *Multiple Sclerosis and Related Disorders*, *33*, 82–87. https://doi.org/10.1016/j.msard.2019.05.017.

Diamond, B. J., Johnson, S. K., Kaufman, M., & Graves, L. (2008). Relationships between information processing, depression, fatigue and cognition in multiple sclerosis. *Archives of Clinical Neuropsychology*, *23*(*2*), 189–199. https://doi.org/10.1016/j.acn.2007.10.002.

Flora, D. B. (2020). Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Advances in Methods and Practices in Psychological Science*, *3*(*4*), 484–501. https://doi.org/10.1177/2515245920951747.

Hind, D., Kaklamanou, D., Beever, D., Webster, R., Lee, E., Barkham, M. et al. (2016). The assessment of depression in people with multiple sclerosis: A systematic review of psychometric validation studies. *BMC Psychiatry*, *16*, 278. https://doi.org/10.1186/s12888-016-0931-5.

Hittner, J. B., May, K., & Silver, N. C. (2003). A Monte Carlo evaluation of tests for comparing dependent correlations. *The Journal of General Psychology*, *130*(*2*), 149–168. https://doi.org/10.1080/00221300309601282.

Hoffmeister, J., Basso, M. R., Reynolds, B., Whiteside, D., Mulligan, R., Arnett, P. A. et al. (2021). Effects of diminished positive mood and depressed mood upon verbal learning and memory among people with multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *43*(2), 117–128. https://doi.org/10.1080/13803395.2020.1853066.

Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, *6*(*1*), 53–60. https://doi.org/10.21427/D7CF7R.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(*1*), 1–55. https://doi.org/10.1080/10705519909540118.

Jones, C. D., Motl, R., & Sandroff, B. M. (2021). Depression in multiple sclerosis: Is one approach for its management enough? *Multiple Sclerosis and Related Disorders*, *51*, 102904. https://doi.org/10.1016/j.msard.2021.102904.

Krupp, L. B., Larocca, N. G., Muir Nash, J., & Steinberg, A. D. (1989). The fatigue severity scale: Application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology*, *46*(*10*), 1121–1123. https://doi.org/10.1001/archneur.1989.00520460115022.

Krupp, L. B., Serafin, D. J., & Christodoulou, C. (2010). Multiple sclerosis-associated fatigue. *Expert Review of Neurotherapeutics*, *10*(*9*), 1437–1447. https://doi.org/10.1586/ern.10.99.

Kursa, M. B., Jankowski, A., & Rudnicki, W. R. (2010). Boruta - a system for feature selection. *Fundamenta Informaticae*, *101*(*4*), 271–285. https://doi.org/10.3233/FI-2010-288.

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software*, *36*(*11*), 1–13. https://doi.org/10.18637/jss.v036.i11.

Lorefice, L., Fenu, G., Trincas, G., Moro, M. F., Frau, J., Coghe, G. C. et al. (2015). Progressive multiple sclerosis and mood disorders. *Neurological Sciences*, *36*(*9*), 1625–1631. https://doi.org/10.1007/s10072-015-2220-3.

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, *52*(*2*), 464–488. https://doi.org/10.3758/s13428-019-01246-w.

Nyenhuis, D. L., Luchetta, T., Yamamoto, C., Terrien, A., Bernardin, L., Rao, S. M. et al. (1998). The development, standardization, and initial validation of the Chicago multiscale depression inventory. *Journal of Personality Assessment*, *70*(*2*), 386–401. https://doi.org/10.1207/s15327752jpa7002_14.

Nyenhuis, D. L., Rao, S. M., Zajecka, J. M., Luchetta, T., Bernardin, L., & Garron, D. C. (1995). Mood disturbance versus other symptoms of depression in multiple sclerosis. *Journal of the International Neuropsychological Society*, *1*(*3*), 291–296. https://doi.org/10.1017/S135561770000028X.

Patten, S. B., Marrie, R. A., & Carta, M. G. (2017). Depression in multiple sclerosis. *International Review of Psychiatry*, *29*(*5*), 463–472. https://doi.org/10.1080/09540261.2017.1322555.

Pittion-Vouyovitch, S., Debouverie, M., Guillemin, F., Vandenberghe, N., Anxionnat, R., & Vespignani, H. (2006). Fatigue in multiple sclerosis is related to disability, depression and quality of life. *Journal of the Neurological Sciences*, *243*(*1–2*), 39–45. https://doi.org/10.1016/j.jns.2005.11.025.

Planche, V., Gibelin, M., Cregut, D., Pereira, B., Clavelou, P., & Montpied, G. (2016). Cognitive impairment in a population-based study of patients with multiple sclerosis: Differences between late relapsing-remitting, secondary progressive and primary progressive multiple sclerosis. *European Journal of Neurology*, *23*, 282–289. https://doi.org/10.1111/ene.12715.

R Core Team (2020). *R: A language and environment for statistical computing*. 2015. R foundation for statistical computation, Vienna, Austria. https://www.R-project.org.

Sepulcre, J., Vanotti, S., Hernández, R., Sandoval, G., Cáceres, F., Garcea, O. et al. (2006). Cognitive impairment in patients with multiple sclerosis using the brief repeatable battery-neuropsychology test. *Multiple Sclerosis Journal*, *12*(*2*), 187–195. https://doi.org/10.1191/1352458506ms1258oa.

Siegert, R. J., & Abernethy, D. A. (2005). Depression in multiple sclerosis: A review. *Journal of Neurology, Neurosurgery & Psychiatry*, *76*(*4*), 469–475. https://doi.org/10.1136/jnnp.2004.054635.

Solari, A., Motta, A., Mendozzi, L., Aridon, P., Bergamaschi, R., Ghezzi, A. et al. (2003). Italian version of the Chicago multiscale depression inventory: Translation, adaptation and testing in people with multiple sclerosis. *Neurological Sciences*, *24*(*6*), 375–383. https://doi.org/10.1007/S10072-003-0192-1.

Solaro, C., Gamberini, G., & Fabio, G. M. (2018). Depression in multiple sclerosis: Epidemiology, aetiology, diagnosis and treatment. *CNS Drugs*, *32*, 117–133. https://doi.org/10.1007/s40263-018-0489-5.

Solaro, C., Trabucco, E., Signori, A., Martinelli, V., Radaelli, M., Centonze, D. et al. (2016). Depressive symptoms correlate with disability and disease course in multiple sclerosis patients: An Italian multi-center study using the Beck depression inventory. *PLoS One*, *11*(*9*), e0160261. https://doi.org/10.1371/journal.pone.0160261.

Spence, J. R., & Stanley, D. J. (2016). Prediction interval: What to expect when You're expecting. A replication. *PLoS One*, *11*(*9*), e0162874. https://doi.org/10.1371/journal.pone.0162874.

Takeda, A., Minatani, S., Ishii, A., Matsuo, T., Tanaka, M., Yoshikawa, T. et al. (2021). Impact of depression on mental fatigue and attention in patients with multiple sclerosis. *Journal of Affective Disorders Reports*, *5*, 100143. https://doi.org/10.1016/j.jadr.2021.100143.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281–324). American Psychological Association. https://doi.org/10.1037/10222-009.