Validation of the French version of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS)

Nicolas Maubeuge[1], Mathilde S.A. Deloire, PhD,[2] Bruno Brochet, MD,[2,3,4,] Nathalie Ehrlé, PhD,[5] Julie Charré-Morin Msc,[2] Aurore Saubusse, Msc,[2] Aurélie Ruet, MD, PhD.[2,3,4]

For the BICAFMS study investigators.[6]

[1] Univ. Poitiers, F-86000, France

[2]CHU de Bordeaux, F-33000 Bordeaux, France

[3] Inserm U1215 - Neurocentre Magendie, F-33000 Bordeaux, France

[4] Univ. Bordeaux, F-33076 Bordeaux, France

[5]CHU de Reims, F-51000, Reims, France

[6]BICAFMS study investigators : Pr Defer, Mme Mondou, Caen; Pr Clavelou, Mme Cregut, Clermont-Ferrand; Pr Moreau, Mme Quaglio, Dijon; Dr Nkendjuo, Mme Leclercq, Dunkerque; Pr Hautecoeur, M. Lenne, Lille; Pr Vermersch, Mme Jougleux, Lille; Pr Pelletier, Mme Reuter, Marseille; Pr Labauge, M. Trauchessec, Montpellier; Pr Debouverie, Mme Morele, Nancy; Pr Lebrun-Frenay, Mme Joly, Nice; Dr Heinzhef, Mme Tissier, Poissy Saint Germain; Pr Bakchine, Mme Lecrique, Reims

Word count : 3186

Abstract word count : 232

Number of tables : 5

Number of figure : 1

Corresponding author: Professor Bruno Brochet, Service de Neurologie, CHU de Bordeaux, 33076 Bordeaux cedex, France. bruno.brochet@u-bordeaux.fr

T : 335567795521

ORCID : Prof B Brochet https://orcid.org/0000-0003-3824-2796

Conflict of Interest Statement/ The authors did not report any conflict of interest in relation to the study.

Dr. RUET reports grants, personal fees and non-financial support from genzyme, grants and other from teva, grants from bayer, personal fees and non-financial support from medday, grants, non-financial support and other from ROCHE, personal fees, non-financial support and other from biogen, personal fees, non-financial support and other from NOVARTIS, non-financial support from merck, outside the submitted work .

Dr. Brochet reports grants, research support, personal fees or non-financial support from Genzyme, Bayer, Medday, Actelion, Roche, Biogen, Celgene, Novartis, and Merck outside the submitted work.

Other authors did not disclose conflict of interest.

KEY WORDS : MULTIPLE SCLEROSIS, COGNITION, NEUROPSYCHOLOGY, MACFIMS, VALIDATION

The English in this document has been checked by at least two professional editors, both native speakers of English. For a certificate, please see:
http://www.textcheck.com/certificate/uk8vcr

HIGHLIGHTS:

- MACFIMS regression-based norms were established in French speaking healthy subjects
- 32.7% of persons with MS had cognitive impairment according to MACFIMS.
- Learning and processing speed were the most affected cognitive functions.

ABSTRACT

Background: The Minimal Assessment of Cognitive Function in Multiple sclerosis (MACFIMS) is an internationally recognised battery of neuropsychological tests for patients with multiple sclerosis (MS).

Objectives: To establish regression-based norms for the MACFIMS in French-speaking healthy subjects (HS) and validate its use in persons with multiple sclerosis (PwMS).

Methods: 136 PwMS, including 43 with relapsing-remitting MS, 46 with secondary progressive MS and 45 with primary progressive MS, as well as 276 HS were enrolled. Regression-based norms and validity were established for the seven tests of the MACIMS: the Symbol Digit Modalities Test (SDMT), the Paced Auditory Serial Addition Test (PASAT), the French learning test (FLT) a French-adapted memory test (or the California Verbal Learning Test (CVLT) at re-testing), the Judgment of Line Orientation Test (JLO), the 'épreuve de classement de cartes de Champagne' (ECCC), a French adaptation of the DKEF-sorting test, the Brief Visuospatial Memory Test (BVMT-R) and the Controlled Oral Word Association Test (COWAT).

Results: Regression-based norms of MACFIMS tests were established in the HS population. The MACFIMS battery was able to identify cognitive impairment (CI) (at least two abnormal tests in different domains) in 32.7% of PwMS. The domains with more frequent impairment were (in descending order): learning followed by IPS, delayed memory, verbal fluency and working memory.

Conclusion: This study established the regression-based norms for French subjects of the French adaptation of the MACFIMS and its validity in PwMS.

1. INTRODUCTION

Despite the influence of cognitive impairment (CI) on activities of daily living in patients with multiple sclerosis (MS) (Ruet et al.,2012, Langdon, 2011), its detection and assessment remain limited in clinical practice (Langdon, 2011; Deloire et al., 2006). The frequency of CI has been reported to be between 30% and 70% of persons with MS (PwMS) (Langdon, 2011; Brochet and Ruet, 2019). Considering its negative impact on several aspects of the lives of PwMS, including working retention (Ruet et al., 2012; Morrow et al., 2010), the detection of CI is important. Moreover, as CI has been shown to predict subsequent physical disability, probably because it is an indicator of the on-going neurodegenerative process (Deloire et al., 2010), its early detection should be part of the clinical work-up in PwMS. Over the last several decades, several internationally recognised batteries of neuropsychological (NP) tests for MS have been validated in English, including the Brief Repeatable Battery of neuropsychological tests (Rao, 1990) and the Minimal Assessment of Cognitive Function in MS (MACFIMS) (Benedict et al., 2002).

The MACFIMS battery was proposed by a panel of experts in the field who met in April 2001, based on scientific literature published at that time (Benedict et al., 2002). It includes seven NP tests covering the main cognitive domains or functions affected by the disease. The NP tests had to fulfil several criteria for inclusion in the battery, i.e. standardisation, normative data, adequate range (absence of ceiling or floor effects), reliability, validity (i.e. able to discriminate PwMS from healthy subjects [HS]), existence of alternate forms and practicality (easy to administer) (Benedict et al., 2002). The seven NP tests assess working memory (WM), information processing speed (IPS), verbal and visual learning and memory, executive function (EF), visual perception, spatial processing and verbal fluency. MACFIMS has been validated for English-speaking subjects ((Benedict et al., 2006) as well as some other

languages (Dusankova et al., 2012; Eshagi et al., 2012; Argento et al., 2018) but not for French-speaking subjects, and there are no normative data for most of the individual tests for French-speaking subjects.

This study was performed to validate the MACFIMS in France, including the establishment of normative data in a sample of HS and its validation for detecting CI in a sample of PwMS.

2. SUBJECTS AND METHODS

2.1 Study design

This study was approved by the institutional review board for human subject research of Bordeaux (Comité de Protection des Personnes (CPP) Sud-Ouest et Outre-Mer; No. 2014/95). This study was part of the BICAFMS study (Clinical Trial.gov NCT02391064) (Maubeuge et al., 2020).

2.2 Subjects

All participants provided written informed consent and were registered in the French Social Security system. They were native French speakers and not under 18 or over 64 years old.

2.2.1 Control population

Normative data were established in a sample of HS recruited in one centre, by advertising, who were paid to participate in the study. Exclusion criteria were history of neurological disease or psychiatric illness, on-going psychotropic drug use, other significant chronic disease, abuse of alcohol or drugs, on-going pregnancy, cognitive complaints or having participated in a cognitive study in the last year. HCs were recruited and divided into 16 groups according to age (18–34, 35–44, 45–54, 55–64), sex, and education level (secondary education [usually 12 years of schooling] and graduated [at least baccalaureate]).

2.2.2 Patients

PwMS fulfilling the MS diagnostic criteria (Polman et al., 2011) were recruited from 15 MS centres in France between February 2015 and June 2017. Exclusion criteria were similar to those for HS with the addition that PwMS had not started, changed dosage or stopped taking psychotropic drugs in the previous 2 months, nor had they started, changed dosage or stopped disease-modifying therapy (DMT) or received steroids in the previous 1 month. They did not have neurological motor, visual or sensory impairment that could preclude their ability to perform cognitive tests. There was no limitation regarding MS phenotype, i.e. relapsing-remitting, primary progressive or secondary progressive MS (RRMS, SPMS or PPMS, respectively) (Lublin et al., 2014).

2.3 Methods

2.3.1 Neuropsychological assessment

PwMS and HS were evaluated by qualified senior neuropsychologists. At baseline, the assessment included all of the tests in the MACFIMS battery. At 1 month for all patients and 50% of HS the assessment included the same cognitive evaluation as applied at baseline.
The MACFIMS battery has been described in detail previously (Benedict et al., 2002; 2006). The battery included the Symbol Digit Modalities Test, oral version (SDMT) (Smith, 1982), which mainly reflects IPS; the Paced Auditory Serial Addition Test with 3.0-s interstimulus interval, PASAT-3) (Gronwall, 1977) to assess working memory and IPS; a French adaptation (French Learning Test [FLT], formerly named the French Verbal Learning Test) (Maubeuge et al., 2020) and the French published version of the California Verbal Learning Test (CVLT) (Poitrenaud et al., 2007) for verbal learning and memory; the Judgment of Line Orientation Test (JLO) (Benton et al., 1994) to test visual perception and spatial processing; a French adaptation of the D-KEFS sorting test (Delis et al., 2001), the ‘épreuve de classement de cartes de Champagne’ (ECCC) developed by the Reims group; the Brief Visuospatial

Memory Test–Revised (BVMTR) ((Benedict et al., 1997) for visuospatial learning; and the Controlled Oral Word Association Test (COWAT) (Benton et al., 1983) for phonemic verbal fluency (letters P, R and S). The edited CVLT French version (Poitrenaud et al., 2007) did not include an alternative form. We used the recently developed FLT (Maubeuge et al., 2020) at the baseline visit and the published CVLT form at the follow-up visit. The Reims group created this alternate form of the French version following the same methodology. This version (courtesy of Dr Ehrlé) was pretested in a group of 20 HS. The ECCC is a measure of conceptual reasoning that permits the differentiation of concept formation from conceptual flexibility and is used to assess higher executive function. This test is a French adaptation of the DKEFS Sorting test. The cards are different but the structure of the test and the scoring are similar. The goal of this executive test is to isolate and measure specific components of problem-solving capabilities. The test has two conditions. In the first, the subject must classify six cards spontaneously in two columns of three cards each according to the most possible different rules (maximum eight) and indicate the rule after each ranking. The cards are varied in many ways, allowing conceptual sorting in accordance with at least eight different principles (e.g. card shape, card colour, semantic associations among words). The ECCC is a timed test with 4 minutes allowed for each of two card sets. The dependent variable considered here was the total number of correct sorts (CS). In the second condition, the subject was required to verbalise the rankings made by the examiner with a maximum duration of 45 s per classification (eight classifications per package). The dependent variable considered here was the verbal description score (DS). We used the Total Learning (TL) and Delayed Recall (DR) indices from the FLT and BVMT-R and both the Total Correct Sorts (CS) and the Description Score (DS) from the ECCC Sorting Test. Alternative forms were use at 1 month for the SDMT, the CVLT, the BVMT and the ECCC.

2.3.2 Neurological evaluation

Neurological disability was measured using the French version of the Expanded Disability Status Scale (EDSS) (Brochet et al., 2009).

2.3.3 Patient-related outcome (PRO)

All subjects completed patient-related outcome (PRO) questionnaires concerning depressive symptoms (Beck Depression Inventory-Fast Screen [BDI-FS]) (Benedict et al.,, 2003), anxiety (State Trait Anxiety Inventory [STAI]) (Spielberger et al., 1993), mood (Echelle d'Humeur Dépressive [EHD-PRO]) (Lamargue-Hamel et al., 2015) and subjective fatigue (French version of the Fatigue Impact Scale [EMIF-SEP]) (Debouverie et al., 2007).

2.3.4 Statistical analysis

All statistical analyses were carried out using SPSS version 23 (SPSS Inc., Chicago, IL, USA). Between-group differences were examined with Student's *t* test and the χ^2 test for continuous and categorical variables, respectively. Effect size was calculated with Cohen's D statistic. The results were considered significant at an α-level of < 0.05. Regression-based norms were created according to the procedures of Parmenter et al. (2010) but with the adjustment described by Berrigan et al. (2014) for centring the age variable (age minus the mean age) to prevent multicollinearity and for ease of interpretation. First, the raw scores of the HS were converted to a scaled score metric (mean 10, SD 3) based on the cumulative frequency distribution of each measure presented in Table 2. This served to normalise all of the test score distributions. The percentile range encompassed by each scaled score was set such that the resultant distribution of scaled scores was as normal as possible. We adopted the data analysis and presentation approach of the Wechsler Intelligence Scale for Children-revised (Ivnik et al., 1992). Next, the resulting scaled scores were regressed on four demographic variables: centred age ($age_c = age - age_{mean}$ (43.83)), age_c^2, sex (female vs.

male) and educational attainment. The factor age_c^2 was added to account for possible non-linear effects of age on test performance.

The participants were divided into four groups according to educational level: low educational level (LEL) group, individuals who did not complete secondary education, which usually requires 12 years of schooling; high educational level (HEL) group, individuals who completed secondary education or graduated with a 'baccalauréat' level (equivalent to A levels in the UK) or higher. The HEL group was further divided into three subgroups: HEL-BAC, subjects who graduated with a 'baccalauréat' degree; HEL-BAC+2-4, subjects with ≥ 2 years but < 5 years of secondary education; and HEL-BAC+5, subjects with ≥ 5 years of secondary education.

Next, predicted scores were calculated by multiple regression analysis based on β weight values for all four demographic variables and their predictive constants:

$\underline{\textit{Scaled score predicted}} = \textit{constant} + \beta age_c\ (age_c) + \beta age_c^{\,2}\ (age_c^2) + \beta sex\ \textit{(sex)} + \beta education\ \textit{(education)}$

The predicted scaled scores were subtracted from each participant's actual obtained scaled score and the difference was divided by the root mean squared error (RMSE) of the HS:

2.3.5 $\underline{\textit{z-score = (scaled score predicted − scaled score actual)/RMSE.}}$

Tests with z-score ≤ −1.5 were considered 'impaired'.

None of the assumptions of regression analysis were violated (no influential cases, normality of the residuals or homoscedasticity).

For the BDI-FS, which was not normally distributed in the MS sample, the following cut-off scores were used, consistent with the test manual: Normal = BDI-FS < 3; Borderline = BDI-FS 3 − 8; Depressed = BDI-FS > 8.

3. RESULTS

3.1 Clinical and demographic characteristics of participants

The clinical characteristics of the 276 HS and 134 PwMS participating in the study are presented in Table 1.

The PwMS population consisted of 43 with RRMS, 46 with SPMS and 45 with PPMS. Comparison of demographic characteristics indicated no significant between-group differences in education or sex but age was significantly different between the groups ($p < 0.001$). Self-reported questionnaires for depression, anxiety and fatigue showed higher scores in PwMS ($p < 0.001$) than in HS.

3.2 MACFIMS impairment rates using regression-based norms

Table 2 shows scale conversion and Table 3 shows the regression model for each test used to derive demographically adjusted z-scores for the PwMS, including age_c (age – 43.84), age_c^2, sex and education as predictive variables.

Using this formula and the coefficients from Table 3, we can for example calculate the predicted SDMT score of a 42-year-old female (male: sex = 1; female: sex = 2) PwMS with a 'baccalauréat' degree (1):

$$8.91 - 0.09 \times (\mathbf{42 - 43.84}) - 0.002 \times (\mathbf{42 - 43.84})^{\mathbf{2}} + 0.65 \times \mathbf{2} + 0.36 \times \mathbf{1} = 10.73$$

Values for participants' educational level in the calculation are 0: Low education level, 1: Bachelor's degree, 2: 2 to 4 years of secondary education, 3: > 4 years of secondary education. We can then convert the patient's actual score on the SDMT (e.g. 59) to a scaled score (10) using Table 2, which allows us to calculate the difference between the predicted and actual scores. The z-score can be calculated by dividing the difference between the actual scaled score and the predicted scaled score by the relative standard error (RSE) of the regression model (Table 3).

In our example:

$$\text{z-score} = (10 - 10.73) / 2.56 = -0.28$$

Regression analyses (Table 3) indicated that age_c was a significant predictor of performance for all tests in HS except PASAT and COWAT. Sex was a significant predictor of performance in FLT, SDMT, PASAT, JLOT and ECCC. Finally, educational level was a significant predictor of performance in all tests except BVMT-R.

The proportion of PwMS with impairments in each test applying these norms and effect sizes are summarised in Table 4. Only 101 PwMS performed all tests of the battery. Compared to HS, PwMS had lower z-scores on the FLT-TL, SDMT, BVMT-R-TL, FLT-DR and BVMT-R-DR (all, $p < 0.001$), COWAT ($p < 0.01$) and PASAT-3 ($p < 0.05$). There were no differences in JLOT and ECCC.

Performance on each measure of the MACFIMS was classified as impaired or intact based on the z-score. The MACFIMS battery was able to identify CI (at least two abnormal tests in different domains) in 32.7% of PwMS. A short version of the MACFIMS, including only tests significantly lower in PwMS than in HC (SDMT, memory tests, COWAT and PASAT, excluding ECCC and JLOT) identified 29.7% of patients with CI (three patients were not classified as CI in the short version but were classified CI by the whole battery).

The proportions of CI patients in all subtypes are shown in Figure 1. There were significant differences in the frequency of impairment on COWAT between RRMS and PPMS ($p < 0.05$). With a CI rate of 42.90%, patients with SPMS showed a significantly higher rate of CI on the MACFIMS than patients with RRMS ($p < 0.05$). CI was detected in 35.3% of patients with PPMS and 18.8% of patients with RRMS, but the difference was not significant ($p = 0.13$).

3.3 Test-retest reliability

The Pearson's correlation coefficients between the first and second tests in 147 HS and in PwMS, 1 month apart, are presented in Table 5.

3.4 Effects of confounding factors

Level of fatigue (EMIF-SEP), level of anxiety (STAI A and B), depression status (BDI-FS) and mood (EHD-PRO) were not significantly associated with any of the outcome measures in our study for PwMS (Pearson's correlation).

3.5 Vocational status and MACFIMS

Unemployment was present in 58.8% of CI patients and 53.7% of non-impaired (p=0.63). Impairment on the MACFIMS was not significantly associated with an increased risk of unemployment (OR= 1.2 (0.5–2.8); p= 0.63).

4. DISCUSSION

The main objectives of this study were to evaluate cognitive assessment using the MACFIMS in a French-speaking sample of PwMS and HS and establish its validity by determining whether it can discriminate between PwMS and HS. This is the first study to validate a French version of the MACFIMS. We aimed to establish formally the norms of the MACFIMS battery for the French-speaking population, utilising the regression-based norms adjusted for age, sex and education in a large sample of HS. The regression-based approach to norms development enables us to account for the influences of demographic factors on test performance and use the entire normative sample rather than divide it into smaller subgroups for the computation of age- or education-stratified means and standard deviations. The sample of PwMS was recruited from multiple centres.

The second objective was to study the test-retest reliability. Finally, the effects of confounding factors, such as depression, anxiety and fatigue, were examined.

Using these norms, the MACFIMS battery was able to identify CI (at least two abnormal tests in different domains) in 33.7% of PwMS. The proportion of CI varied according to the disease subtype. The most affected patients were those with SPMS (42.9%) followed by PPMS

(35.3%) and then RRMS (18.8%). These results were consistent with those in the literature, in different countries and languages (Ruet et al., 2012; Benedict et al., 2006; Dusankova et al., 2012; Eshagi et al., 2012 ; Argento et al., 2018 ; Migliore et al., 2017).

The domains with more frequent impairment were learning (FLT-TL, BVMT-R-TL) followed by IPS (SDMT), delayed memory (FLT-DR, BVMT-DR), verbal fluency and working memory (PASAT). We did not observe significant differences between PwMS and HS in the JLOT test, as reported in some other studies (Argento et al., 2018; Migliore et al., 2017). The mean scores of the ECCC were not different between PwMS and HS. The absence of significant results for these tests may have been due to a lack of power, small number of patients with SPMS in our population, clinical phenotype associated with more frequent EF involvement[1] or to low sensitivity of these tests.

However, unlike the original validation study (Benedict et al., 2006) in which significant differences were observed for all tests between PwMS and HS, other studies showed the absence of significant differences for some tests, e.g. in recent Italian studies, with regard to the SDMT (Argento et al., 2018; Migliore et al., 2017), CVLT TL (Argento et al., 2018), BVMT TL (Migliore et al., 2017) and the PASAT 3 (Migliore et al., 2017).. These discrepancies may be explained by differences between the groups in these studies. Indeed, in the first validation study (Benedict et al., 2006), a large proportion of the PwMS were referred due to suspicion of cognitive difficulties, unlike in the present study. Considering these results, we evaluated the sensitivity of a shorter form of the MACFIMS battery not including the JLOT and the ECCC. This short battery identified 29.7% of patients with CI. The usefulness of this shorter battery must be assessed in other samples in future studies. Recently, an abbreviated form of the MACFIMS has been proposed as a short assessment (Gromisch et al., 2018).

The effect size of the tests was assessed by the Cohen's D showing a larger effect for IPS

(SDMT) followed by episodic memory. This was consistent with the MS literature (Ruet et al., 2012; Langdon, 2011).

The validity of the battery was confirmed by the capacity of the MACFIMS to distinguish PwMS from HS.

Only the SDMT shows good test-retest reliability (>0.8), the test-retest reliability is acceptable for FLT and COWAT (>0.7). The BVMTR reliability is particularly poor in comparison with literature.It is a limitation of the study.. The use of alternative forms for the memory tests and the ECCC may explain the lower reliability observed for these tests. The results of the tests were unaffected by confounding factors, as determined using self-reported questionnaires concerning mood, depressive symptoms, anxiety and fatigue, although the scores for these scales were significantly different between PwMS and HS. Another limitation of the study is the lack of good external criterion for validity. Risk of unemployment was not associated with CI on the MACFIMS. Other standards to judge the adequacy of the testing could have been added to the study such as assessment of activities of daily living.

In conclusion, this study established the norms of the MACFIMS battery for the French-speaking population and provided evidence for the validity of the MACFIMS as a large battery for assessment of cognition in MS.

Acknowledgements:

Funding: The study was sponsored by the Bordeaux University Hospital CHU de Bordeaux. The study was funded by an unrestricted grant from Bayer Healthcare Pharma. The funder did not contribute to the study design, data collection, analysis or reporting of the study.

REFERENCES

1. Ruet A, Deloire M, Hamel D, Ouallet JC, Petry K, Brochet B. Cognitive impairment, health-related quality of life and vocational status at early stages of multiple sclerosis: a 7-year longitudinal study. J. Neurol. 2013; 260: 776–784.

2. Langdon, D. W. Cognition in multiple sclerosis. Curr. Opin. Neurol. 2011; 24: 244–249.

3. Deloire MS, Bonnet MC, Salort E, Arimone Y, Boudineau M, Petry KG, Brochet B. How to detect cognitive dysfunction at early stages of multiple sclerosis? Mult Scler. 2006; 12: 445-52.

4. Brochet, B. & Ruet, A. Cognitive Impairment in Multiple Sclerosis With Regards to Disease Duration and Clinical Phenotypes. Front Neurol 10, (2019). Doi: 10.3389/fneur.201900261. ecollection 2019.

5. Morrow SA, Drake A, Zivadinov R, Munschauer F, Weinstock-Guttman B, Benedict RH. Predicting loss of employment over three years in multiple sclerosis: clinically meaningful cognitive decline. Clin Neuropsychol. 2010; 24: 1131-45.

6. Deloire, M., Ruet, A., Hamel, D., Bonnet, M. & Brochet, B. Early cognitive impairment in multiple sclerosis predicts disability outcome several years later. Mult. Scler. 2010; 16: 581–587.

7. Rao SM. Cognitive Function Study Group, National Multiple Sclerosis Society. A manual for the Brief Repeatable Battery of neuropsychological tests in multiple sclerosis . National Multiple Sclerosis Society, 1990.

8. Benedict, R. H., Fischer JS, Archibald CJ, et al. Minimal neuropsychological assessment of MS patients: a consensus approach. Clin Neuropsychol. 2002; 16: 381–397.

9. Benedict, R. H., Cookfair D, Gavett R, et al. Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). J Int Neuropsychol Soc. 2006; 12: 549–558.

10. Dusankova, J. B., Kalincik, T., Havrdova, E. & Benedict, R. H. B. 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: 1186–1200.

11. Eshaghi A, Riyahi-Alam S, Roostaei T, et al., Validity and reliability of a Persian translation of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS). Clin Neuropsychol. 2012; 26: 975-84.

12. Argento, O. et al. The Italian validation of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS) and the application of the Cognitive Impairment Index scoring procedure in MS patients. Neurol. Sci. 2018 ; 39, 1237–1244.

13. Maubeuge N., Deloire, M., Brochet B., t :. French Validation of the Brief International Cognitive Assessment for Multiple Sclerosis. Rev. Neurol. (Paris).2020 S003537872030610X doi:10.1016/j.neurol.2020.04.028. online ahead of print.

14. Polman CH, Reingold SC, Banwell B., et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol. 2011; 69: 292-302.

15. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology. 2014; 83: 278-86.

16. Smith, A. Symbol digit modalities test: Manual. Los Angeles: Western Psychological Services (1982).

17. Gronwall, D. M. Paced auditory serial-addition task: a measure of recovery from concussion. Percept Mot Skills 1977 ; 44 : 367–373.

18. Poitrenaud, J., Deweer, B., Kalafat, M. & Van der linden, M. CVLT - Test d'apprentissage et de mémoire verbale (adaptation française). (2007).

19. Benton, A. L. Sivan A.B., Hamsher K.D., Varnet N.R., Spreen O.. Contributions to Neuropsychological Assessment: A Clinical Manual. (Oxford University Press, 1994).

20. Delis , D.C. , Kaplan , E. , & Kramer , J.H. Delis-Kaplan Executive Function System . San Antonio, Texas : Psychological Corporation. (2001).

21. Benedict, RHB. The Brief Visuospatial Memory Test Revised (BVMT-R). (1997).

22. Benton, A. L., Hamsher, de S. K., & Sivan, A. B. Multilingual aplasia examination (2nd ed.). Iowa City, IA: AJA Associates]. (1983).

23. Brochet B. Assessing incapacity at early stages of multiple sclerosis using the EDSS. Rev Neurol (Paris). 2009; 165(Suppl 4): S173-S179.

24. Benedict RH, Fishman I, McClellan MM, Bakshi R, Weinstock-Guttman B. Validity of the Beck Depression Inventory-Fast Screen in multiple sclerosis. Mult Scler. 2003; 9: 393–6.

25. Spielberger, C. D., Bruchon-Schweitzer, M. & Paulhan, I. Inventaire d'anxiété état-trait forme Y (STAI-Y): [manuel. (Ed. du Centre de psychologie appliquée). (1993).

26. Lamargue Hamel D, Deloire M, Ruet A, Charré-Morin J, Saubusse A, Ouallet J-C, et al.(2015) Deciphering Depressive Mood in Relapsing-Remitting and Progressive Multiple Sclerosis and Its Consequence on Quality of Life. PLoS ONE 2015 10(11):e0142152. doi:10.1371/journal.pone.0142152

27. Debouverie, M., Pittion-Vouyovitch, S., Louis, S. & Guillemin, F. Validity of a French version of the fatigue impact scale in multiple sclerosis. Multiple sclerosis. 2007; 13: 1026–32.

28. Parmenter, B. A., Testa, S. M., Schretlen, D. J., Weinstock-Guttman, B. & Benedict, R. H. B. The utility of regression-based norms in interpreting the minimal assessment of

cognitive function in multiple sclerosis (MACFIMS). J Int Neuropsychol Soc. 2010; 16: 6–16.

29. Berrigan, L. I. et al. Reliability of regression-based normative data for the oral symbol digit modalities test: an evaluation of demographic influences, construct validity, and impairment classification rates in multiple sclerosis samples. Clin Neuropsychol. 2014 ; 28 : 281–299.

30. Ivnik R J, Malec J F, Smith G E et al. Mayo's Older Americans Normative Studies: WAIS—R norms for ages 56 to 97. Clinical Neuropsychologist. 1992; 6(Suppl): 1–30.

31. Migliore S, Ghazaryan A, Simonelli I, et al. Cognitive Impairment in Relapsing-Remitting Multiple Sclerosis Patients with Very Mild Clinical Disability. Behav Neurol. 2017-7404289.

32. E.S. Gromisch, J.G. Portnoy, F.W. Foley, Comparison of the abbreviated minimal assessment of cognitive function in multiple sclerosis (aMACFIMS) and the brief international cognitive assessment for multiple sclerosis (BICAMS), J. Neurol. Sci. 2018 ; 388 : 70–75.

Legend of Figure 1: Proportion of PwMS with cognitive impairment

*$P < 0.05$, compared to RRMS (χ^2 test); **$P < 0.01$, compared to RRMS (Fisher's exact test).

Table 1: Clinical and demographic characteristics of PwMS

	Patients (*n* = 134)	**HS (*n* = 276)**	***p***
Age	49.62 (9.34)	43.84 (12.42)	< 0.001
Sex			ns
Female (%)	85 (63.4%)	158 (57.3%)	
Education			ns
LEL (%)	60 (44.8%)	114 (41.3%)	
HEL-BAC (%)	23 (17.2%)	22 (8.0%)	
HEL-BAC+2-4 (%)	42 (31.3%)	105 (38.0%)	
HEL-BAC+5 (%)	9 (6.7%)	35 (12.7%)	
Disease duration (years)	15.23 (9.62)		
EDSS (median)	4.0 [0 – 8]		
Disease subtype			
RR (*n*)	43		
SP (*n*)	46		
PP (*n*)	45		
BDI-FS	3.64 (3.2)	1.47 (2.1)	< 0.001
EHD-PRO	20.00 (5.25)	16.03 (3.92)	< 0.001
STAI A	39.35 (11.74)	29.07 (8.13)	< 0.001
STAI B	42.44 (9.03)	34.53 (9.56)	< 0.001
EMIF-SEP-physical	23.56 (8.71)	9.09 (8.42)	< 0.001
EMIF-SEP cognitive	15.85 (9.57)	9.28 (8.43)	< 0.001
EMIF-SEP-social	3.85 (2.57)	2 (2.18)	< 0.001
Employment (%)	60 (44.8%)		

Results are expressed as means unless otherwise specified.

BDI-FS, Beck Depression Index-Fast Screen; EDSS, Expanded Disability Status Scale; EHD-PRO, Echelle d'Humeur Dépressive-patient reported outcome; EMIF-SEP, French adaptation of the modified fatigue Impact Scale; HEL, high education level (see Methods); HS, healthy subjects; LEL, low education level; not significant; ns, PP, primary progressive; RR, relapsing remitting; SD, standard deviation; SP, secondary progressive; STAI, State-Trait Anxiety Inventory (A = trait, B = state).

Table 2: Raw score to scaled score conversions

	Raw score									
Scaled Score	**SDMT**	**FLT**		**BVMT-R**		**PASAT-3**	**JLOT**	**COWAT**	**ECCC**	
		Total learning	**Delayed Recall**	**Total learning**	**Delayed Recall**				**Correct sorting**	**Description**
2	< 35	< 35	< 3	< 9	< 3	< 16	< 13	< 17	< 4	< 4
3	35–36	35–37	3–5	9–10	3	16–19	13–14	17–18	4	4
4	37–38	37–40	6	11–12	4	20	15	19–20		
5	39–44	41–43	7	13–16	5–6	21–23	16–17	21–24	5	5
6	45–48	44–45	8–9	17–19	7	24–29	18–19	25–26	6	6
7	49–50	46–48	10	20–21	8	30–34	20–21	27–30	7	7
8	51–53	49–52	11	22–23	9	35–40	22	31–33		8
9	54–56	53–55	12	24–26	10	41–45	23–24	34–35	8	9
10	57–60	56–61	13	27–28	11	46–49	25–26	36–39	9	10
11	61–63	62–63	14	29–30		50–53	27	40–42	10	11
12	64–65	64–65	15	31		54	28	43–45	11	12
13	66–68	66–68		32–33	12	55–56	29	46–48	12	
14	69–71	69–70	16	34		57–58		49–52		13
15	72–73	71		35		59	30	53–55	13	14
16	74–75	72						56–59		
17	76–77	73		36		60		60		15
18	> 77	> 73				> 60		> 60	> 14	> 15

BVMT-R, Brief Visuospatial Memory Test-revised; COWAT, Controlled Oral Word Association Test; ECC, épreuve de classement de cartes de Champagne; FLT, French

Learning Test; JLOT, Judgment of Line Orientation Test; PASAT-3, Paced Auditory Serial Addition Test with 3.0-s interstimulus interval; SDMT, Symbol Digit Modalities Test.

Table 3: Final regression models for MACFIMS measures

Measure	Predictor	B	Standard error B	Standardised B	t	*p*	Total R square	RMSE
SDMT	(constant)	8.91	0.603		14.761	< 0.001		
	age_c	−0.09	0.013	−0.397	−7.169	< 0.001		
	age_c^2	−0.002	0.001	−0.097	−1.746	0.082		
	Sex	0.65	0.317	0.113	2.049	0.041		
	Education	0.36	0.14	0.143	2.586	0.01	**0.185**	2.56
FLT-TL	(constant)	7.52	0.608		12.354	< 0.001		
	age_c	−0.05	0.013	−0.219	−3.932	< 0.001		
	age_c^2	−0.003	0.001	−0.132	−2.348	0.02		
	Sex	1.33	0.319	0.231	4.151	< 0.001		
	Education	0.68	0.141	0.27	4.832	< 0.001	**0.173**	2.58
BVMT-R-TL	(constant)	10.26	0.617		16.637	< 0.001		
	age_c	−0.08	0.013	−0.333	−5.798	< 0.001		
	age_c^2	−0.002	0.001	−0.09	−1.567	0.118		
	Sex	−0.17	0.324	−0.03	−0.517	0.605		
	Education	0.25	0.143	0.098	1.712	0.088	**0.122**	2.62
FLT-DR	(constant)	7.57	0.604		12.546	< 0.001		
	age_c	−0.04	0.013	−0.194	−3.357	0.001		
	age_c^2	−0.001	0.001	−0.033	−0.566	0.572		
	Sex	1.24	0.317	0.227	3.921	< 0.001		
	Education	0.46	0.14	0.191	3.291	0.001	**0.113**	2.56
BVMT-R-DR	(constant)	10.48	0.56		18.694	< 0.001		
	age_c	−0.06	0.012	−0.28	−4.816	< 0.001		
	age_c^2	−0.002	0.001	−0.125	−2.139	0.033		

	Sex	−0.29	0.294	−0.057	−0.978	0.329		
	Education	0.20	0.13	0.09	1.54	0.125	**0.098**	2.38
PASAT-3	(constant)	11.28	0.624		18.067	< 0.001		
	age_c	−0.001	0.013	−0.003	−0.06	0.952		
	age_c^2	−0.002	0.001	−0.081	−1.405	0.161		
	Sex	−1.15	0.327	−0.203	−3.527	< 0.001		
	Education	0.66	0.145	0.263	4.562	< 0.001	**0.131**	2.62
JLOT	(constant)	11.67	0.577		20.238	< 0.001		
	age_c	−0.04	0.012	−0.181	−3.261	0.001		
	age_c^2	0	0.001	−0.019	−0.332	0.74		
	Sex	−1.54	0.303	−0.283	−5.081	< 0.001		
	Education	0.53	0.134	0.221	3.965	< 0.001	**0.164**	2.45
COWAT	(constant)	8.73	0.646		13.51	< 0.001		
	age_c	0.02	0.014	0.09	1.527	0.128		
	age_c^2	0.003	0.001	0.149	2.525	0.012		
	Sex	0.06	0.339	0.01	0.163	0.87		
	Education	0.62	0.15	0.244	4.143	< 0.001	**0.081**	2.74
ECCC (CS)	(constant)	9.99	0.54		18.478	< 0.001		
	age_c	−0.06	0.011	−0.258	−5.065	< 0.001		
	age_c^2	−0.001	0.001	−0.075	−1.46	0.145		
	Sex	−0.72	0.283	−0.13	−2.547	0.011		
	Education	1.15	0.125	0.468	9.166	< 0.001	**0.316**	2.28
ECCC (D)	(constant)	9.99	0.54		18.363	< 0.001		
	age_c	−0.06	0.011	−0.297	−5.639	< 0.001		
	age_c^2	−0.001	0.001	−0.057	−1.07	0.286	0.266	
	Sex	−0.61	0.285	−0.113	−2.145	0.033		2.29

	Education	0.95	0.126	0.397	7.51	< 0.001		

BVMT-R, Brief Visuospatial Memory Test-Revised; COWAT, Controlled Oral Word Association Test; CS, correct sorts; D, description score; DR, Delayed recall; ECC, épreuve de classement de cartes de Champagne; FLT, French Learning Test; JLOT, Judgment of Line Orientation Test; PASAT-3, Paced Auditory Serial Addition Test with 3.0-s interstimulus interval; SDMT, Symbol Digit Modalities Test; TL, total learning.

Table 4: Comparison between PwMS and HS in each metric using individual z-scores derived from regression-based model

	PwMS			**HS (*n* = 276)**		***P*(1)**	**Effect size**
	n	z-score	Impaired(2)	z-score	Impaired(2)		
SDMT	125	−0.77 (1.20)	28.5%	0.01 (1.00)	6.9%	$p < 0.001$	0.73
FLT (TL)	134	−1.14 (1.44)	38.1%	−0.33 (1.05)	13.4%	$p < 0.001$	0.68
FLT (DR)	134	−0.58 (1.15)	21.6%	0.02 (1.00)	6.9%	$p < 0.001$	0.57
BVMT-R (TL)	134	−0.68 (1.18)	29.1%	−0.22 (1.02)	12.3%	$p < 0.001$	0.43
BVMT-R (DR)	133	−0.53 (1.15)	19.5%	−0.01 (1.00)	7.2%	$p < 0.001$	0.49
PASAT-3	119	−0.22 (1.15)	10.9%	0.03 (1.00)	6.6%	$p < 0.05$	0.24
JLOT	134	−0.07 (1.00)	9.7%	−0.04 (1.00)	6.9%	$p = 0.770$	
ECCC (CS)	125	−0.05 (1.34)	12.0%	−0.03 (1.00)	10.6%	$p = 0.858$	
ECCC (D)	121	−0.13 (1.21)	9.9%	−0.001(1.00)	5.5%	$p = 0.299$	
COWAT	134	−0.40 (1.32)	17.2%	−0.01 (1.00)	5.8%	$p < 0.01$	0.40

(1): p values between PwMS and HS. (2): Subjects were considered impaired if z-score ≤ −1; BVMT-R, Brief Visuospatial Memory Test-revised; COWAT, Controlled Oral Word Association Test; CS, correct sorts; D, description score; DR, Delayed recall; ECC, épreuve de classement de cartes de Champagne. TL, total learning; FLT, French Learning Test; JLOT, Judgment of Line Orientation Test; PASAT-3, Paced Auditory Serial Addition Test with 3.0-s interstimulus interval; SDMT, Symbol Digit Modalities Test.

Table 5: Correlation between the two sessions

	SDMT	FLT/CVLT-TL	FLT/CVLT DR	BVMT-R IR	BVMT-R DR	PASAT	JLOT	ECCC	COWAT
Correlation coefficient in HS	0.846	0.670	0.730	0.518	0.583	0.688	0.574	0.490	0.728
Correlation coefficient in PwMS	0.872	0.790	0.762	0.574	0.560	0.599	0.522	0.685	0.794

All correlations were significant ($p < 0.001$). BVMT-R, Brief Visuospatial Memory Test-revised; COWAT, Controlled Oral Word Association Test; D, description score; CS, correct sorts; DR, Delayed recall; ECC, épreuve de classement de cartes de Champagne. TL, total learning; FLT, French Learning Test; JLOT, Judgment of Line Orientation Test; PASAT-3, Paced Auditory Serial Addition Test with 3.0-s interstimulus interval; SDMT, Symbol Digit Modalities Test.

Figure 1: Proportion of PwMS with cognitive impairment