

Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS) in the Italian population

Simone Migliore^{1,2} · Anna Ghazaryan³ · Iaria Simonelli^{3,4} · Patrizio Pasqualetti^{3,4} ·
Doriana Landi⁵ · Maria Giuseppina Palmieri⁵ · Filomena Moffa³ · Pasquale Rinaldi⁶ ·
Fabrizio Vernieri⁷ · Maria Maddalena Filippi³

Received: 20 November 2015 / Accepted: 7 April 2016 / Published online: 20 April 2016
© Springer-Verlag Italia 2016

Abstract Cognitive dysfunction involves 40–65 % of multiple sclerosis (MS) patients. It can be detected in all MS phenotypes from the early stages of the disease, and it tends to progress over time. Minimal Assessment of Cognitive Function in MS (MACFIMS) has been proved to be the most sensitive and comprehensive battery available for MS cognitive assessment in the English population. In Italy, MACFIMS applicability is limited in everyday clinical practice since the overall validity of this battery in the Italian MS population has never been demonstrated. The aim of this study was to translate/cross-culturally adapt and validate an Italian version of the MACFIMS. A total of 130 MS patients and 60 healthy controls (HCs) were enrolled and evaluated with an Italian version of the MACFIMS. All tests discriminated MS patients from HCs;

according to the literature, approximately more than half of MS patients (70.8 %) exhibit cognitive impairment. Principal component analysis showed four distinct components: visual–spatial memory/processing speed, working memory, executive functions and verbal memory. Our study is the first to validate an Italian version of the MACFIMS. Several aspects of validity have been demonstrated: criterion and, partially, construct. Future work will investigate the longitudinal course of neuropsychological dysfunction in Italian MS patients using these measures.

Keywords Multiple sclerosis · Processing speed · Memory · Neuropsychological testing · Validity

✉ Simone Migliore
s.migliore@unicampus.it

Anna Ghazaryan
a.ghazaryan@unicampus.it

Iaria Simonelli
ilaria.simonelli@afar.it

Patrizio Pasqualetti
patrizio.pasqualetti@afar.it

Doriana Landi
doriana.landi@gmail.com

Maria Giuseppina Palmieri
mgpalmieri@inwind.it

Filomena Moffa
filomena.moffa@fbf-isola.it

Pasquale Rinaldi
pasquale.rinaldi@istc.cnr.it

Fabrizio Vernieri
f.vernieri@unicampus.it

Maria Maddalena Filippi
mariamaddalena.filippi@fbf-isola.it

¹ Clinical Psychology Service, University Campus Bio-Medico, Via Alvaro de Portillo 200, Rome, RM, Italy

² LIRH Foundation, Via dei Mille 41, Rome, Italy

³ Department of Neuroscience, Fatebenefratelli Hospital-Isola Tiberina, Rome, Italy

⁴ Neurodegeneration Lab and Department of Imaging, IRCCS San Raffaele Pisana, Rome, Italy

⁵ Department of Neuroscience, Policlinico “Tor Vergata”, Rome, Italy

⁶ Institute of Cognitive Sciences and Technologies, National Research Council of Italy, Rome, Italy

⁷ Neurology Unit, University Campus Bio-Medico of Rome, Rome, Italy

Introduction

Since the early description of the disease, cognitive disorders have been known to be a common feature of multiple sclerosis (MS); the precise nature of these deficits, their frequency and their evolution throughout the disease process are yet to be completely established. Recent evidence suggests that cognitive impairment (CImp) may be detectable at any stage of MS, with an overall prevalence ranging from 40 to 65 % [1] and even after a single demyelinating episode [2, 3]. Prevalence of cognitive disturbance varies among MS subtypes, and it can be the main or only clinical symptom [4, 5]. Studies comparing selected samples of patients with relapsing–remitting MS and progressive MS (PMS—secondary and primary) demonstrated that CImp is more frequent and severe in progressive than in the relapsing–remitting group [6, 7].

Attention, visuospatial abilities, learning and memory, information processing speed and executive functions [1, 8, 9] are the most involved domains. Processing speed and visual learning and memory seem to be most commonly affected in MS [1, 10]. Areas of cognition that are not usually affected are “simple” attention (e.g. repeating digits) and essential verbal skills (e.g. word naming and comprehension) [1, 11].

To evaluate neuropsychological deficits, several tools are available. Two neuropsychological batteries are mostly used the Rao Brief Repeatable Neuropsychological Battery (BRBN) [12] and the Minimal Assessment of Cognitive Function in MS (MACFIMS) [10, 13]. MACFIMS was developed by an expert panel of neuropsychologists convened by the Consortium of MS centre in April 2001 [13]. It requires approximately 90 min and consists of seven tests: Controlled Oral Word Association Test (COWAT; [14]), Judgement of Line Orientation Test (JLO; [14]), California Verbal Learning Test II edition (CVLT-II; [15]), Brief Visuospatial Memory Test—Revised (BVRT-R; [16]), Paced Auditory Serial Addition Test (PASAT; [17]), Symbol Digit Modalities Test (SDMT; [18]), Delis Kaplan Executive Function System Sorting Test (DKEFS sorting test; [19]). These tests have good reliability [20] as well as validity, in that they generally correlate well with brain MRI metrics [21, 22] and employment status [10]. A recent study [23] comparing BRBN and MACFIMS batteries found that they are comparable in their discriminative validity. However, the BVRT-R appears to be more sensitive than the 10/36 Spatial Recall Test, used for the assessment of visual memory; moreover, the MACFIMS battery investigates executive functions that the BRBN does not explore. However, further studies are necessary to prove and enhance these results.

Despite well-documented psychometrics in English, the MACFIMS is not very easily applied in non-English-speaking countries. Currently, MACFIMS is not applicable to the Italian population both for research purposes and for everyday clinical practice since the validity of this battery in Italy has never been demonstrated.

The aim of this study was to explore the validity of the MACFIMS in the Italian population; several aspects of validity were assessed as follows: criterion (ability to discriminate between MS patients and healthy controls) and construct (factorial structure of the MACFIMS).

Materials and methods

Subjects

One hundred and thirty patients with clinically diagnosed MS (according to McDonald’s criteria; [24]) and 60 healthy controls (HCs) matched by sex, age and education were enrolled. The demographic and clinical characteristics of the MS and HC groups are reported in Table 1.

All subjects were recruited at the MS centre of the Department of Neuroscience of San Giovanni Calibita “Fatebenefratelli” Hospital (Rome), Policlinico “Tor Vergata” (Rome) and Campus Bio-Medico University (Rome). The research participants (including all HCs) were contacted by mail/telephone or were approached during the course of their usual clinical care at the MS centre. All patients referred to MS centres from January 2012 to June 2013 and meeting the inclusion/exclusion criteria were enrolled. About 35 % of them were assessed for routine monitoring of cognitive functioning, 30 % for specific clinical reasons (i.e. disability evaluation, differential diagnosis of CImp vs depressive disorder, suspected cognitive impairment), and about 35 % were research volunteers. HCs were selected from hospital employees (physicians, nurses, clerks, cleaners, porters) and their relatives.

All subjects met the following inclusion criteria: age 18 years or older; fluent in Italian; and able to provide informed consent to all procedures. Exclusion criteria were neurological disorder other than MS; psychiatric disorder other than mood, personality or behaviour change following the onset of MS; medical condition that might influence cognition; history of developmental disorder (e.g. ADHD, learning disability); history of substance or alcohol dependence, or current abuse; motor or sensory defect that might interfere with cognitive test performance; relapse and/or corticosteroid use within 4 weeks of assessment (only patients). Patients were under routine treatment, and we did not include or exclude patients based on medications.

Table 1 Demographics and clinical characteristics of study samples

	MS patients		Total patients (<i>n</i> = 130)	Healthy controls (<i>n</i> = 60)	<i>p</i> *
	RR (<i>n</i> = 100) + CIS (<i>n</i> = 5)	SP (<i>n</i> = 16) + PP (<i>n</i> = 9)			
Age, years (mean, \pm SD)	42 (\pm 10.7)	55.9 (\pm 7.5)	44.7 (\pm 11.5)	44.3 (\pm 10.6)	0.834
Gender (F, %)	76 (72.4 %)	14 (56 %)	90 (69.2 %)	41 (68.3 %)	0.901
Years of education, years (median—min/max)	13 (5–18)	13 (5–18)	13 (5–18)	13 (8–18)	0.674
Disease duration, years (median—min/max)	9.9 (0.3–30.1)	18.2 (1.2–46.2)	11.8 (0.3–46.2)		
EDSS (median—min/max)	1 (0–7.5)	6 (1.5–8)	1.5 (0–8)		

Age: ANOVA (HC vs CIS/RR vs. SP/PP), $p < 0.001$

Post hoc comparison. HC vs CIS/RR, $p = 0.513$; HC vs SP/PP, $p < 0.001$; CIS/RR vs SP/PP, $p < 0.001$

MS multiple sclerosis, RR relapsing–remitting, CIS clinical isolated syndrome, SP secondary progressive, PP primary progressive

* p value is referring to t test or χ^2 test comparing total patients to healthy controls

In data analysis, we considered together relapsing–remitting and clinically isolated syndrome patients (Relapsing group) and secondary progressive and primary progressive patients (Progressive group).

A detailed clinical interview was performed to verify the inclusion and exclusion criteria; all patients underwent a complete neurological examination. Each HC and MS patients was asked to sign an informed written consent (previously approved by the local ethical committee) to participate in the study.

Neuropsychological assessment

Tests were administered in a standardized manner, during daytime, in a quiet room, and in a fixed order, in accordance with consensus panel recommendations [10, 13]. The MACFIMS includes seven tests. The COWAT is a measure of verbal fluency and mental flexibility; it was administered in the standard manner, following the method of Arthur Benton [14]. In successive 1-min trials, participants generated as many words as possible, beginning with each of three designated letters (F-A-S). The dependent measure was the total number of correct words over the three trials. The JLO is a measure of the visual–spatial ability; it required participants to identify the angle defined by two stimulus lines from among those defined by a visual array of lines covering 180°. The dependent variable was the total number of correct responses over 30 items. The CVLT-II is a measure of verbal learning and memory; the ability to learn a 16-word list was first examined over the course of five trials. The examiner reads 16 words and asked participants to repeat as many words as possible. The entire List A was repeated each time. After 25 min, participants are asked to recall the list again without another exposure. Outcome measures are total learning over five trials and number of correct recalls following the delay.

The BVMT-R is a measure of visuospatial learning and memory; it consists of three learning trials and delayed recall of a matrix of six designs. Each design received a score of 0, 1 or 2 based on accuracy and location scoring criteria. There were three free-recall trials followed by a 25-min delayed recall. The PASAT is a test of auditory working memory and mental speed and included 60 trials presented at inter-stimulus intervals of three (PASAT-3) and two (PASAT-2) seconds. The dependent measure was the number of correct responses from each of the two trials. The SDMT is a test of visual processing speed consisting of nine abstract symbols paired with a number from one to nine. Participants responded by voicing the digit associated with each symbol as quickly as possible. The dependent measure was the number of correct responses in 90 s. The DKEFS sorting test was employed to assess higher executive function. The examinee is asked to sort six different cards into two groups in as many different ways as possible. With each sort, the rationale for sorting is explicated by the patient and scored for accuracy. We recorded the total number of correct sorts (CS) and calculated the verbal description score (DS), which was based on the abstractness and accuracy of the sort descriptions.

Moreover, the participants completed the Beck Depression Inventory (BDI; [25]) to evaluate depression. The entire test battery required 90/100 min of face-to-face testing time.

Every neuropsychological test is considered impaired if the score is \leq 5th percentile with respect to normative data; patients impaired on two or more tests were defined as cognitively impaired. Italian normative data of COWAT [26], JLO [27], PASAT [28], SDMT [29] and the DKEFS sorting test [30] are available; for normative data of CVLT-II and BVMT-R, we considered the US sample [15, 16].

On the basis of the number of test failures, patients were classified as mildly (two test failures), moderately (three

test failures) or severely (four or more test failures) impaired, following the paper of Nocentini et al. [31]. This classification reflects different levels of cognitive deterioration in order to highlight different degrees of severity of cognitive dysfunction.

In addition to the number of tests failed, the number of cognitive domains impaired was considered. More specifically, the domain was considered damaged when at least one test in the domain was impaired. MS patients were considered as multi-domain cognitive impaired (mDCI) when at least two domains were found altered. We take into consideration five cognitive domains: verbal memory [CVLT total learning (TL); CVLT-delay recall (DR)], visual memory (BVMT-total learning (TL); BVMT-delay recall (DR)), information processing speed (PASAT-3; PASAT-2; SDMT), executive functions (DKEFS-CS; DKEFS-DS; COWAT) and visuospatial perception (JLO).

Translation and cross-cultural adaptation

Italian versions of COWAT [26], JLO [27], PASAT [28], SDMT [29] and DKEFS sorting test [30] were available.

To preserve semantic equivalence in the remaining test (BVMT-R, CVLT-II), first bilingual translators with a good understanding of Italian and English translated the tests into Italian. Next, back translation into English was performed by independent translators with a good understanding of Italian and English. To assure their equivalence, original and back-translated versions were compared.

BVMT-R is a visual–spatial test, thus only the instructions needed translation into Italian. Moreover, there are no semantic associations with stimuli in the Italian culture or language. For CVLT-II, an iterative process of modification, including cultural adaptation of some words, was undertaken, as these tests emphasize verbal stimuli. In CVLT-II, each word of the English version of List A has been replaced by an Italian word, so as to preserve the test's properties, match word frequency and appropriate similarity of meaning.

For every test, care was devoted to keep the original items in every case in which it was possible, given the resemblance between the two languages.

Statistical analysis

Data were presented as mean (SD) or, if necessary, as median (min–max). A log-transformation was applied to some variables to gain a better fit to Gaussianity, to limit the dangerous effect of extreme values and to reduce heteroscedasticity in the residuals. Group differences on demographic and clinical data were assessed using parametric tests (Student's *t* test or univariate ANOVA) or, if necessary, nonparametric tests (Mann–Whitney *U* test or

Kruskal–Wallis test, chi-square test). The correlations among all neuropsychological subtests and demographic and clinical data [BDI log scale (log-BDI)] were evaluated using Pearson's or Spearman correlation coefficients. Analyses of covariance (ANCOVA) were performed to evaluate the differences in test performance among HC and diagnosis groups (Relapsing and Progressive group) adjusting for sex, age and log-BDI. To describe the effect size, Cohen's *d* was calculated; it is the difference between means divided by the pooled SD, and its magnitude is assessed using the thresholds provided in Cohen [32].

Because data about four tests (BVMT-TL and BVMT-DR; PASAT-3; and PASAT-2) presented a right-skewed distribution inflated at their minimum value, adequate statistical models were used in addition to the linear general models previously applied. Specifically, zero-truncated negative binomial (ZTNB) regression was used to model BVMT-TL and BVMT-DR data (more precisely, the value of 0 could not occur and their distributions were inflated on their minimum value equals to 20).

Instead, PASAT-3 and PASAT-2 data were semi-continuous presenting a continuous distribution except for a probability mass at 0. A two-part model was applied, so as to separate the modelling into two stages. In the first stage, a binary model was used for the dichotomous event of having 0 or positive values, such as the logistic regression model; conditional on a positive value, in the second part, data were modelled using a log-gamma generalized linear model. In the second part of the model, sex, age and log-BDI were considered as covariates.

A logistic regression model was applied to analyse the association between CImp (two or more tests below the cut-off) and diagnosis groups (HC, Relapsing and Progressive group), adjusting for sex, age and log-BDI.

Construct validity analysis was performed using principal component analysis (PCA) with varimax rotation. All neuropsychological (NP) tests were included. As a criterion to choose the number of factors, we considered that at least 80 % of the total variance should be explained. We considered the NP tests with a loading ≥ 0.5 for that variable most representative of a specific component. Overall, a *p* value of less than 0.05 was considered significant and a Bonferroni adjustment was applied when performing multiple comparisons. All statistical analyses were performed using SPSS 16 except for the ZTNB which was performed using STATA 10.

Results

Overall, MS patients and HCs were matched by age, sex and years of education. Considering relapsing and progressive groups, a significant difference was observed; in

Table 2 Neuropsychological tests scores comparing MS patients and HC

MACFIMS test	HC <i>n</i> = 60	MS patients <i>n</i> = 130	ANOVA [†] (HC vs MS) <i>p</i>	ldl Cohen ^{†§}	Pts group 1 (CIS/RR) <i>n</i> = 105	Pts group 2 (SP/PP) <i>n</i> = 25	ANOVA [†] (HC vs group 1 vs group 2) <i>p</i>	<i>p</i> adjusted*
CVLT-TL	53.6 (7.3)	46.8 (11.1)	<i>p</i> < 0.001	0.76	49.4 (10)	36.2 (9.1)	<i>p</i> < 0.001	<i>p</i> < 0.001
CVLT-DR	12.2 (2.5)	9.9 (3.5)	<i>p</i> < 0.001	0.79	10.5 (3)	7 (4.1)	<i>p</i> < 0.001	<i>p</i> < 0.001
BVMT-TL	26 (10–36)	23 (2–36)	<i>p</i> < 0.001 [‡]	0.63	24 (2–36)	12 (3–26)	<i>p</i> < 0.001 [‡]	<i>p</i> < 0.001 [‡]
BVMT-DR	11 (6–12)	10 (0–12)	<i>p</i> < 0.001 [‡]	0.66	10 (0–12)	6 (0–12)	<i>p</i> < 0.001 [‡]	<i>p</i> < 0.001 [‡]
PASAT 3	45 (23–58)	37.5 (0–60)	<i>p</i> = 0.047°	0.33	41 (0–60)	20 (0–58)	<i>p</i> = 0.034°	<i>p</i> = 0.091°
PASAT 2	34.5 (0–55)	26 (0–59)	<i>p</i> = 0.005°	0.47	29 (0–59)	22 (0–42)	<i>p</i> = 0.015°	<i>p</i> = 0.018°
SDMT	49.4 (10.0)	43.4 (14.8)	<i>p</i> = 0.004	0.45	46 (13.8)	32.2 (13.8)	<i>p</i> = 0.001	<i>p</i> = 0.011
DKEFS-CS	9.6 (2.1)	7.2 (2.2)	<i>p</i> < 0.001	1.15	7.4 (2.1)	6 (2.3)	<i>p</i> < 0.001	<i>p</i> < 0.001
DKEFS-DS	37.1 (8.7)	26.4 (9.9)	<i>p</i> < 0.001	1.18	27.8 (9.3)	20.6 (10.6)	<i>p</i> < 0.001	<i>p</i> < 0.001
JLO	24.9 (4.4)	22.4 (5.1)	<i>p</i> < 0.001	0.52	22.9 (4.7)	20.2 (5.8)	<i>p</i> < 0.001	<i>p</i> = 0.007
COWAT	41.4 (13.2)	32.8 (12.1)	<i>p</i> = 0.001	0.69	34 (12)	27.4 (11.1)	<i>p</i> < 0.001	<i>p</i> = 0.002
BDI	6.5 (0–19)	12 (0–42)	<i>p</i> = 0.003	0.63	10 (0–37)	14.5 (0–42)	<i>p</i> = 0.004	
MACFIMS test	Multiple comparisons ^{†#}							
	HC – (CIS+RR)				(CIS+RR) – (SP+PP)			
	ldl Cohen ^{†§*}	<i>p</i>	<i>p</i> adjusted*	ldl Cohen ^{†§**}	<i>p</i>	<i>p</i> adjusted*	ldl Cohen ^{†§**}	<i>p</i> adjusted*
CVLT-TL	0.65	<i>p</i> = 0.001	<i>p</i> < 0.001	1.37	<i>p</i> < 0.001	<i>p</i> < 0.001	0.71	<i>p</i> < 0.001
CVLT-DR	0.73	<i>p</i> < 0.001	<i>p</i> < 0.001	1.02	<i>p</i> < 0.001	<i>p</i> < 0.001	0.28	<i>p</i> = 0.060
BVMT-TL	0.51	<i>p</i> = 0.012 [‡]	<i>p</i> = 0.001 [‡]	1.41	<i>p</i> < 0.001 [‡]	<i>p</i> < 0.001 [‡]	0.47	<i>p</i> < 0.001 [‡]
BVMT-DR	0.57	<i>p</i> = 0.006 [‡]	<i>p</i> = 0.006 [‡]	1.52	<i>p</i> < 0.001 [‡]	<i>p</i> < 0.001 [‡]	0.48	<i>p</i> < 0.001 [‡]
PASAT 3	0.34	<i>p</i> = 0.327°	<i>p</i> = 0.039°	0.32	<i>p</i> = 0.030°	<i>p</i> = 0.04°	0.05	<i>p</i> = 0.282°
PASAT 2	0.53	<i>p</i> = 0.042°	<i>p</i> = 0.030°	0.21	<i>p</i> = 0.024°	<i>p</i> = 0.034°	0.33	<i>p</i> = 0.768°
SDMT	0.39	<i>p</i> = 0.057	<i>p</i> = 0.04	0.69	<i>p</i> = 0.001	<i>p</i> = 0.025	0.29	<i>p</i> = 0.036
DKEFS-CS	1.02	<i>p</i> < 0.001	<i>p</i> < 0.001	1.25	<i>p</i> < 0.001	<i>p</i> < 0.001	0.23	<i>p</i> = 0.999
DKEFS-DS	0.99	<i>p</i> < 0.001	<i>p</i> < 0.001	1.23	<i>p</i> < 0.001	<i>p</i> < 0.001	0.24	<i>p</i> = 0.999
JLO	0.44	<i>p</i> = 0.031	<i>p</i> = 0.028	0.69	<i>p</i> < 0.001	<i>p</i> = 0.024	0.25	<i>p</i> = 0.037
COWAT	0.55	<i>p</i> < 0.001	<i>p</i> = 0.004	0.66	<i>p</i> < 0.001	<i>p</i> = 0.034	0.11	<i>p</i> = 0.536

Table 2 continued

MACFIMS test	Multiple comparisons ^{†‡}					
	HC – (CIS+RR)		HC-(SP+PP)		(CIS+RR) – (SP+PP)	
	IdI Cohen ^{§§*}	<i>p</i>	<i>p</i> adjusted*	IdI Cohen ^{§§*}	<i>p</i>	<i>p</i> adjusted*
BDI	0.55	<i>p</i> = 0.062		1.13	<i>p</i> = 0.005	<i>p</i> = 0.240

Data are presented as mean (SD) or if the variables are not normally distributed, Median (min–max)

MS multiple sclerosis, RR relapsing–remitting, CIS clinical isolated syndrome, SP secondary progressive, CVLT-TL California verbal learning test-total learning, CVLT-DR CVLT-delay recall, BVMT-TL Brief visuospatial memory test-total learning, BVMT-DR BVMT-delay recall, SDMT symbol digit modalities test, PASAT–3 and 2 paced auditory serial addition test 3 and 2 s, JLO Judgment of line orientation, DKEFS-CS Delis Kaplan executive function system–sorting test–correct sort, DKEFS-DS DKEFS-description score, COWAT controlled oral words association test

* Adjusted for gender age and log-BDI

§ The magnitude is assessed using the thresholds provided in (Cohen [32]), i.e. IdI = 0.2 “small”, IdI = 0.5 “medium”, IdI = 0.8 “large”

Bonferroni correction

† Statistical tests and Cohen’s *d* were applied considering the standard scores for each test and after log-transformation of BDI to obtain a better fit to Gaussianity and a reduction of heteroscedasticity

‡ The results were obtained applying ZTNB (zero-truncated negative binomial) model

° The results were obtained by the second step of the two-part model applied

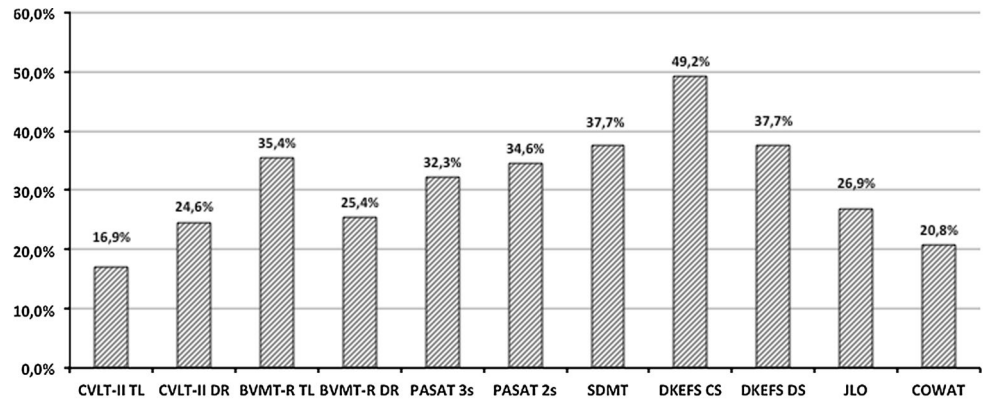
particular, the Progressive group seemed to be older than other groups (HC vs Progressive group, $p < 0.001$; Relapsing vs Progressive group, $p < 0.001$). For sex and years of education, no differences were observed (see Table 1). A significant difference was found in log-BDI scores comparing MS patients and HCs ($p = 0.003$). We also evaluated correlation between MACFIMS and depression (BDI score on a log scale). We observed only a slight significant correlation between PASAT-2 ($p = 0.045$), DKEFS-CS ($p = 0.006$), DKEFS-DS ($p = 0.015$), JLO ($p = 0.019$) and COWAT ($p = 0.035$). The statistical analyses took into account this difference (see “Statistical analysis” section).

Cognitive performance for all tests administrated among the MS patients and HCs is shown in Table 2a. ANOVA showed significant effects between group for all tests. Effect sizes ranged from 0.33 for PASAT-3 to 1.18 for DKEFS-DS. Moreover, to highlight any differences between subgroups, we considered multiple comparisons (see Table 2b). ANCOVA showed that all tests significantly discriminate between HCs and the Relapsing group; moreover, all tests significantly discriminate HCs from the Progressive group; CVLT-TL ($p = 0.014$) and BVMT-TL ($p < 0.001$) significantly discriminates the Relapsing from the Progressive group. However, when adjusted for EDSS, the differences of BVMT-TL and CVLT-TL did not remain significant ($p = 0.587$ and $p = 0.362$, respectively).

Since the statistical distribution of PASAT and BVMT-R presented the peculiarity of inflation at the minimum of their distribution (PASAT at zero and BVMT-R at 20), the semi-continuous and zero-truncated models (see “Statistical analysis” section) indicated that the disease had a significant effect on BVMT-TL scores. According to this model, the expected BVMT-TL score for the Relapsing group was equal to 0.87 (95 % confidence interval (CI) = 0.78–0.96) times the reference value of HC group (T score = 50). Thus, we estimated that BVMT-TL in the Relapsing group was 43.4 (95 % CI = 42.8–52.6). Similarly, the BVMT-TL expected score in the Progressive group was 37.6 (95 % CI = 32.2–45), holding constant all other variables in the model. Also, for the BVMT-DR score, the model revealed an effect of the disease. The expected BVMT-DR score for the Relapsing group was equal to 0.86 (95 % CI = 0.78–0.95) times the reference value of the HC group (T score = 54.7), so the estimated BVMT-DR score in the Relapsing group was 46.9 (95 % CI = 42.5–51.8). The expected BVMT-DR score in the Progressive group was 37.6 (95 % CI = 32.0–44.2), holding constant all other variables in the model.

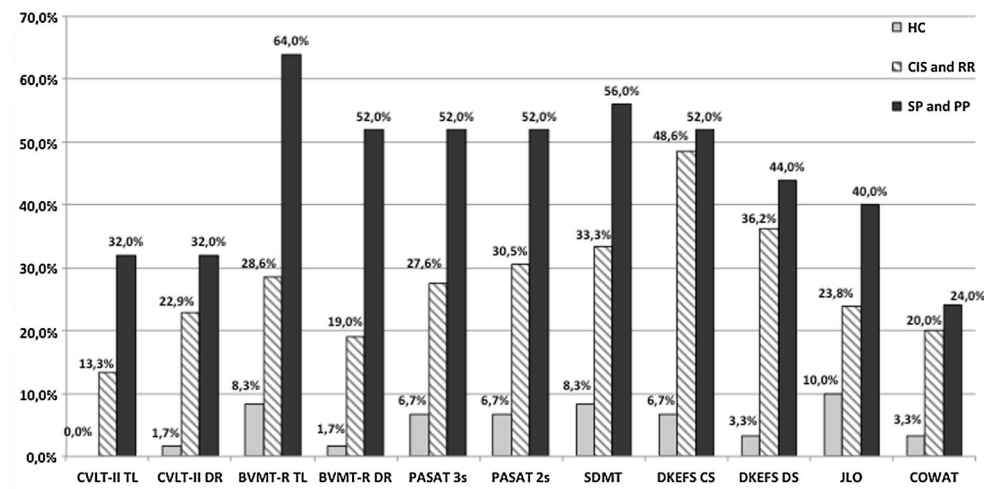
About PASAT-3, applying the two-part model, we observed a significant association between the probability of a zero score and the diagnosis, in particular with the Progressive group (OR_{Relapsing vs HC} = 3.42, 95 %

Fig. 1 Frequencies of impairment using a cut-off of 5th percentile



CVLT-TL. California Verbal Learning Test-Total Learning; CVLT-DR. CVLT-Delay Recall; BVMT-TL. Brief Visuo-spatial Memory Test-Total Learning; BVMT-DR. BVMT-Delay Recall; SDMT. Symbol Digit Modalities Test; PASAT - 3 and 2. Paced Auditory Serial Addition Test 3 and 2 seconds; JLO. Judgment of Line Orientation; DKEFS-CS. Delis Kaplan Executive Function System-Sorting Test-Correct Sort; DKEFS-DS. DKEFS-Description Score; COWAT. Controlled Oral Words Association Test.

Fig. 2 Frequencies of cognitive impairment in MS patients and HC



RR=Relapsing-Remitting; CIS=Clinical Isolated Syndrome; SP=Secondary Progressive; PP=Primary Progressive
 CVLT-TL. California Verbal Learning Test-Total Learning; CVLT-DR. CVLT-Delay Recall; BVMT-TL. Brief Visuo-spatial Memory Test-Total Learning; BVMT-DR. BVMT-Delay Recall; SDMT. Symbol Digit Modalities Test; PASAT - 3 and 2. Paced Auditory Serial Addition Test 3 and 2 seconds; JLO. Judgment of Line Orientation; DKEFS-CS. Delis Kaplan Executive Function System-Sorting Test-Correct Sort; DKEFS-DS. DKEFS-Description Score; COWAT. Controlled Oral Words Association Test.

CI = 0.95–12.25, $p = 0.177$; $OR_{\text{Progressive vs HC}} = 14.93$, 95 % CI = 3.67–60.79, $p < 0.001$, but when a positive score was obtained, no difference could be attributed to the disease.

With regard to PASAT-2, the two-part model suggested that the diagnosis was associated with a high probability of a zero response ($OR_{\text{Relapsing vs HC}} = 6.41$, 95 % CI = 1.44–28.79, $p = 0.045$; $OR_{\text{Progressive vs HC}} = 22.79$, 95 % CI = 4.53–114.65, $p < 0.001$) and also was significantly related to positive values, although only two HCs obtained a 0 score. Particularly, the Relapsing group compared with HCs had a lower probability of obtaining an elevated positive score to PASAT-2; in the Progressive group, a significant effect did not emerge because a high percentage of MS patients obtained 0 scores. The patients with a 0 score at the PASAT-2 were not able to perform the test for general cognitive deficits. The two HCs, instead,

refused to perform the task due to the difficulty to make the calculations and the resulting stress.

Based on scores equal or below the 5th percentile, frequencies of impairment in MS patients were as follows: BVMT-TL (total learning) 35.4 %, PASAT-3 32.3 % and PASAT-2 34.6 %, SDMT 37.7 %, BVMT-DR (delayed recall) 25.4 %, CVLT-DR (delayed recall) 24.6 %, JLO 26.9 %, CVLT-TL (total learning) 16.9 %, DKEFS-DS 49.2 %, DKEFS-CS 37.7 %, COWAT 20.8 % (Fig. 1). Generally, the Relapsing group performed better than the progressive one, as shown in Fig. 2.

According to the definition of CImp previously reported, 70.8 % of patients are impaired. In order to reveal different degrees of severity in cognitive dysfunction, 18.5 % showed a mild level of impairment (two tests failed), 10.0 % a moderate level (three tests failed) and 42.3 % a severe level (four or more tests failed). Regarding the

number of cognitive domains impaired, 57.7 % of MS patients showed at least two domains compromised (out of them, 20 % were mildly, 16.9 % moderately and 20.8 % severely impaired).

The logistic regression analysis, with CImp as dependent variables, revealed a significant effect of the disease ($OR_{MS\ patients\ vs\ HC} = 14.79$, 95 % CI = 6.02;36.31, $p < 0.001$); considering diagnosis groups (HC, relapsing and progressive groups), we observed that the probability of failure on two or more tests increases with the extent of the disease course ($OR_{relapsing\ vs\ HC} = 13.88$, 95 % CI = 5.47;35.19, $p < 0.001$; $OR_{progressive\ vs\ HC} = 19.00$, 95 % CI = 4.83;74.70, $p < 0.001$). Considering only MS patients, the association between the probability of failure on two or more tests and the extent of the disease course did not remain after the adjustment for disease duration ($OR_{relapsing\ vs\ progressive} = 1.37$, 95 % CI = 0.35; 5.43, $p = 0.651$).

Principal component analysis (PCA)

In order to evaluate the construct validity, PCA was performed. Four components were obtained for MS patients (Table 3): (1) visual–spatial memory and processing speed,

Table 3 Results of Principal Component Analysis

MS Patients ($n = 130$)					
	C1	C2	C3	C4	h^2
BVMT-R TL	0.863	0.197	0.083	0.256	0.86
BVMT-R DR	0.856	0.283	0.094	0.200	0.86
SDMT	0.723	0.294	0.165	0.190	0.67
COWAT	0.509	0.077	0.370	0.275	0.48
PASAT 3s	0.370	0.826	0.185	0.147	0.88
PASAT 2s	0.321	0.845	0.136	0.148	0.86
JLO	0.070	0.623	0.253	0.355	0.58
DKEFS-CS	0.116	0.203	0.946	0.123	0.97
DKEFS-DS	0.185	0.201	0.921	0.155	0.95
CVLT-II TL	0.317	0.233	0.142	0.846	0.89
CVLT-II DR	0.317	0.234	0.179	0.815	0.85
% variance explained	51.7	12.8	8.4	7.4	80.3

Bold reflects the tests that load in the same component.

Table contains component loadings using varimax rotation for each component (C). Communalities (h^2) are the proportion of each variable's variance that can be explained by the principal components. Final row shows per cent of variance explained by all 11 variables for factor and the total model.

MS multiple sclerosis, CVLT-TL California verbal learning test-total learning, CVLT-DR CVLT-delay recall, BVMT-TL Brief visuospatial memory test-total learning, BVMT-DR BVMT-delay recall, SDMT symbol digit modalities test, PASAT-3 and 2 paced auditory serial addition test, JLO Judgment of line orientation, DKEFS-CS Delis Kaplan executive function system-sorting test-correct sort, DKEFS-DS DKEFS-description score, COWAT controlled oral words association test

(2) working memory, (3) executive functions and (4) verbal memory. Overall, the four components explained 80.3 % of the variance of the 11 NP subtests included in the analysis and alone the first component explained 51.7 % of the total variance. PASAT-3 and PASAT-2, in addition to the second component, were discreetly loading also on the first component (0.370–0.321), being a measure of processing speed. Similarly, the two CVLT-II variables were discreetly loaded onto the first component (0.317–0.317), as a measure of memory. COWAT loaded quietly onto the third component (0.370) as a measure of executive functions. Overall, the cumulability for JLO and COWAT was lower with respect to the others (0.58 and 0.48, respectively) indicating that the four components explained a modest percentage of the variance of these two tests.

Discussion

Our study has shown the validity of the Italian version of the MACFIMS battery. Our data confirmed the group differences between MS patients and HCs in cognitive performance as assessed by the MACFIMS. Overall, with respect to the Relapsing group, the progressive one had lower performances in every test considered [33, 34] (Fig. 2). Effect sizes ranged from medium to very large, and verbal memory and executive functions tests were most sensitive to the disease.

According to previous literature [1, 7, 10], more than half of MS patients (70.8 %) exhibit CImp (at least two tests failed); out of them 18.5 % were mildly, 10 % were moderately, and 42.3 % were severely compromised. Differences in the frequency of impairment can be expected to differ across studies because of variability in testing methods and sample composition. This sample is representative of patients attending a hospital-based MS centre. Lower frequencies of impairment would be expected in a population-based sample, and higher frequencies might be found in samples of patients seen only for specified clinical purposes.

The cognitive functions mostly affected were visual memory and information processing speed/executive functions (Fig. 1). Considering the number of cognitive domains impaired, we found that 57.7 % of MS patients showed at least two domains compromised (out of them 20.0 % mildly, 16.9 % moderately and 20.8 % severely). This classification may be more specific to identify MS patients with a clear cognitive impairment; in fact, patients with two tests failed in the same domain are not considered mDCI.

In evaluating the PCAs, we identified four components: visual–spatial memory/processing speed; working memory; executive functions; verbal memory. The first

component explained 51.7 % of the total variance. This high percentage could be due to the number of variables with non-negligible loadings on this component: BVMT-R, SDMT and COWAT (characterized by highest loadings on C1), PASAT and CVLT-II showed loadings higher than 0.3. Our data identified two clear verbal episodic memory and executive function components, and a more general component involving processing speed, visual memory and spatial functions. It is important to clarify that the MACFIMS test represents a minimal assessment of cognitive functions in MS [13] and does not assess cognitive functioning in a comprehensive manner. Therefore, these results should not be viewed as a representation of cognitive functioning in MS. However, we observed two single components representing verbal memory and executive functions, two tests that in our study were most sensitive to the disease.

There are some limitations in our study. We enrolled MS patients in three MS centre, all in Rome; to increase the external validity of the study would be desirable in the future to recruit MS patient in MS centre in other Italian cities. Moreover, higher percentage of cognitive impairment highlighted could be due to hospital MS centres particularly careful to the cognitive assessment. Another limitation was the lack of data about MS patients' cognitive reserve and its correlation with the neuropsychological results; cognitive reserve, indeed, could be an important factor to explain cognitive profile in MS. Also, the limited size of our HC sample not allow to provide normative Italian data for the CVLT-II and the BVMT-R so we used for these tests US normative data.

In conclusion, our study is the first to validate an Italian version of the MACFIMS. Several aspects of validity have been demonstrated: criterion (all MACFIMS subtests discriminate MS patients from HCs) and, partially, construct (factorial structure of the MACFIMS). Future work will investigate the test–retest reliability of these tests, vocational status since cognitive dysfunction has a major burden on level of employment reached and the longitudinal course of NP dysfunction in MS using these measures.

Compliance with ethical standards

Funding This study not received any funding.

Conflict of interest No conflict declared by any authors.

References

- Chiaravalloti ND, DeLuca J (2008) Cognitive impairment in multiple sclerosis. *Lancet Neurol* 7(12):1139–1151
- Amato MP, Ponziani G, Pracucci G, Bracco L, Siracusa G, Amaducci L (1995) Cognitive impairment in early-onset multiple sclerosis: pattern, predictors, and impact on everyday life in a 4-year follow-up. *Arch Neurol* 52(2):168–172
- Deloire MSA, Salort E, Bonnet M, Arimone Y, Boudineau M, Amieva H et al (2005) Cognitive impairment as marker of diffuse brain abnormalities in early relapsing remitting multiple sclerosis. *J Neurol Neurosurg Psychiatry* 76(4):519–526
- Amato MP, Zipoli V, Goretti B, Portaccio E, De Caro MF, Ricchiuti L et al (2006) Benign multiple sclerosis. *J Neurol* 253(8):1054–1059
- Pittock SJ, McClelland RL, Mayr WT, Jorgensen NW, Weinschenker BG, Noseworthy J et al (2004) Clinical implications of benign multiple sclerosis: a 20-year population-based follow-up study. *Ann Neurol* 56(2):303–306
- Huijbregts SCJ, Kalkers NF, de Sonneville LMJ, de Groot V, Polman CH (2006) Cognitive impairment and decline in different MS subtypes. *J Neurol Sci* 245(1):187–194
- Jongen PJ, Ter Horst AT, Brands AM (2012) Cognitive impairment in multiple sclerosis. *Minerva Med* 103(2):73–96
- DeLuca J, Chelune GJ, Tulsky DS, Lengenfelder J, Chiaravalloti ND (2004) Is speed of processing or working memory the primary information processing deficit in multiple sclerosis? *J Clin Exp Neuropsychol* 26(4):550–562
- Denney DR, Sworowski LA, Lynch SG (2005) Cognitive impairment in three subtypes of multiple sclerosis. *Arch Clin Neuropsychol* 20(8):967–981
- Benedict RHB, Cookfair D, Gavett R, Gunther M, Munschauer F, Garg N et al (2006) Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *J Int Neuropsychol Soc* 12(04):549–558
- Cook SD (2001) *Handbook of multiple sclerosis*: CRC Press, Boca Raton
- Rao SM, Leo GJ, Bernardin L, Unverzagt F (1991) Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology* 41(5):685–691
- Benedict RHB, Fischer JS, Archibald CJ, Arnett PA, Beatty WW, Bobholz J et al (2002) Minimal neuropsychological assessment of MS patients: a consensus approach. *Clin Neuropsychol* 16(3):381–397
- Benton AL (1994) *Contributions to neuropsychological assessment: a clinical manual*, Oxford University Press, Oxford
- Delis DC, Kramer JH (2000) *California Verbal Learning Test: CvLT-II; Adult Version; Manual*
- Benedict R (1997) *Brief visuospatial memory test-revised: professional manual*. Psychological Assessment Resources, Inc, Odessa
- Gronwall DMA (1977) Paced auditory serial-addition task: a measure of recovery from concussion. *Percept Mot Skills* 44(2):367–373
- Smith A (2002) *Symbol digit modalities test: Manual*, Western Psychological Corporation, Torrance
- Delis DC, Kaplan E, Kramer JH (2001) *Delis-Kaplan executive function system (D-KEFS)*, Psychological Corporation
- Benedict RHB (2005) Effects of using same-versus alternate-form memory tests during short-interval repeated assessments in multiple sclerosis. *J Int Neuropsychol Soc* 11(06):727–736
- Houtchens MK, Benedict RHB, Killiany R, Sharma J, Jaisani Z, Singh B et al (2007) Thalamic atrophy and cognition in multiple sclerosis. *Neurology* 69(12):1213–1223
- Mike A, Glanz B, Hildenbrand P, Meier D, Bolden K, Liguori M et al (2011) Identification and clinical impact of multiple sclerosis cortical lesions as assessed by routine 3T MR imaging. *American Journal of Neuroradiology*. 32(3):515–521
- Strober L, Englert J, Munschauer F, Weinstock-Guttman B, Rao S, Benedict RHB (2009) Sensitivity of conventional memory tests in multiple sclerosis: comparing the Rao Brief

- Repeatable Neuropsychological Battery and the Minimal Assessment of Cognitive Function in MS. *Mult Scler* 15(9):1077–1084
24. Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L et al (2005) Diagnostic criteria for multiple sclerosis: 2005 revisions to the “McDonald Criteria”. *Ann Neurol* 58(6):840–846
 25. Beck AT, Steer RA, Ball R, Ranieri WF (1996) Comparison of Beck Depression Inventories-IA and-II in psychiatric outpatients. *J Pers Assess* 67(3):588–597
 26. Carlesimo GA, Caltagirone C, Gainotti G, Fadda L, Gallassi R, Lorusso S et al (1996) The mental deterioration battery: normative data, diagnostic reliability and qualitative analyses of cognitive impairment. *Eur Neurol* 36(6):378–384
 27. Ferracuti S, Cannoni E, Sacco R, Hufty AM (2000) Contributi per un assessment neuropsicologico. *Manuale clinico Giunti Organizzazioni Speciali*, Firenze
 28. Amato MP, Portaccio E, Goretti B, Zipoli V, Ricchiuti L, De Caro MF et al (2006) The Rao’s Brief Repeatable Battery and Stroop Test: normative values with age, education and gender corrections in an Italian population. *Mult Scler* 12(6):787–793
 29. Nocentini U, Giordano A, Di Vincenzo S, Panella M, Pasqualetti P (2006) The symbol digit modalities test-oral version: Italian normative data. *Funct Neurol* 21(2):93–96
 30. Mattioli F, Stampatori C, Bellomi F, Scarpazza C, Galli P, Guarneri C et al (2014) Assessing executive function with the D-KEFS sorting test: normative data for a sample of the Italian adult population. *Neurol Sci* 35(12):1895–1902
 31. Nocentini U, Pasqualetti P, Bonavita S, Buccafusca M, De Caro MF, Farina D et al (2006) Cognitive dysfunction in patients with relapsing-remitting multiple sclerosis. *Mult Scler* 12(1):77–87
 32. Cohen J (1992) A power primer. *Psychol Bull* 112(1):155
 33. Comi G, Filippi M, Martinelli V, Campi A, Rodegher M, Alberoni M et al (1995) Brain MRI correlates of cognitive impairment in primary and secondary progressive multiple sclerosis. *J Neurol Sci* 132(2):222–227
 34. Kraus JA, Schütze C, Brokate B, Kröger B, Schwendemann G, Hildebrandt H (2005) Discriminant analysis of the cognitive performance profile of MS patients differentiates their clinical course. *J Neurol* 252(7):808–813