A multifaceted approach to understand cognitive impairment in MS: exploring the nonlinearity of cognition

Dam, M. van

2

The potential of diagnostic biomarkers for cognition

2.1

A multimodal marker for cognitive functioning in multiple sclerosis: the role of NfL, GFAP and conventional MRI in predicting cognitive functioning in a prospective clinical cohort

M. van Dam, B.A. de Jong, E.A.J. Willemse, I.M. Nauta, M. Huiskamp, M. Klein, B. Moraal, S. de Geus-Driessen, J.J.G. Geurts, B.M.J. Uitdehaag, C.E. Teunissen & H.E. Hulst.

Journal of Neurology (2023).

ABSTRACT

Background: Cognitive impairment in people with MS (PwMS) has primarily been investigated using conventional imaging markers or fluid biomarkers of neurodegeneration separately. However, the single use of these markers do only partially explain the large heterogeneity found in PwMS.

Objective: To investigate the use of multimodal (bio)markers: i.e., serum and cerebrospinal fluid (CSF) levels of neurofilament light chain (NfL) and glial fibrillary acidic protein (GFAP) and conventional imaging markers in predicting cognitive functioning in PwMS.

Methods: Eighty-two PwMS (56 females, disease duration = 14 ± 9 years) underwent neuropsychological and neurological examination, structural magnetic resonance imaging, blood sampling and lumbar puncture. PwMS were classified as cognitively impaired (CI) if scoring ≥ 1.5SD below normative scores on ≥ 20% of test scores. Otherwise, PwMS were defined as cognitively preserved (CP). Associations between fluid and imaging (bio)markers were investigated, as well as binary logistic regressions to predict cognitive status, Finally, a multimodal marker was calculated using statistically important predictors of cognitive status.

Results: Only higher NfL levels (in serum and CSF) correlated with worse processing speed (r = -0.286, p = 0.012 and r = -0.364, p = 0.007, respectively). sNfL added unique variance in the prediction of cognitive status on top of grey matter volume (NGMV), p = 0.002). A multimodal marker of NGMV and sNfL yielded most promising results in predicting cognitive status (sensitivity = 85%, specificity = 58%).

Conclusion: Fluid and imaging (bio)markers reflect different aspects of neurodegeneration and cannot be used interchangeably as markers for cognitive functioning in PwMS. The use of a multimodal marker, i.e. the combination of grey matter volume and sNfL, seems most promising for detecting cognitive deficits in MS.

INTRODUCTION

Cognitive impairment is one of the most disabling symptoms of multiple sclerosis (MS), significantly hampering day-to-day functioning.[1] In an effort to monitor cognitive functioning in MS, understanding its underlying neurobiological correlates is of utmost importance. To date, most studies investigated magnetic resonance imaging (MRI) characteristics in relation to cognitive performance, which has taught us that cognitive impairment is associated with neurodegeneration such as cortical and deep grey matter atrophy,[2, 3] as well as with functional impairment of neuronal networks.[4] However, the clinical implementation of these prognostic biomarkers is limited, as these markers cannot fully account for the large heterogeneity found between people with MS (PwMS).[5] The complex pathology of MS, including inflammation, demyelination and neurodegeneration warrants a multimodal biomarker linking both molecular and imaging biomarkers.[6]

Recent studies focused on the combination of neurofilament light chain levels in serum (sNfL) and conventional imaging markers (e.g., lesion load and grey matter volume) for predicting cognitive functioning in PwMS.[7, 8] NfL reflects the major intermediate cytoskeletal protein of axons and is considered to be a marker for neuro-axonal damage.[9] Indeed, increased levels of NfL in the serum and cerebrospinal fluid (CSF) of PwMS have been related to cognitive impairment[10] and decreased performance on multiple cognitive domains,[7, 11] in various disease stages,[8, 12] showing promising predictive value over time.[10] Another potentially interesting biomarker for the assessment of neurodegeneration in MS is glial fibrillary acidic protein (GFAP), the intermediate cytoskeletal protein of astrocytes.[13, 14] Both serum and CSF levels of GFAP have been shown to relate to disease type (i.e., increased levels of GFAP in progressive PwMS)[15] and disease severity (i.e., increased GFAP levels were associated with higher physical disability and longer disease durations).[13, 16] However, the association between GFAP and cognitive functioning in MS has yet to be established.

The aim of current study was to compare, confirm and combine (bio)markers of neurodegeneration (i.e., serum and CSF levels of both NfL and GFAP and conventional imaging markers) for its role in cognition in a clinical sample of PwMS that visited our outpatient clinic because of perceived cognitive complaints.

MATERIALS AND METHODS

Study population

In total, 129 PwMS that visited the Second Opinion Multiple Sclerosis and Cognition (SOMSCOG) outpatient clinic of the MS Center Amsterdam between February 2017 and November 2020 (82 females, mean age: 48 ± 11 years) were included in this

cross-sectional study. Individuals were referred to our SOMSCOG outpatient clinic because of cognitive complaints by their general physician or neurologist, and underwent an extensive diagnostic workup for cognitive impairment, including neuropsychological and neurological examination, MRI, blood sampling and lumbar puncture (all administered on the same day). PwMS were included if they gave written informed consent and had a clinically definite diagnosis of MS according to the McDonald MS criteria (2017 – revised)[17] or clinically isolated syndrome. Additionally, PwMS were only included if they underwent a performance validity test (Amsterdam Short-Term Memory (ASTM) test)[18, 19] and had reached a sufficient score on this test, resulting in the exclusion of 35 PwMS. After applying the in- and exclusion criteria, a total of 82 PwMS remained eligible for data-analysis (Figure 1). This is the second paper including data of this cohort, the first paper focused on performance validity.[19]

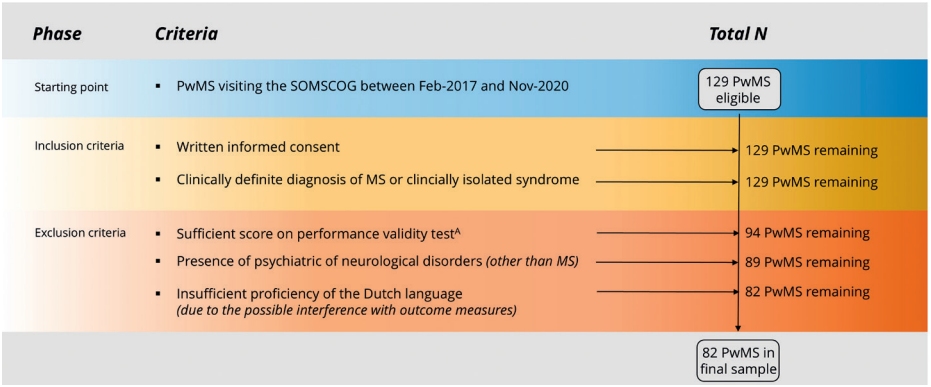

Figure 1. An overview of the included PwMS after applying the in- and exclusion criteria.
[A] PwMS were only included if they underwent a performance validity test (Amsterdam Short-Term Memory (ASTM) test) and had reached a sufficient score on this test. *Abbreviations: PwMS = People with MS. SOMSCOG = Second Opinion Multiple Sclerosis and Cognition.*

Demographics and clinical functioning

The demographic characteristics included age, sex and level of education (coded according the Verhage classification, a Dutch classification system for education).[20] Information on MS type, disease duration, and disease-modifying therapy (DMT; yes/no and if yes, first-line or second-line DMT) was collected from the medical charts. The level of physical disability was assessed by a certified examiner using the Expanded Disability Status Scale (EDSS).[21]

Neuropsychological examination

Cognitive functioning was measured using an Dutch adaptation of the MACFIMS,[22] consisting of the following five (sub)-domains: processing speed (Symbol Digit Modalities Test[23] and Stroop Color-Word Test cards I and II),[24] verbal memory (Dutch

version of the California Verbal Learning Test version 2),[25] visuospatial memory (Brief Visuospatial Memory Test – Revised),[26] executive function (EF)-verbal fluency (Controlled Oral Word Association Test)[27] and EF-inhibition (Stroop Color-Word Test Interference score).[24] The neuropsychological examination was administered by a certified clinical neuropsychologist at the department of Medical Psychology of the hospital. Cognitive test scores were corrected for age, sex and educational level (if predictor was below $\alpha = 0.1$ in regression analysis),[28] and transformed into five domain-specific Z-scores, based on a normative sample of Dutch healthy controls ($N = 407$). PwMS were classified as cognitively impaired (CI) if ≥ 20% of the cognitive test scores were ≥ 1.5SD below normative scores (corresponding to ≥ 3 out of 11 test scores), or otherwise as cognitively preserved (CP).[29]

Since multiple psychological factors are known for its impact on cognition, henceforth referred to as patient-reported outcomes (PROMS), current study protocol included: self-perceived cognitive problems (using the MS Neuropsychological Questionnaire–Patient version, MSNQ-P),[30] symptoms of anxiety and depression (using the Hospital Anxiety and Depression Scale, HADS),[31] levels of fatigue (using the Checklist Individual Strength-20 Revised, CIS20-R),[32] and sleep-related problems (using the Athens Insomnia Scale, AIS).[33] For all aforementioned questionnaires, higher scores indicate more symptoms.

Imaging markers

Seventy-eight PwMS (~91%) underwent MR scanning on a 3-Tesla whole-body scanner (General Electric Signa-HDxt, Milwaukee, WI, USA), with an 8-channel head coil (see the Supplemental Methods for the detailed MRI protocol). White matter lesions on FLAIR images were segmented after which lesions on T1-weighted images were filled using an automated lesion-filling technique (LEAP).[34] The SIENAX pipeline was used to obtain estimates of global white matter and grey matter volumes. FIRST was then applied for the automatic segmentation and calculation of bilateral hippocampus and thalamus volumes, areas known to be associated with cognitive decline in MS. All volumes were corrected for head size using the V-scaling factor obtained by SIENAX.

Fluid biomarkers

Blood samples were collected in serum tubes through venipuncture for 78 PwMS (~95%), whereas CSF samples were obtained by lumbar puncture for 54 PwMS (~66%; see Supplemental Methods).[35] Serum and CSF NfL and GFAP levels were quantified in parallel on Single Molecule Array (Simoa) HD-1 analyzers (Quanterix) using the Simoa NF-light Advantage Kit (Quanterix) and the Simoa GFAP Discovery Kit (Quanterix).[36] Paired CSF and serum samples per PwMS were analyzed within one run. The average intra-assay coefficient of variation (CV) of sample duplicates was 5.3 ± 4.1% for serum GFAP (sGFAP) and 4.1 ± 3.72% for CSF GFAP (cGFAP); NfL

measurements were performed in singlicates. One serum NfL (sNfL) measure failed due to a debris error (total *N* sNfL = 77). For sGFAP, two samples were measured in singlicate and for cGFAP, one sample was measured in singlicate.

Statistical analysis

Normality of variables was explored using the Shapiro-Wilk test and histogram inspection of the residuals. In case of non-normally distributed data, logarithmic (for NLV, sNfL and CSF NfL (cNfL)) or square root transformations (for disease duration, sGFAP and cGFAP) were applied. Fluid and imaging (bio)markers were corrected, regression based, for sex and age (if *p*-value of demographic variable was smaller than $\alpha = 0.1$ in the model). The following correction formulas were applied:

$$\text{sNfL (corrected)} = \text{LG10(sNfL)} - 0.009 * \text{age (in years)}.$$

$$\text{NGMV (corrected)} = \text{NGMV} + 2.622 * \text{age (in years)} - 46.680 * \text{sex (0 = female)}.$$

To investigate differences between CP and CI, independent samples *t*-tests were applied (for age, disease duration, PROMS, corrected fluid and imaging (bio)markers), or Mann-Whitney a *U*-test for EDSS. Chi-square tests were used to investigate CP vs. CI differences in sex, educational level type of MS, use of DMT and type of DMT (first-line vs. second-line). Pearson correlations were used to investigate the association (1) between corrected fluid biomarkers and cognitive domains, (2) between corrected imaging markers and cognitive domains and (3) between corrected fluid and imaging (bio)markers. Outcomes were Bonferroni corrected: the α-level of 0.05 was divided by the number of fluid ($p < 0.0125$) or imaging (bio)markers ($p < 0.01$).

Two binary logistic regression models (using either serum or CSF markers) with forward selection were run to identify the predictors of cognitive status. The choice of running two separate regressions was made as a significant part of the sample did not receive the lumbar puncture. To reduce the number of variables in the prediction models, the imaging markers were only inserted as predictor if significant group differences were present.

In a post-hoc analysis, the predictive value of the predictors alone was explored, while a weighted composite score of the significant predictors was calculated using the standardized betas. The composite score was further evaluated by drawing receiver operator characteristic curves and calculating the areas under the curves (AUCs) to determine diagnostic accuracy. Interpretation of the AUC was as follows: an AUC of 0.60 and 0.70 could be considered 'poor', an AUC of 0.70 to 0.80 could be considered 'acceptable' or 'fair', and AUC of 0.80 to 0.90 could be considered 'good' and above 0.9 an AUC could be considered 'excellent'.[37] Significance level was set at α-level <0.05 and the statistical analyses were performed in SPSS 28.0 (IBM, Armonk, NY, USA).

RESULTS

Study population

The final sample consisted of 82 PwMS (56 females, mean age = 47 ± 9 years, mean disease duration = 14 ± 9 years). Information on demographics, disease related variables, PROMS and imaging and fluid biomarkers can be found in Table 1. In 75% of the PwMS, the MSNQ-P was above the threshold of 27,[38] indicating the presence of self-perceived cognitive problems at the time of the visit. Performance on individual cognitive tests is included in Supplementary Table 2.

Table 1. Information on demographics, disease related variables, patient-reported outcome measures, imaging markers (in ml) and fluid biomarkers (in pg/ml) displayed for cognitive groups.

	CP (*N* = 36)	**CI (*N* = 46)**	***p*-value**
Demographics			
Sex (female : male)	30 : 6	20 : 26	**.010***
Age	47.36 ± 9.95	47.07 ± 9.00	.888
Educational level	6 [5-6]	6 [5-6]	.913
Clinical functioning			
Disease duration[a]	13.04 ± 9.27	13.93 ± 8.92	.693
EDSS	3.5 [2.5–4.0]	4.0 [3.0–4.5]	**.012***
MS type (CIS/RRMS/PPMS/SPMS/UN)	(2/26/4/4/0)	(2/28/3/12/1)	.398
Use of DMT (yes : no)	18 : 18	21 : 25	.696
Type of DMT (first-line: second-line)	13 : 5	15 : 10	.407
Patient-reported outcome measures			
HADS anxiety	8.43 ± 4.13	8.73 ± 4.58	.765
HADS depression	6.40 ± 4.09	7.64 ± 4.23	.194
CIS20-R (fatigue)	90.21 ± 21.82	94.86 ± 17.20	.295
MSNQ-P (cognitive complaints)	32.77 ± 10.28	33.53 ± 8.56	.738
AIS (sleep-related problems)	6.80 ± 4.63	8.22 ± 4.60	.175
Imaging markers (ml)[b]			
NGMV	805.03 ± 63.30	756.45 ± 59.89	**.002***
NWMV	684.12 ± 46.41	666.98 ± 48.43	.141
NLV[c]	22.60 ± 20.38	35.27 ± 27.09	**.010***
Hippocampi	9.35 ± 1.11	8.57 ± 1.34	.051
Thalami	19.28 ± 2.25	17.18 ± 2.76	**.002***
Fluid biomarkers (pg/ml)			
sNfL[c,d]	8.45 [5.19–12.67]	10.54 [8.41–15.25]	**.010***

Table 1. Continued

	CP (*N* = 36)	**CI (*N* = 46)**	***p*-value**
sGFAP[a,e]	103.37 [79.59–147.88]	129.52 [92.57–184.89]	**.035***
cNfL[c,f]	561.11 [342.63–739.91]	579.74 [502.50–1109.46]	.267
cGFAP[a,f]	7467.53 [5452.89–9223.88]	8039.92 [6818.75–9307.84]	.073

Displayed are the mean and standard deviation of continuous variables, the median and interquartile range of ordinal or non-normally distributed data. Imaging markers and fluid biomarkers were corrected for age and sex (if appropriate) before tested. ***p* < 0.05. [a]Variable was square root-transformed before tested. [b]*N* = 78 (*N* CP = 36, *N* CI = 42). All volumes were normalized using the V-scaling factor. [c]Variable was log-transformed before tested. [d]*N* = 77 (*N* CP = 33, *N* CI = 44). [e]*N* = 78 (*N* CP = 33, *N* CI = 45). [f]*N* = 54 (*N* CP = 23, *N* CI = 31). *Abbreviations*: *CP = cognitively preserved; CI = cognitively impaired; EDSS = Expanded Disability Status Scale; CIS = clinically isolated syndrome; RRMS = relapsing-remitting MS; PPMS = primary progressive MS; SPMS = secondary progressive MS; UN = Unknown; DMT = disease-modifying therapy; HADS = Hospital Anxiety and Depression Scale; CIS20-R = Checklist Individual Strength 20 – Revised; MSNQ = MS Neuropsychological Questionnaire. AIS = Athens Insomnia Scale; NGMV = normalized grey matter volume; NWMV = normalized white matter volume; NLV = normalized lesion volume; sNfL = serum neurofilament light (NfL); sGFAP = serum glial fibrillary acidic protein (GFAP); cNfL = CSF NfL; cGFAP = CSF GFAP.*

Differences between cognitive groups

An overview of the differences between groups can be found in Table 1 (in Supplementary Table 3 group differences with only complete data is included). Compared to CP PwMS (*N* = 36), the group of CI PwMS (*N* = 46) had worse physical disability ($p = 0.012$) and consisted of more men (16.7% versus 43.8%; $p = 0.010$). PROMS results were similar for both groups. CI PwMS had lower NGMV ($p = 0.002$, $d = 0.741$, 95% confidence interval (*95%CI*) = [0.278:1.199]), lower thalamic volume ($p = 0.002$, $d = 0.727$, *95%CI* = [0.265:1.185]) and higher NLV ($p = 0.010$, $d = -0.603$, *95%CI* = [-1.056:-0.146]), compared to CP PwMS. As depicted in Figure 2, increased levels of sNfL and sGFAP were found in CI compared to CP PwMS ($p = 0.010$, $d = -0.605$, *95%CI* = [-1.064:-0.141]; $p = 0.035$, $d = -0.492$, *95%CI* = [-0.947:-0.035], respectively). No differences were found for cNfL and cGFAP. In a final step, a sensitivity analysis was performed. By only including PwMS who have CSF measures available (*N* = 54), it was checked whether differences between CP and CI PwMS regarding sNfL and sGFAP were still present as confirmation. Although levels of sGFAP were similar between cognitive groups ($p = 0.088$), levels of sNfL were increased in CI PwMS (mean sNfL =12.95 ± 7.75 pg/ml) compared to CP PwMS (mean sNfL = 11.03 ± 10.95 pg/ml; $p = 0.043$).

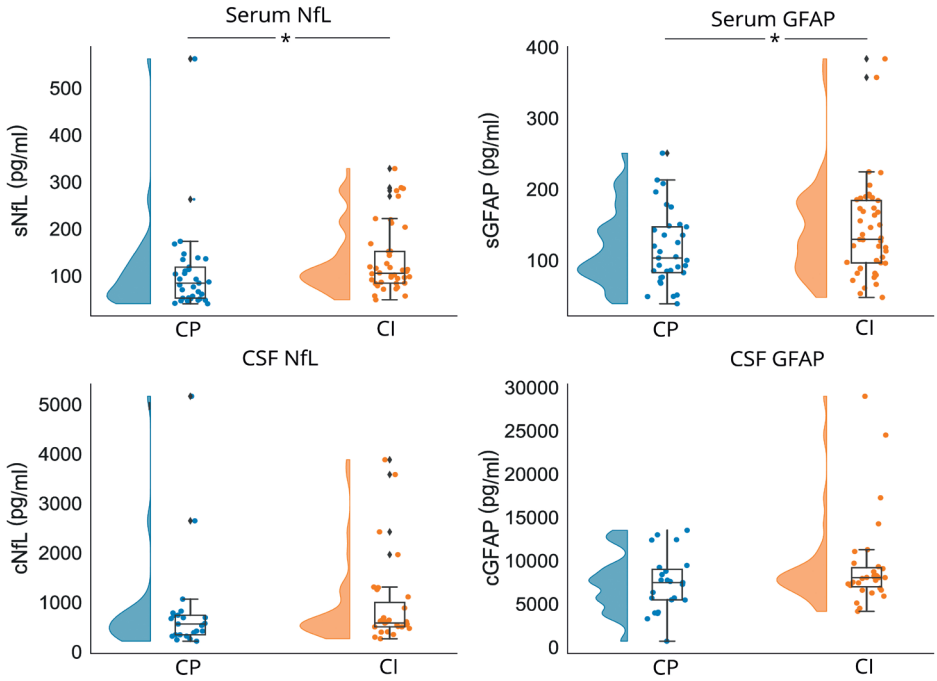

Figure 2. Differences in fluid biomarkers between cognitively preserved (CP) and cognitively impaired (CI) PwMS.
Results indicate that sNfL and sGFAP are increased in CI PwMS, compared to CP PwMS (* = $p < 0.05$). For illustrative purposes, the raw (non-transformed, not corrected) values of fluid biomarkers are shown. *Abbreviations*: *PwMS = People with MS; CP = cognitively preserved; CI = cognitively impaired; sNfL = serum neurofilament light (NfL); sGFAP = serum glial fibrillary acidic protein (GFAP); cNfL = CSF NfL; cGFAP = CSF GFAP.*

Associations between fluid and imaging (bio)markers and cognitive functioning

Fluid biomarkers and cognitive domains. As depicted in Figure 3, reduced processing speed was associated with increased levels of sNfL (r = -0.286, p = 0.012) and cNfL (r = -0.364, p = 0.007).

Imaging markers and cognitive domains. Reduced processing speed was associated with lower NGMV, thalamic and hippocampal volume (range of coefficients: 0.371-0.422), and increased NLV (r = -0.376). Reduced verbal and visuospatial memory was associated with lower NGMV and thalamic volume, with only an association with increased NLV for visuospatial memory. Correlation coefficients are included in Figure 3.

Fluid and imaging (bio)markers. Increased levels of cNfL were associated with reduced thalamic volume (r = -0.389, p = 0.004) and borderline increased NLV (r = 0.345, p = 0.011, Figure 3). Finally, increased levels of cGFAP were associated with reduced hippocampal volume (r = -0.347, p = 0.010), although this finding was

borderline significant. No other correlation between fluid and imaging (bio)markers and cognitive functioning survived correction for multiple comparisons (Figure 3).

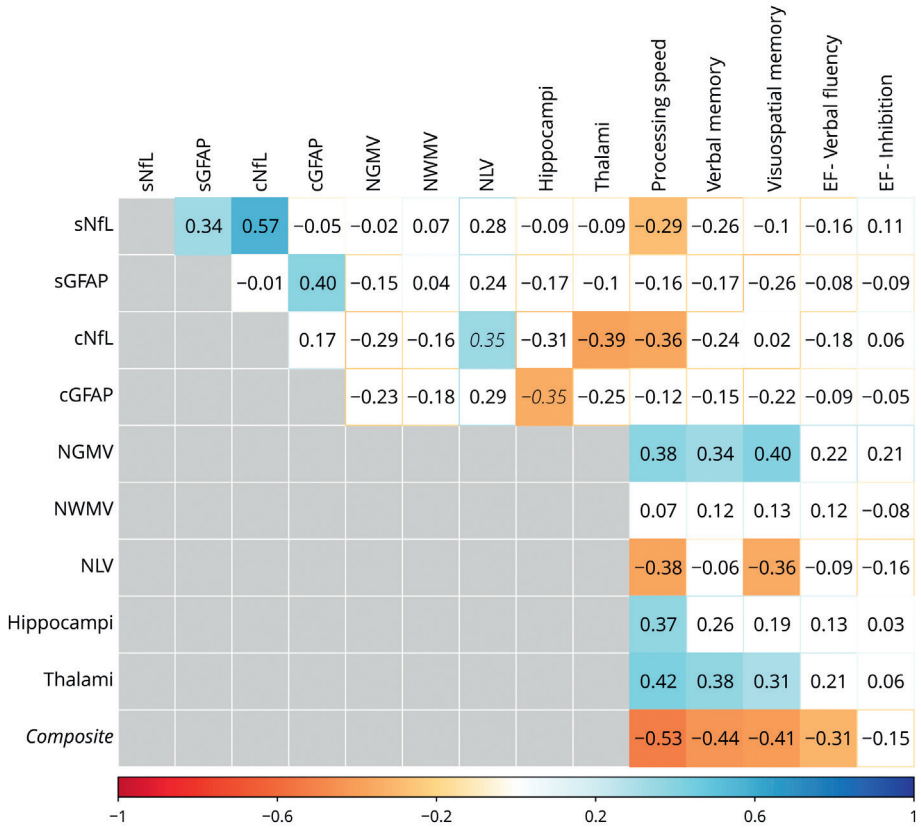

Figure 3. An overview of the correlations between fluid biomarkers, imaging markers and cognitive domains.
The correlation coefficient is displayed inside the blocks. Only significant correlations are shown in color (after correction for multiple comparisons; in italic if borderline significant). Correlation between cognitive domains and a post-hoc calculated composite score (a combination of significant predictors (sNfL and NGMV)) is depicted on the bottom row. *Abbreviations*: *sNfL = serum neurofilament light (NfL); sGFAP = serum glial fibrillary acidic protein (GFAP); cNfL = CSF NfL; cGFAP = CSF GFAP; NGMV = Normalized grey matter volume; NWMV = normalized white matter volume; NLV = normalized lesion volume; EF = executive function.*

Prediction of cognitive status

In the first logistic regression model, imaging markers (i.e., NGMV, NLV and thalami), sNfL and sGFAP were included as predictors of cognitive status (N = 73). Only NGMV and sNfL were able to predict cognitive status (Table 2). When added to the model, sNfL significantly improved the prediction of cognitive status compared to NGMV alone (sensitivity increased from 70% to 77.5%, whereas the specificity remained 60.6%; p = 0.025). The final model, including NGMV and sNfL, resulted in a sensitivity

of 77.5%, a specificity of 60.6%, a positive predictive value (PPV) of 70.5%, a negative predictive value (NPV) of 69.0% and an accuracy of 69.9%. In a post-hoc analysis, the independent value of sNfL in the prediction of cognitive status was explored (Table 2). A correct classification of CI PwMS was found in 84.1% PwMS (specificity = 48.5%).

In a second model, imaging markers (i.e., NGMV, NLV and thalami), cNfL and cGFAP were included as predictors of cognitive status (*N* = 54), with only NGMV resulting as significant predictor of cognitive status (Table 2). The final model resulted in a sensitivity of 74.2%, a specificity of 65.2%, a PPV of 74.2%, a NPV of 65.2% and an accuracy of 70.4%. Both models yielded similar explained variances of ~25% (Table 2).

Table 2. Results of binary logistic regressions of predicting cognitive status (CP vs. CI).

Models		Predictors	B	SE	Wald	*p*-value	Odds-ratio [95% CI]
1. Serum & imaging markers							
	Step 1[a]	Constant	11.97	4.16	8.27	.004	
		NGMV	-0.01	0.05	7.93	**.004***	0.987 [0.977-0.996]
	Step 2[b]	Constant	10.91	4.31	6.41	.011	
		NGMV	-0.01	0.05	8.18	**.004***	0.986 [0.976-0.996]
		sNfL	2.96	1.32	5.01	**.025***	19.204 [1.443-255.635]
1.1 Post-hoc							
	Step 1[c]	Constant	-1.45	0.74	3.86	.050	
		sNfL	3.08	1.27	5.88	**.015***	21.847 [1.808-264.044]
2. CSF & imaging markers							
	Step 1[d]	Constant	15.14	5.06	8.95	.003	
		NGMV	-0.02	0.06	8.71	**.003***	0.983 [0.972-0.994]

Model 1: *N* = 73; R^2 = 0.190 (Cox & Snell); 0.253 (Nagelkerke) [a]X^2 = 9.49 (1), *p* = 0.002 (step addition: *p* = 0.002). [b]X^2 = 15.34 (2), *p* < 0.001 (step addition: *p* = 0.016). Model 1.1: In a post-hoc comparison, the significant predictors of model 1 were compared by running a logistic regression on separate predictors. [c]*N* = 77; R^2 = 0.09 (Cox & Snell); 0.12 (Nagelkerke); X^2 = 7.04 (1), *p* = 0.008 (step addition: *p* = 0.008). Model 2: *N* = 54; R^2 = 0.182 (Cox & Snell); 0.245 (Nagelkerke); [d]X^2 = 10.85 (1), *p* < 0.001 (step addition: *p* < 0.001). *Abbreviations*: *CP = cognitively preserved; CI = cognitively impaired; B = regression coefficient; SE = standard error; Wald = Wald statistic; 95% CI = 95% confidence interval; sNfL = serum neurofilament light (NfL) in pg/ml; NGMV = normalized grey matter volume in ml.*

Multimodal marker for cognitive functioning

In a post-hoc analysis, the two significant predictors of cognitive status (NGMV and sNfL) were combined into a composite score as a multimodal marker for cognitive functioning in PwMS. Using the standardized betas as weights, the following formula was applied:

$$\text{Composite} = \text{NGMV (corrected)} * -0.014 + \text{sNfL (corrected)} * 2.955 + 20.$$

The AUC was larger for the composite score (AUC=0.751, classification = fair), compared to NGMV (AUC=0.696, classification = poor-sufficient) and sNfL (AUC=0.680, classification = poor-sufficient). Classification of cognitive status using the composite score resulted in a sensitivity of 85.0%, a specificity of 57.6%, a PPV of 70.8%, a NPV of 76.0% and an accuracy of 72.6%. A higher composite score was associated with increased performance on multiple cognitive domains: processing speed (r = -0.528, p < 0.001), verbal memory (r = -0.436, p < 0.001), visuospatial memory (r = -0.411, p < 0.001) and EF-verbal fluency (r = -0.314, p = 0.008; Figure 3).

DISCUSSION

This study investigated the relation of NfL and GFAP measured in serum and CSF and cognitive performance in PwMS presenting with cognitive complaints, and their added predictive value compared to conventional imaging markers. Based on the levels of sNfL and sGFAP we were able to distinguish cognitively preserved from cognitively impaired PwMS, albeit with limited diagnostic accuracy. Increased levels of both serum NfL and GFAP were observed in cognitively impaired PwMS compared to cognitively preserved PwMS. NfL levels (in serum and CSF) were inversely associated with processing speed, indicating that increased levels of sNfL and cNfL were associated with decreased processing speed. No correlations could be detected between GFAP (measured in either serum or CSF) and cognitive functioning in PwMS. Finally, sNfL added unique variance in the prediction of cognitive status on top of NGMV. A composite score of both measures (a multimodal marker) resulted in a fair classification of cognitive status, stressing the need for a multimodal approach when predicting cognitive functioning.

Consistent with previous literature, increased levels of sNfL were found for cognitively impaired PwMS.[7, 10, 11] Furthermore, increased levels of sNfL and cNfL were associated with reduced processing speed. Slowed processing speed, has been hypothesized to be the major driver of cognitive impairment in MS,[1] thereby possibly explaining why correlations with this specific domain are more prevalent in studies investigating NfL and cognitive functioning in MS.[10, 12] Yet, mixed results have been reported for increased levels of NfL and the performance in other cognitive domains.[11] Differences in sample size, administered neuropsychological tests, study population (i.e. focus on newly diagnosed PwMS[12] or SPMS,[8] combination of MS types, or PwMS with mild cognitive impairment)[39] and specific focus on treatment are most likely explaining these differences.[40] In our sample, the distribution of PwMS on DMT at the time of the visit (but also the distribution of PwMS on first-line DMT vs. second-line DMT) was similar between cognitive groups, thereby reducing the likelihood of impacting our findings. Nonetheless, it could have played a role on an individual level as has been shown before.[41] Although it was beyond the scope of current research, the impact of DMTs on cognitive functioning in MS warrants

further investigation.[1] Finally, although levels of sGFAP were increased in cognitively impaired PwMS, no correlations between GFAP and the cognitive domains survived correction for multiple comparisons, hereby limiting its potential as a clinical biomarker for cognitive functioning in MS.

Jakimovski et al., demonstrated in two previous studies a relatively weaker correlation between sNfL and cognition,[10] compared to correlations between sNfL and MRI outcomes.[42] As potential explanation they put forward the role of adaptive processes to significantly influence the relationships between the released NfL and cognitive test results. Subsequently, PwMS who demonstrate preserved functional connectivity, despite ongoing structural pathology, can maintain high levels of cognitive performance.[43] In the current study, associations between sNfL and imaging markers were absent. However, associations of cNfL and sNfL between both processing speed and between cNfL and imaging markers were comparable in effect size with previous studies, thereby confirming aforementioned difference.[10, 42] Interestingly, in our study, imaging markers displayed a higher number of associations with multiple cognitive domains (not only processing speed, but also verbal and visuospatial memory) compared to fluid NfL and GFAP levels, highlighting that structural pathology was present and related to several cognitive test scores. As fluid biomarkers provide a real-time evaluation of the amount of pathology compared to the less dynamic imaging markers,[9] it can be hypothesized that cognitive changes are not resulting from acute disturbances but rather from a more global effect over time on the brain in certain areas.

When added to the model, sNfL improved the prediction of cognitive status compared to NGMV alone. Especially when combining biomarkers, in our case NGMV and sNfL (the "multimodal marker") a large effect was found for processing speed, whereas medium effects were reported for verbal and visuospatial memory. Even a medium sized effect for EF-verbal fluency was found when using the multimodal marker, which was absent when investigating individual markers. The current study is one of the first studies to combine both neuroimaging and fluid biomarkers of interest to detect cognitive impairment in MS. Investigating the role of a multimodal marker for cognitive functioning in PwMS is of high importance since these different modalities might reflect different aspects of neurodegeneration, which also has been reported in Alzheimer's disease[44] and recently in MS as well.[7, 45] More specifically, previous studies investigating cross-modal fluid and imaging (bio) markers indeed show a "additive" effect of sNfL compared to cortical thickness[45] or lesion load and grey matter volume[7] in recently diagnosed PwMS. Together with our results, the added effect of sNfL highlights the necessity of using multiple sources of information to create a diagnostic marker for something as highly complex as cognitive performance, but also how these markers of neurodegeneration cannot be used interchangeably.

Nonetheless, clinical interpretation may be optimized when the full prognostic potential of sNfL for cognitive functioning will be evaluated over time, which is an important limitation of the current cross-sectional study design. The inclusion of control group would have further aided the disentanglement between normal and abnormal levels of fluid and imaging (bio)markers. Also, contrary to measurements in serum, both NfL and GFAP measured in CSF were unable to discriminate between cognitive status. The most plausible explanation for this lack of detecting a difference is the limited power (*N* CSF = 54 versus *N* = 78 for serum). Performing a lumbar puncture is rather invasive and not all PwMS wanted to partake in this procedure. Importantly, without a post-contrast sequence being available in current study protocol, it was not possible to determine whether PwMS had active lesions at the time of evaluation. As a consequence, the investigation of the effect of recent disease activity on serum and CSF levels was limited and could be considered an important avenue for future research. Finally, the inclusion of a clinical, real-life sample is one of the biggest strengths, as the PwMS are reflective of our population at the outpatient clinic with perceived cognitive complaints. At same time, being a real-life sample is also one of the main limitations. A homogenous sample is often desirable when investigating differences between groups, although data on other types of MS than RRMS is often lacking. Furthermore, given the fact that PwMS visited the outpatient clinic because of cognitive complaints, a slight bias towards cognitive impairment may have been present. The main clinical aim of the outpatient clinic is to investigate whether these complaints (or impairments) are due to MS pathology or, for instance, psychological or social factors (known to influence cognitive performance).[46] Results on PROMS measuring mood, anxiety, fatigue and sleep were therefore reported in this manuscript, showing similar scores between cognitively preserved and impaired PwMS. Consequently, the impact of these factors on cognition was also considered similar.

In conclusion, we provided novel insights into the relationship between fluid biomarkers of neurodegeneration and their relation to cognitive functioning and conventional imaging measures in PwMS. The main finding of this study is the result that sNfL explains additional variance in cognitive performance on top of NGMV. A novel insight that was further explored in our study was the potential for combining two (bio)markers from a different origin when predicting cognitive status, instead of focusing on single measures of NfL or imaging outcomes. Combining multimodal biomarkers may be the way forward in order to enable timely identification of cognitive decline in MS.

Ethical standards statement

The Medical Ethics Research Committee of Amsterdam UMC concluded that the Medical Research Involving Human Subjects Act (WMO) did not apply to this study, as the data collection was part of clinical care (number METC-2016.395).

REFERENCES

1. Benedict RHB, Amato MP, DeLuca J, Geurts JJG. Cognitive impairment in multiple sclerosis: clinical management, MRI, and therapeutic avenues. The Lancet Neurology 2020;19:860-871.
2. Eijlers AJ, van Geest Q, Dekker I, et al. Predicting cognitive decline in multiple sclerosis: a 5-year follow-up study. Brain 2018;141:2605-2618.
3. Schoonheim MM, Pinter D, Prouskas SE, et al. Disability in multiple sclerosis is related to thalamic connectivity and cortical network atrophy. Multiple Sclerosis Journal 2021:13524585211008743.
4. Di Filippo M, Portaccio E, Mancini A, Calabresi P. Multiple sclerosis and cognition: synaptic failure and network dysfunction. Nature Reviews Neuroscience 2018;19:599-609.
5. Zivadinov R, Jakimovski D, Gandhi S, et al. Clinical relevance of brain atrophy assessment in multiple sclerosis. Implications for its use in a clinical routine. Expert review of neurotherapeutics 2016;16:777-793.
6. Stefano ND, Sormani MP. Combining biomarkers to profile multiple sclerosis patients. Nature Reviews Neurology 2020;16:463-464.
7. Brummer T, Muthuraman M, Steffen F, et al. Improved prediction of early cognitive impairment in multiple sclerosis combining blood and imaging biomarkers. Brain Communications 2022;4:fcac153.
8. Williams T, Tur C, Eshaghi A, et al. Serum neurofilament light and MRI predictors of cognitive decline in patients with secondary progressive multiple sclerosis: Analysis from the MS-STAT randomised controlled trial. Multiple Sclerosis Journal 2022;28:1913-1926.
9. Teunissen CE, Khalil M. Neurofilaments as biomarkers in multiple sclerosis. Multiple Sclerosis Journal 2012;18:552-556.
10. Jakimovski D, Zivadinov R, Ramanthan M, et al. Serum neurofilament light chain level associations with clinical and cognitive performance in multiple sclerosis: A longitudinal retrospective 5-year study. Multiple Sclerosis Journal 2020;26:1670-1681.
11. Ramani S, Berard JA, Walker LA. The relationship between neurofilament light chain and cognition in neurological disorders: A scoping review. Journal of the neurological sciences 2021;420:117229.
12. Gaetani L, Salvadori N, Lisetti V, et al. Cerebrospinal fluid neurofilament light chain tracks cognitive impairment in multiple sclerosis. Journal of neurology 2019;266:2157-2163.
13. Högel H, Rissanen E, Barro C, et al. Serum glial fibrillary acidic protein correlates with multiple sclerosis disease severity. Multiple Sclerosis Journal 2020;26:210-219.
14. Escartin C, Galea E, Lakatos A, et al. Reactive astrocyte nomenclature, definitions, and future directions. Nature Neuroscience 2021;24:312-325.
15. Ayrignac X, Le Bars E, Duflos C, et al. Serum GFAP in multiple sclerosis: correlation with disease type and MRI markers of disease severity. Scientific Reports 2020;10:10923.
16. Abdelhak A, Huss A, Kassubek J, Tumani H, Otto M. Serum GFAP as a biomarker for disease severity in multiple sclerosis. Scientific Reports 2018;8:14798.
17. Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. The Lancet Neurology 2018;17:162-173.
18. Schagen S, Schmand B, Sterke Sd, Lindeboom J. Amsterdam Short-Term Memory Test: A new procedure for the detection of feigned memory deficits. Journal of Clinical and Experimental Neuropsychology 1997;19:43-51.

19. Nauta I, Bertens D, van Dam M, et al. Performance validity in outpatients with multiple sclerosis and cognitive complaints. Multiple Sclerosis Journal;0:13524585211025780.

20. Verhage F. Intelligentie en leeftijd: Onderzoek bij Nederlanders van twaalf tot zevenenzeventig jaar. : Assen: Van Gorcum, 1964.

21. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33:1444-1444.

22. Benedict RH, Cookfair D, Gavett R, et al. Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). Journal of the International Neuropsychological Society: JINS 2006;12:549.

23. Smith A. Symbol digit modalities test (SDMT) manual (revised) Western Psychological Services. Los Angeles 1982.

24. Hammes J. The STROOP color-word test: manual. Amsterdam: Swets and Zeitlinger, 1973.

25. Mulder J, Dekker, R, Dekker, DH. . Verbale Leer- & Geheugen test: Handleiding [Verbal Learning & Memory Test: Manual]. Lisse: Swets & Zeitlinger, 1996.

26. Benedict RH, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. Psychological assessment 1996;8:145.

27. Benton L, Hamsher K, Sivan A. Controlled oral word association test, multilingual aphasia examination. Iowa City, IA: AJA Associates 1994.

28. Parmenter BA, Testa SM, Schretlen DJ, Weinstock-Guttman B, Benedict RH. The utility of regression-based norms in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). Journal of the International Neuropsychological Society 2010;16:6-16.

29. Fischer M, Kunkel A, Bublak P, et al. How reliable is the classification of cognitive impairment across different criteria in early and late stages of multiple sclerosis? Journal of the neurological sciences 2014;343:91-99.

30. Benedict RH, Munschauer F, Linn R, et al. Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. Multiple Sclerosis Journal 2003;9:95-101.

31. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta psychiatrica scandinavica 1983;67:361-370.

32. Vercoulen JH, Swanink CM, Fennis JF, Galama JM, van der Meer JW, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. Journal of psychosomatic research 1994;38:383-392.

33. Soldatos CR, Dikeos DG, Paparrigopoulos TJ. Athens Insomnia Scale: validation of an instrument based on ICD-10 criteria. Journal of psychosomatic research 2000;48:555-560.

34. Chard DT, Jackson JS, Miller DH, Wheeler-Kingshott CA. Reducing the impact of white matter lesions on automated measures of brain gray and white matter volumes. Journal of magnetic resonance imaging 2010;32:223-228.

35. Teunissen C, Petzold A, Bennett JL, et al. A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. Neurology 2009;73:1914-1922.

36. Kuhle J, Barro C, Andreasson U, et al. Comparison of three analytical platforms for quantification of the neurofilament light chain in blood samples: ELISA, electrochemiluminescence immunoassay and Simoa. Clinical Chemistry and Laboratory Medicine (CCLM) 2016;54:1655-1661.

37. Hosmer Jr DW, Lemeshow S, Sturdivant RX. Applied logistic regression: John Wiley & Sons, 2013.

38. Nauta IM, Balk LJ, Sonder JM, et al. The clinical value of the patient-reported multiple sclerosis neuropsychological screening questionnaire. Multiple Sclerosis Journal 2019;25:1543-1546.

39. Aktas O, Renner A, Huss A, et al. Serum neurofilament light chain. No clear relation to cognition and neuropsychiatric symptoms in stable MS 2020;7:e885.

40. Mattioli F, Bellomi F, Stampatori C, et al. Longitudinal serum neurofilament light chain (sNfL) concentration relates to cognitive function in multiple sclerosis patients. Journal of Neurology 2020;267:2245-2251.

41. Disanto G, Barro C, Benkert P, et al. Serum neurofilament light: a biomarker of neuronal damage in multiple sclerosis. Annals of neurology 2017;81:857-870.

42. Jakimovski D, Kuhle J, Ramanathan M, et al. Serum neurofilament light chain levels associations with gray matter pathology: a 5-year longitudinal study. Annals of clinical and translational neurology 2019;6:1757-1770.

43. Fuchs TA, Benedict RH, Bartnik A, et al. Preserved network functional connectivity underlies cognitive reserve in multiple sclerosis. Human brain mapping 2019;40:5231-5241.

44. Ebenau JL, Pelkmans W, Verberk IM, et al. Association of CSF, Plasma, and Imaging Markers of Neurodegeneration With Clinical Progression in People With Subjective Cognitive Decline. Neurology 2022;98:e1315-e1326.

45. Cruz-Gomez ÁJ, Forero L, Lozano-Soto E, et al. Cortical Thickness and Serum NfL Explain Cognitive Dysfunction in Newly Diagnosed Patients With Multiple Sclerosis. Neurology-Neuroimmunology Neuroinflammation 2021;8.

46. Kalb R, Beier M, Benedict RH, et al. Recommendations for cognitive screening and management in multiple sclerosis care. Multiple Sclerosis Journal 2018;24:1665-1680.

SUPPLEMENTARY MATERIALS

Supplemental methods

Imaging protocol

A total of 78 PwMS (~91%) underwent MR scanning on a 3-Tesla whole-body scanner (General Electric Signa-HDxt, Milwaukee, WI, USA), with an 8-channel head coil. The details of the MRI protocol consisted of (1) a 3D-T1 weighted fast spoiled gradient echo sequence for brain volumetry and (2) a 3D fluid-attenuated inversion recovery sequence for white matter lesion detection (see supplemental methods Table 1 for acquisition parameters). Subsequent analyses were performed using FSL5 (http://fsl.fmrib.ox.ac.uk/).

Supplementary Table 1. Information on acquisition parameters of the sequences included in the imaging protocol.

		3D-T1 FSPGR sequence	3D-FLAIR sequence
Parameters			
	Acquisition time	256s	224s
	Repetition time (TR)	8.22ms	8000ms
	Echo time (TE)	3.22ms	128ms
	Inversion time (TI)	450ms	2343ms
	Flip angle	12°	12°
	Orientation	Sagittal	Sagittal
	Voxel size	1.0mm	1.2mm

Abbreviations: FSPGR = Fast Spoiled Gradient Echo; FLAIR = Fluid-Attenuated Inversion Recovery.

Fluid biomarker preprocessing and analysis protocol

After 30 minutes and within two hours, collected blood and CSF (all between 12.00 and 15.00hs) were centrifuged at 1800g for ten minutes and all samples were stored at -20°C for the first 24.00hs and thereafter at -80°C until analysis according to consensus protocols.[1]

Supplemental results

Supplementary Table 2. The performance on individual neuropsychological tests

		Mean (± standard deviation)
Neuropsychological tests		
Stroop		
	Card I	-0.62 (± 1.41)
	Card II	-1.28 (± 1.27)
	Interference	-1.32 (± 1.28)
CVLT-2		
	Direct recall	-0.93 (± 1.16)
	Delayed recall	-0.93 (± 1.46)
	Recognition	-0.53 (± 1.13)
BVMT-R		
	Direct recall	-1.05 (± 1.03)
	Delayed recall	-1.53 (± 1.85)
	Recognition	-0.24 (± 0.94)
SDMT		
	Total score	-1.59 (± 1.11)
COWAT		
	Total score	-0.94 (± 0.73)

Abbreviations: CVLT-2 = California Verbal Learning Test version 2; BVMT-R = Brief Visuospatial Memory Test-Revised; SDMT = Symbol Digit Modalities Test; COWAT = Controlled Oral Word Association Test

Supplementary Table 3. Information on demographics, disease related variables, patient reported outcome measures, imaging markers (in ml) and fluid biomarkers (in pg/ml) displayed for cognitive groups.

	CP (*N* = 33)	**CI (*N* = 40)**	***p*-value**
Demographics			
Sex (female : male)	29 : 4	24 : 16	**.008***
Age	48.09 ± 9.02	47.20 ± 8.06	.658
Educational level	6 [5-6]	6 [5-6]	.972
Clinical functioning			
Disease duration[a]	13.66 ± 9.39	14.12 ± 8.79	.830
EDSS	3.5 [2.5–4.0]	4.0 [3.0–4.5]	**.032***
MS Type (CIS/RRMS/PPMS/SPMS)	(2/24/3/4)	(2/25/1/12)	.146
Use of DMT (yes : no)	17 : 16	18 : 22	.750
Type of DMT (first-line: second-line)	11 : 5	14 : 8	.999
Patient-reported outcome measures			
HADS anxiety	8.31 ± 4.15	8.39 ± 4.79	.940
HADS depression	6.16 ± 3.90	7.62 ± 4.48	.341
CIS20-R (fatigue)	88.39 ± 21.86	94.03 ± 17.94	.240
MSNQ-P (cognitive complaints)	32.11 ± 10.44	33.82 ± 7.74	.465
AIS (sleep-related problems)	6.69 ± 4.83	7.62 ± 4.48	.404
Imaging markers (ml)[b]			
NGMV	805.00 ± 63.95	753.25 ± 56.95	**.002***
NWMV	683.27 ± 46.69	668.58 ± 48.48	.194
NLV[c]	22.14 ± 21.10	35.25 ± 27.85	**.007***
Hippocampi	9.36 ± 1.12	8.55 ± 1.38	.086
Thalami	19.29 ± 2.34	17.13 ± 2.85	**.004***
Fluid biomarkers (pg/ml)			
sNfL[c]	8.45 [5.19–12.67]	10.33 [8.37–13.94]	**.021***
sGFAP[a]	103.37 [79.59–147.88]	124.77 [88.48–181.63]	.071
cNfL[c]	546.99 [336.23–736.64]	579.74 [502.50–1109.46]	.134
cGFAP[a]	7360.16 [5100.24–8887.99]	8039.92 [6818.75–9307.84]	.086

Displayed are the mean and standard deviation of continuous variables, the median and interquartile range of ordinal or non-normally distributed data. Imaging markers and fluid biomarkers were corrected for age and sex (if appropriate) before tested. [a]Variable was square root-transformed before tested. [b]All volumes were normalized using the V-scaling factor. [c]Variable was log-transformed before tested. *Abbreviations*: *CP = cognitively preserved; CI = cognitively impaired; EDSS = Expanded Disability Status Scale; CIS = clinically isolated syndrome; RRMS = relapsing remitting MS; PPMS = primary progressive MS; SPMS = secondary progressive MS; UN = Unknown; DMT = Disease-modifying Therapy; HADS = Hospital Anxiety and Depression Scale; CIS20-R = Checklist Individual Strength 20 – Revised; MSNQ = MS Neuropsychological Questionnaire. AIS = Athens Insomnia Scale; NGMV = normalized grey matter volume; NWMV = normalized white matter volume; NLV = normalized lesion volume; sNfL = serum neurofilament light (NfL); sGFAP = serum glial fibrillary acidic protein (GFAP); cNfL = CSF NfL; cGFAP = CSF GFAP.*

SUPPLEMENTAL REFERENCES

1. Teunissen C, Petzold A, Bennett JL, et al. A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. Neurology 2009;73:1914-1922.