UNIVERSITA' DEGLI STUDI DI VERONA

DEPARTMENT OF NEUROSCIENCES, BIOMEDICINE AND MOVEMENT SCIENCES

GRADUATE SCHOOL OF LIFE AND HEALTH SCIENCES

DOCTORAL PROGRAM IN NEUROSCIENCES, PSYCHOLOGICAL AND PSYCHIATRIC SCIENCES

CYCLE XXXII / 2016

COGNITIVE ALTERATIONS IN MULTIPLE SCLEROSIS PATIENTS: DIAGNOSTIC, PROGNOSTIC, AND REHABILITATION ASPECTS

S.S.D. M-PSI/02

Coordinator: Prof. Gianluigi Zanusso

Supervisor: Prof. Massimiliano Calabrese

Doctoral Student: Stefano Ziccardi

Cognitive alterations in Multiple Sclerosis patients: diagnostic, prognostic, and rehabilitation aspects
Stefano Ziccardi
PhD thesis
Verona, 6 April 2020
ISBN

SOMMARIO

Il deterioramento cognitivo è frequente nella maggior parte dei pazienti affetti da Sclerosi Multipla (SM), riguarda molteplici domini cognitivi e comporta un notevole impatto sulla loro qualità di vita e sui loro aspetti personali, sociali e lavorativi. Una precoce e approfondita valutazione del profilo cognitivo del paziente può avere importanti risvolti diagnostici, prognostici e riabilitativi.

Nel primo capitolo saranno evidenziati gli aspetti diagnostici e prognostici, attraverso la descrizione di un progetto multicentrico, condotto in collaborazione con i centri SM di Bergamo, Montichiari e Modena, nel quale sono stati inclusi pazienti con SM neo-diagnosticati al fine di valutare i loro esiti neurologici, neuropsicologici, neuroradiologici e bioumorali. I risultati di questo progetto hanno permesso la stesura di vari lavori con importanti risultati: il primo studio ha evidenziato come i pazienti con SM al momento della diagnosi, nonostante siano classificati senza un danno cognitivo come tradizionalmente definito in letteratura, mostrano una funzionalità cognitiva ridotta rispetto ad un gruppo di controllo formato da persone neurologicamente sane, sia a livello di funzionamento globale, sia considerando gli specifici domini cognitivi. Il secondo studio ha permesso l'identificazione di due biomarcatori liquorali associati alle alterazioni di funzionamento cognitivo: il primo biomarcatore (LIGHT) è associato alla fase infiammatoria di malattia, mentre il secondo (parvalbumina) è associato anche alla fase neurodegenerativa della malattia e correla con l'assottigliamento della corteccia cerebrale e con la disabilità fisica, oltretutto con una associazione più forte rispetto a quella rilevata con il livello della catena di neurofilamenti leggeri (NF-L, noto marcatore di danno neurodegenerativo). Il terzo studio ha permesso di descrivere il ruolo predittivo di alcune citochine liquorali (CXCL13, CXCL12, IFNγ, TNF, TWEAK, LIGHT, sCD163) nel discriminare, già al momento della diagnosi, i

profili di quei pazienti che dopo 4 anni svilupperanno, con una maggior probabilità, un peggioramento neurologico e neuroradiologico.

Il secondo capitolo affronta l'importanza non solo di effettuare una valutazione neuropsicologica con gli strumenti tradizionali, ma anche di utilizzare protocolli sperimentali. Il primo studio, condotto in collaborazione con l'Università di Firenze e con l'Università di Padova, ha indagato il fenomeno delle false memorie mediante l'utilizzo di un paradigma che induce distorsioni mnestiche a causa della forte connessione tra parole associate ad una stessa categoria semantica. I risultati dimostrano che i pazienti con SM non mostrano la distorsione mnestica attesa, probabilmente per la scarsa associazione tra i nodi che costituiscono la memoria semantica, soggetta a fenomeni neurodegenerativi. Il secondo studio, condotto in collaborazione con la Kessler Foundation (West Orange, Stati Uniti), si è focalizzato sulle abilità di cognizione sociale: in un gruppo di pazienti SM senza danno cognitivo come classicamente definito è emersa una prestazione significativamente inferiore rispetto a controlli neurologicamente sani in test di riconoscimento delle emozioni facciali, di teoria della mente e di empatia. Inoltre, è stato dimostrato che queste alterazioni di cognizione sociale sono correlate in modo specifico al volume delle lesioni corticali nelle amigdale dei pazienti, rispetto ad altre misure di danno cerebrale incluse nello studio (assottigliamento corticale e carico lesionale di tutta la corteccia cerebrale).

Il terzo e ultimo capitolo, infine, si concentra sugli aspetti riabilitativi, mostrando i risultati di uno studio condotto presso il Buffalo Neuroimaging Analysis Center (Buffalo, Stati Uniti) su un gruppo di pazienti con SM che hanno svolto un training cognitivo in regime di teleriabilitazione. L'obiettivo del progetto era identificare quali variabili neurologiche, psicologiche e neuroradiologiche caratterizzassero i pazienti con un maggior miglioramento cognitivo dopo il trattamento riabilitativo. I risultati hanno mostrato come una forma di malattia recidivante-remittente (rispetto ai pazienti con forme progressive), una personalità caratterizzata da elevata

coscienziosità, un maggior volume della sostanza grigia cerebrale, una ridotta compromissione dei tratti di sostanza bianca in un network cerebrale centrato nel precuneo e nel cingolato posteriore, e un maggior scostamento di connettività funzionale rispetto a quella dei controlli sani, siano tutti elementi risultati associati a un maggior effetto positivo della riabilitazione cognitiva effettuata.

ABSTRACT

Cognitive impairment is frequent in most patients with Multiple Sclerosis (MS) and affects several cognitive domains, having a significant impact on their quality of life and on their personal, social and work dimensions. An early and comprehensive neuropsychological assessment may provide relevant diagnostic, prognostic, and rehabilitative implications.

The first chapter highlights the diagnostic and the prognostic aspects, with the description of a multicentric project, conducted in collaboration with MS centers of Bergamo, Montichiari, and Modena, in which were included newly-diagnosed MS patients and were evaluated their neurological, neuropsychological, neuroradiological and bioumoral outcomes. Results of this project have allowed the preparation of several sub-studies with important results: the first study highlighted how MS patients at the time of diagnosis, even in the absence of an evident cognitive impairment as clinically defined, are characterized by slight cognitive alterations as compared to healthy controls, both considering global cognitive functioning level and also specific cognitive domains. The second study has allowed the identification of two biomarkers present in the cerebrospinal fluid that are associated with cognitive alterations: the first (LIGHT) is associated with the inflammatory phase of the disease, while the second (parvalbumin) is associated with the neurodegenerative phase of the disease and also correlates with cortical thinning and physical disability, moreover with a stronger association compared to the one found with the level of neurofilament light chain (NF-L, a well-known biomarker of neurodegeneration). The third study has allowed to describe the predictive role of some inflammatory cytokines in the cerebrospinal fluid (CXCL13, CXCL12, IFNγ, TNF, TWEAK, LIGHT, sCD163) in discriminating, since the time of diagnosis, those MS patients that were more likely to develop neurologic and neuroradiologic worsening after 4-years follow-up.

The second chapter addresses the importance of assessing MS patients not only with the classical neuropsychological tests but also with experimental paradigms. The first study, conducted in collaboration with the University of Florence and the University of Padua, investigated the phenomena of false memories, using a paradigm that induces memory distortions due to the strong connection between words associated with a same semantic category. Results showed that MS patients were not characterized by the expected memory distortions, probably due to weak association between nodes that compose semantic memory, because of neurodegenerative events. The second study, conducted in collaboration with the Kessler Foundation (West Orange, NJ, USA), focused on social cognition abilities: in a group of MS patients without evidence of cognitive impairment as traditionally defined was observed a performance significantly lower compared to healthy controls in tests of facial emotion recognition, theory of mind, and empathy. Moreover, it was demonstrated that these social cognition alterations were correlated specifically with the cortical lesions volume in both the amygdalae of MS patients, while no significant correlation was found with other measures of brain damage included in the study (cortical thickness and cortical lesion load in all the cerebral cortex).

The third and last chapter focuses on the rehabilitative aspects, showing results from a study carried at the Buffalo Neuroimaging Analysis Center (Buffalo, NY, USA) on a group of MS patients that performed a cognitive training by using a telerehabilitation approach. The project aimed to identify neurological, psychological and neuroradiological variables able to characterize patients that can benefit more from the rehabilitation. Results showed that a relapsing-remitting disease phenotype (as compared with progressive patients), a higher personality trait of conscientiousness, a higher gray matter volume, a lower tract disruption in a network centered

on precuneus and posterior cingulate, and a higher deviation in functional brain connectivity compared to healthy controls, play a key role to achieve a greater cognitive amelioration after the rehabilitative treatment.

TABLE OF CONTENTS

GENERAL INTRODUCTION

MULTIPLE SCLEROSIS

Multiple sclerosis (MS) is a chronic inflammatory neurodegenerative disease of the central nervous system (CNS) (Sospedra & Martin, 2005), which represents the most common non-traumatic cause of disability in young adults (Browne et al., 2014). Its pathogenesis is complex: the involvement of genetic and environmental factors has been demonstrated (Handel et al., 2010; Sawcer et al., 2011;), however it is not yet fully understood (Olsson, Barcellos, & Alfredsson, 2017).

MS is characterized by the attack of the immune system to the myelin sheath, leading to accumulation of chronic white matter (WM) demyelination, axonal injury, and diffuse inflammatory changes in the normal-appearing WM (Noseworthy, Lucchinetti, Rodriguez, & Weinshenker, 2000; Graumann et al., 2003; Howell et al., 2010; Kutzelnigg & Lassmann, 2014). Moreover, both focal and diffuse cortical gray matter (GM) demyelination and neurodegeneration is typical in MS (Kutzelnigg et al. 2005; Wegner et al. 2006; Chard & Miller 2009; Frischer et al. 2009; Geurts et al. 2009; Howell et al. 2011) and can occur early in the disease and worsen over time, accounting for the accumulation of severe disability (Magliozzi et al., 2010; Calabrese et al., 2015a; Scalfari et al., 2018). In particular, cortical lesions and GM atrophy are the major drivers of physical disability (Rovaris et al. 2000; Fisniku et al. 2008; Filippi et al. 2013;; Zivadinov et al. 2013; Zivadinov et al. 2016b) having a predominant role since the earliest stages of the disease, (De Stefano et al. 2003; Lucchinetti et al. 2011) and accumulate in particular in the progressive phase (Calabrese et al., 2015a; Rojas et al., 2016) contributing to the increase of neurodegeneration and disease progression (Calabrese et al., 2012; Geurts et al., 2012). Brain-imaging studies have confirmed that cortical thinning, as an early event in MS pathogenesis, occurs even in mildly disabled patients (Calabrese et al. 2007, Charil et al. 2007) and prominently in areas of the

brain that have extensive cortico-cortical connections, such as the cingulate gyrus and the insular, frontal, temporal, and parietal cortices (Charil et al. 2007).

MS pathology is characterized by both GM and WM involvement, with GM damage more predominant in rapidly progressive MS compared to WM damage (Howell et al., 2011). GM lesions are associated with meningeal infiltrates in the subarachnoid space, while WM lesions are associated with perivascular infiltrates (Michel et al., 2015). Perivascular B-cell and T-cell infiltrates have also been found in both intracortical and subpial lesions, especially in patients with progressive MS (Serafini et al., 2004). Meningeal inflammation may be associated with cortical demyelination, influencing cerebrospinal fluid (CSF) inflammation (Calabrese et al., 2015a). Increased levels of pro-inflammatory cytokines and chemokines, such as CXCL13, CXCL12 (Krumbholz et al. 2006), TNF (Sharief & Hentges 1991; Baraczka et al. 2004), and IFNγ (Duan et al. 2013) have been found in the CSF of MS patients.

MS affects approximately 2.5 million people worldwide (Browne et al., 2014) and is more common in females compared to males (2:1; Noseworthy et al., 2000; Orton et al., 2006), even though in males were described more severe prognoses (Kantarci & Wingerchuk, 2006), while disease onset occurs typically between 18 and 40 years (Calabresi, 2004). This age period usually overlaps with the most productive period for an individual and is characterized by personal future planning in terms of work, family and life in general: therefore, to receive the diagnosis in this life period may lead to critical psychological difficulties for patients, in particular regarding the acceptance of the disease (Macias Islas & Ciampi, 2019). Despite life expectancy of patients with MS is decreased by only an average of 10 years compared with general population, MS could be associated with physical disability and with almost three times in the risk of death (Weinshenker, Issa & Baskerville, 1996; Brønnum-Hansen et al., 2006).

Different phenotypes of MS have been proposed, based on the progression of the disease and the clinical course (Lublin & Reingold, 1996): Relapsing-Remitting MS (RRMS), Secondary-Progressive MS (SPMS), Primary-Progressive MS (PPMS), and Relapsing-Progressive MS (RPMS). RRMS refers on average to the 85-90% of the patients and is characterized by clearly defined attacks (relapses) with complete or partial recovery, with no evidence of clinical progression between intercritical periods; SPMS is characterized by a gradual worsening after an initial relapsing disease course; PPMS is characterized by a continuous worsening since the onset of the disease; RPMS is characterized by a progressive accumulation of disability from onset but defined acute clinical relapses with or without full recovery. Recently, this classification has been reviewed by Lublin et al. (2014). The clinical courses have been grouped in relapsing phenotypes or progressive phenotypes and there have been described new clinical courses: clinically isolated syndrome (CIS), characterized by the first clinical presentation of the disease suggesting MS but not enough to fulfil the diagnostic criteria, and radiologically isolated syndrome (RIS), characterized by the first imaging finding suggestive of MS but in absence of clinical signs or symptoms (Lublin, 2014).

Several disease-modified treatments (DMTs) are used trying to slow MS progression. Currently approved DMTs are mainly immunomodulatory or immunosuppressive drugs that showed a significant (although variable) effect in reducing the frequency of attacks in the relapsing forms of the disease and delaying the disability progression; however, actual DMTs have limited effect in preventing the transition to the progressive form of the disease (Weinstock-Guttman, 2013). Clinicians are called to the challenging task of choosing the best suited DMTs for each patients, selecting on the basis of the clinical status of each patient either a first-line treatment (moderate efficacy but high safety) or a second-line treatment (more effective but also with more safety risks) (Gajofatto & Benedetti, 2015).

Clinical manifestations of MS include various episodes of neurological dysfunction, that differ considerably among patients, depending on the CNS site involved. Common clinical presentations are optic neuritis, transverse myelitis, brainstem syndromes, motor symptoms, sensitive involvement, imbalance, bladder and intestinal dysfunction, heat sensitivity, and epileptic attacks (Gelfand, 2014). Physical disability is typically evaluated by using the Expanded Disability Status Scale (EDSS; Kurtzke, 1983), a scale ranging from 0 to 10 that quantifies the accumulation of physical impairment. Moreover, many other clinical manifestations may occur in MS patients: emotional disturbances (depression, anxiety, stress), both physical and mental fatigue symptoms, and cognitive impairment (Kos et al., 2008; Feinsten, 2011; Calabrese & Pitteri, 2018).

COGNITIVE IMPAIRMENT IN MULTIPLE SCLEROSIS

In addition to physical disability, cognitive impairment (CI) is common in MS patients (Amato, Zipoli & Portaccio, 2006; Grzegorski & Losy, 2017) with frequencies ranging from 43% to 70% (Chiaravalloti & DeLuca, 2008), depending on the studied population, the tests used, and the applied cut-off scores (Fischer et al., 2014; Amato et al., 2018). CI can alter MS patients' behavior and quality of life (Rao, 2004; Amato, Zipoli, & Portaccio, 2006), leading to dysfunctional social and personal difficulties, despite minimal physical disability (Benedict & Zivadinov, 2011).

The main affected cognitive domains are verbal learning and memory, attention, information processing speed, and executive functions (Benedict et al., 2006); however, alterations in specific domains may lead to global cognitive impairment. To assess cognitive functioning of MS patients, some batteries of neuropsychological tests have been proposed, i.e. the Brief Repeatable Battery of Neuropsychological Test (BRB-NT; Rao, 1990), the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS;

Benedict et al., 2006) and the Brief International Cognitive Assessment for MS (BICAMS; Benedict et al., 2012).

CI can occur since the earliest stages of the disease course (Amato et al., 2010; Calabrese et al., 2015b) and has been strongly associated with both focal and diffuse GM damage (Calabrese et al., 2009; Eijlers et al., 2018) and WM lesion measures (Rocca et al., 2015; Preziosa et al., 2016). It has been demonstrated an association between early frontotemporal neurodegeneration in patients with mild CI, while patients with severe CI are characterized by more widespread neurodegeneration phenomena (Calabrese et al., 2010).

For these reasons, assessing CI since the early phase of the disease is of paramount importance (Kalb et al., 2018; Sumowski et al., 2018). However, cognitive functioning assessment is usually not included in the standard clinical evaluation of MS patients, as well as cognitive outcomes are not included in all MS clinical trials (Macias Islas & Ciampi, 2019). As a consequence, cognitive alterations could be dramatically neglected with tremendous implications in clinical monitoring of disease progression and in quality of life of MS patients and their caregivers. Moreover, cognition is related to the functioning of several neural pathways and connections that are involved in processing information in the brain: since all neurological diseases, independently from their predominant inflammatory or neurodegenerative main processes, are characterized by CNS damage and neuronal suffering in several brain structures (both at cortical and subcortical levels), this will reflect on cognitive alterations. A comprehensive neuropsychological assessment should be then provided to all neurological patients due to its importance in helping to define a better diagnostic patient's profile and its contribution as prognostic factor of long-term disease severity.

Firstly, cognitive functioning assessment leads to important implications in terms of diagnosis, in order to identify different levels of

cognitive impairment. During assessing cognitive functioning, it is important not to focus only on traditional neuropsychological batteries but also utilize experimental paradigms, because a more detailed investigation might better elucidate specific cognitive alterations. All patients, independently from their baseline cognitive functioning, should undergo follow-up neuropsychological assessment to be compared with the baseline measure: monitoring cognitive functioning allows to monitor indirectly disease progression and, consequently, brain damage phenomena.

Secondly, the cognitive functioning level might also have a prognostic value for disease and disability progression, as also for future accumulation of cortical and subcortical brain damage. Vice versa, neurologic, neuroradiological, and bioumoral outcomes at baseline might be used as prognostic markers for the development of cognitive impairment years later.

Thirdly, cognitive impairment can be thwarted by planning specific cognitive training and/or neuropsychological rehabilitation sessions. It is important to provide cognitive rehabilitation to the patients without waiting for their cognitive impairment to become severe, but otherwise to intervene as soon as possible. The early assessment of cognitive functioning displays which cognitive domain should be targeted during the training and offers a baseline status that can be used as a comparison for pre- and post-intervention observation, in order to measure the progress earned through cognitive rehabilitation sessions.

PROJECT 1 – LONGITUDINAL MULTI-PARAMETRIC CLINICAL, COGNITIVE, CSF AND MRI STUDY AIMED AT PREDICTING DISEASE PROGRESSION IN MULTIPLE SCLEROSIS

The diagnostic process of patients with suspected MS patients is a core event, not free from psychological consequences. Following the first episode of the disease, that could be either neurologic or neuroradiological, the diagnostic purpose is to identify indisputably the presence of MS. Diagnostic criteria has been proposed: most updated revisions of diagnostic criteria were recently completed (Thompson et al., 2018) and considered together clinical ("no better explanation" than MS), neuroradiological (lesions dissemination in space and time) and biological (presence of CSF-specific oligoclonal bands OCBs) variables. Magliozzi et al. (2018) showed an association between MRI and CSF outcomes at the time of MS diagnosis: an intrathecal profile in CSF characterized by a regulatory pattern of chemokines (IFN-α2, IFN-λ2, CCL25) is typical of MS patients with low cortical damage (low number and volume of cortical lesions) and low meningeal inflammation, while a more severe inflammatory CSF profile (CXCL13, TNF, IFNγ, CXCL12, IL6, IL10, BAFF, APRIL, LIGHT, TWEAK, IL8, MMP-2, Pentraxin3, sCD163 and NF-L) is common in MS patients with a higher level of cortical damage (high number and volume of cortical lesions) and high meningeal inflammation.

Considering the remarkable social and health efforts that have been made for challenging MS, both in terms of therapeutic and diagnostic outcomes (Pozzilli, Romano & Cannoni, 2002; Patwardhan et al., 2005), a high-specificity and comprehensive diagnostic process is important to provide clinicians as many elements as possible to understand the disease, formulate the correct diagnosis, and select the best disease-modifying treatment (DMT) for each patient. However, particularly the choice of the

pharmacological treatment, based on the patient's expected clinical outcome, remains a challenging task.

Moreover, so far it was impossible to predict when, how rapidly, and how severely the disease will progress, in terms of accumulation of physical and cognitive disability, and also to predict the possible transition to a secondary progressive MS phase; however, being able to know the disease progression in advance is of fundamental importance for the patients, considering the dramatic impact that MS evolution has on psychological status and quality of life of both patients and caregivers. The possible prognostic value of variables early collected might improve the ability to know in advance how the disease will progress, in order to be helpful to the clinician for the choice of the best suited DMT for each patient.

Neuroimaging data have shown that the extent of GM damage is a good predictor of long-term disability (Calabrese et al., 2012; Bergsland et al., 2018; Eshaghi et al., 2018). However, the non-conventional MRI analyses required to detect GM demyelination and brain atrophy are not straightforward, their assessment still needs to be standardized and their reliability has to be confirmed (Filippi, Preziosa, & Rocca, 2018).

Moreover, also CSF outcomes have a key role as biomarkers to predict future disease progression: our research group showed that the presence of OCBs at the time of diagnosis predicted a more severe outcome 10 years later, in particular the association was found with more severe GM pathology, physical disability, and cognitive impairment (Farina et al., 2017). Several studies about bioumoral markers have revealed the differential expression of CSF proteins in MS compared to controls. Furthermore, level of neurofilament light chain (NF-L) are also well-known and widely used as a promising fluid biomarker for neurodegenerative processes in neurological diseases (Jakimovski et al., 2019): NF-L is part of axonal cytoskeleton that, following neuronal deaths, is released in the CSF. In MS patients, neuronal suffering measured with the level of NF-L in the

CSF is associated with a greater risk of conversion to clinically definite MS (Fialová et al. 2013), increased brain and spinal cord atrophy (Petzold et al. 2016), as well as cognitive impairment (Quintana et al., 2018; Gaetani et al., 2019; Kalatha et al., 2019).

Cognitive functioning itself might be also a helpful prognostic marker. Longitudinal studies have shown that CI detected at the time of diagnosis can predict the conversion from clinically isolated syndrome to definite MS (Zipoli et al., 2010), the progression of physical disability (Deloire et al., 2010), and the transition to the secondary progressive phase (Moccia et al., 2016). A recent publication from our group showed how early CI is helpful in the prognosis of MS patients at high risk of conversion from CIS to definite MS, as well as progression of disability, transition to SPMS and especially brain cortical atrophy 8-years later (Pitteri et al., 2017).

Considering the importance and the high potential of the predictive role of the variables abovementioned, there is a great need for additional biomarkers that could be identified early in the disease course to classify patients destined to have more severe outcomes. There would be a distinct therapeutic advantage in being able to predict which profiles of MS patients are likely to progress earlier and more rapidly, in order to appropriately manage the most suitable therapeutic intervention since the time of diagnosis. For this reason, a research project titled "*Longitudinal multi-parametric clinical, CSF, and MRI study aimed at predicting disease progression in Multiple Sclerosis*" was conducted. In particular, my doctoral project was focused on identifying the cognitive profile of newly-diagnosed MS patients and its association with MRI and CSF data.

Project design

This multi-centric observational study was conducted in collaboration with 3 other Italian MS centers: Bergamo, Montichiari, and Modena. Patients included in the project were selected among possible relapsing-remitting MS

patients (RRMS) about to undergo examinations for diagnostic purposes, according to ordinary clinical practice. At baseline (T0), each patient underwent the following examinations:

- CSF withdrawal from lumbar puncture (LP);
- Blood withdrawal;
- 3T MRI with advanced imaging sequences;
- Neurologic examination;
- Comprehensive neuropsychological assessment with an extensive battery of tests.

Once concluded the diagnostic pathway, the patients who received the diagnosis of MS underwent the same set of examinations also at T1 (1-year follow-up) and at T2 (2-years follow-up). In Figure 1 are provided detailed information about the assessment pipeline.

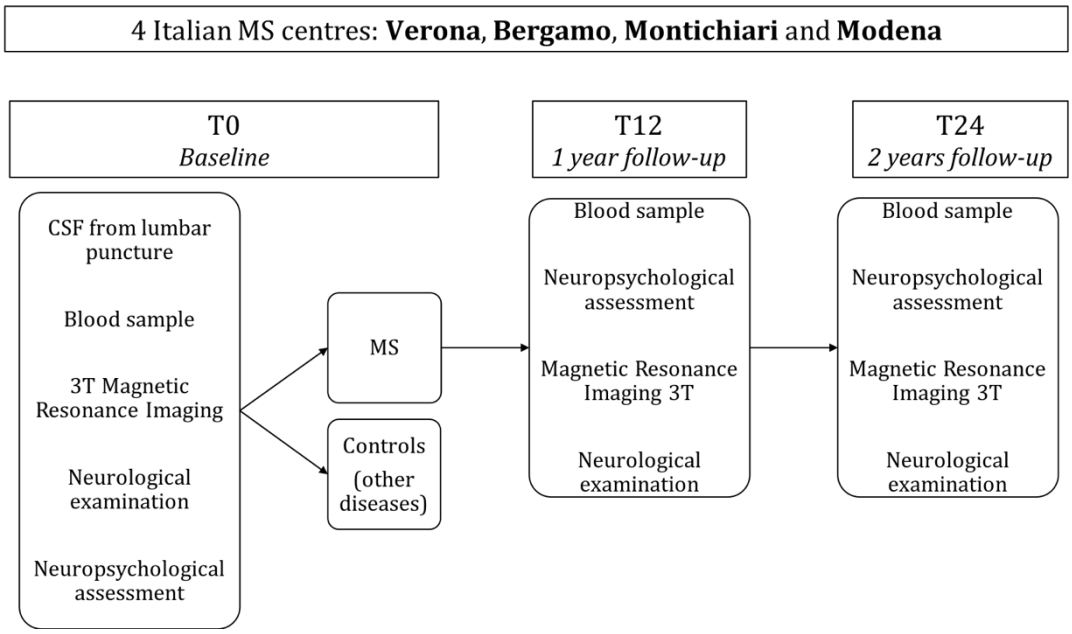

Figure 1. Assessment pipeline of the project.

At the end of September 2019, 165 patients were overall recruited: 126 (76%) from Verona, 15 (9%) from Bergamo, 8 (5%) from Montichiari and 16 (10%) from Modena. Figure 2 shows a pie chart about patients'

recruitment and Table 1 shows demographic and clinical characteristics of the patients recruited.

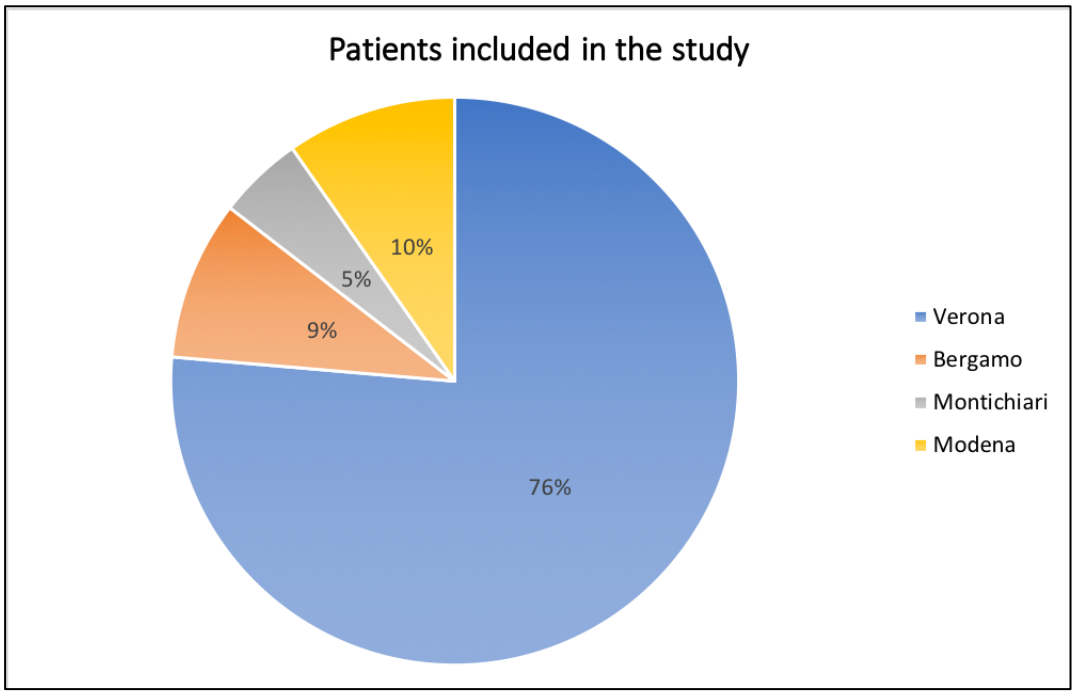

Figure 2. Pie chart about patients recruited during the project.

	Whole sample (=165)
Gender (M/F)	50/115
Age (years)	37.9±12.1
MS type (RMS vs. PMS)	154/11
EDSS	1.5 (0-6)
Time from onset to LP (years)	2.1±3.5

Table 1. Demographic and clinical characteristics of MS patients recruited. Continuous variables are provided with Mean±SD. EDSS is provided with Median (Range). Age, EDSS, and disease duration are provided at the time of lumbar puncture. MS type is showed comparing patients with a relapsing type of MS (CIS, RRMS, and SPMS) vs patients with a progressive type of MS (PPMS).
MS = Multiple Sclerosis; RMS = Relapsing Multiple Sclerosis; PMS = Progressive Multiple Sclerosis; EDSS = Expanded Disability Status Scale; LP = Lumbar Puncture; CIS = Clinically Isolated Syndrome; RRMS = Relapsing Remitting Multiple Sclerosis; SPMS = Secondary Progressive Multiple Sclerosis; PPMS = Primary Progressive Multiple Sclerosis

Inclusion criteria comprised patients of age between 18 and 65 years that, after the diagnostic pathway, received the diagnosis of MS (according to Polman et al., 2010), with no other concomitant neurological or other pathological health conditions, and with no substance abuse or other concomitant medications. The local Ethics Committee approved the project and informed consent was obtained from all the patients.

1.1 COGNITIVE FUNCTIONING AT THE TIME OF DIAGNOSIS IN MS PATIENTS

Despite the different batteries of neuropsychological tests used to assess cognitive performance, the classification of CI is undoubtedly affected by the chosen cut-off applied (Binder, Iverson, & Brooks, 2009; Fischer et al., 2014) and by the number of neuropsychological tests administered. Usually, MS patients are classified either as having “normal cognition” or “impaired cognition” in a perspective of dichotomous classification (unimpaired vs. impaired). This approach, however, is far from being meaningful considering the real-life (Amato et al., 2018), in which measures of functional aspects, such as cognitive functioning, resemble continuous variables, as also underlined in other neurological populations (e.g., Azouvi, Jacquin-Courtois & Luaute, 2017; Pitteri et al., 2018).

Dichotomizing continuous variables, such as cognitive functions, carries the risk of losing information that might increase the number of false-positive results, as well as the risk to reduce power of statistical analysis and to underestimate the extent of variation in performance (Altman & Royston, 2006). For this reason, it would be more appropriate to use psychometric methods switching from a “cognitive impairment-based” to a “cognitive functioning-based” approach, considering cognitive functioning as a continuum variable as it is in real-life, ranging from a minimum to a maximum performance. This is of particular interest given that cognitive decline may develop as a result of gradual progression related to neurodegeneration and brain atrophy, or of acute disease activity, for which decline in cognitive performance can be often followed by incomplete recovery, thus contributing to the burden of cognitive impairment in the long term (Rocca et al., 2015).

In order to investigate the usefulness of this approach, the cognitive performance of a group of newly-diagnosed MS patients with relapsing-remitting course were assessed and compared to a group of healthy controls

(HCs). It is expected that also newly-diagnosed MS patients, classified as being “cognitively normal” with the traditional approach, would instead show reduced cognitive functioning with respect to HCs.

Materials and methods

Participants

Fifty newly-diagnosed RRMS patients (37 females, mean±SD age=38.2±11.6; mean±SD education=14.2±2.7; mean±SD disease duration=3.5±5.2; median[range] EDSS=1.5 [0-4]) were tested with an extensive battery of neuropsychological tests close to the time of MS diagnosis (average of 6 months). At the time of neuropsychological testing, 31 RRMS patients were untreated, whereas 14 were treated with dymethilfumarate, 1 with fingolimod, 1 with natalizumab, 1 with interferon beta1-a, 1 with peg-interferon beta1-a, and 1 with azathioprine. Inclusion criteria for MS patients comprised diagnosis of RRMS (Polman et al., 2011), no relapse or steroid treatment within 30 days before the neuropsychological assessment, no concomitant neurological or other pathological health conditions, no substance abuse or other MS concomitant medication (as benzodiazepines or antidepressant drugs), no visual impairment.

A group of 36 matched healthy controls (HCs) was recruited and tested with the same battery of neuropsychological tests used to assess RRMS patients. Inclusion criteria for HCs comprised no cognitive deficits measured with the Montreal Cognitive Assessment test (MoCA; Santangelo et al., 2015), no neurologic, psychiatric, or other concomitant pathologies; normal or corrected to normal vision; no substance abuse or other concomitant medications.

All participants were recruited at the MS Center of the Verona University Hospital (Verona, Italy). The study was approved by the local Ethics Committee and informed consent was collected from all participants.

Neuropsychological assessment

RRMS patients and HCs were tested with an extensive battery of neuropsychological tests, which included the Brief Repeatable Battery of Neuropsychological Test (BRB-NT; Amato et al., 2006a), the Stroop Test (Caffarra et al., 2002), the Phonological, Semantic, and Alternate Verbal Fluency tests (VF; Costa et al., 2014), and the Modified Five Point Test (MFPT; Cattelani et al., 2011). Depression, anxiety, and stress were evaluated with the Depression Anxiety Stress Scale 21-items (DASS-21; Bottesi et al., 2015) and subjective fatigue with the Fatigue Severity Scale (FSS; Krupp et al., 1989). According to the most used method for classifying CI (Amato et al., 2006b), RRMS patients were classified as "cognitive normal" (CN) if they scored below the cut-off (5^{th} percentile; z-score=-1.65) on 0, 1, or 2 neuropsychological tests; otherwise (i.e., score below the cut-off on more than 2 neuropsychological tests), the patients were classified as having CI.

For each neuropsychological test and for each RRMS patient and HCs, the Z-score index was calculated in which the normative data of the Italian validation of each test were not used but, rather, the mean and the standard deviation (SD) of scores of both the RRMS and HCs *together*. Considering the mean and SD of both groups together, in which MS patients and HCs composed the same population, allowed to normalize the dependent variable (Z-score index) in a unique gaussian distribution with overlapped curves, mimicking a more real-life condition. With this procedure were calculated: 1) a global cognitive functioning index (Z-global) considering the average of the Z-score of each neuropsychological test; and 2) three domain-specific Z-score indexes: memory (Z-MEM), attention/information processing speed (Z-ATT/IPS), and executive functions (Z-EF). Detailed classification of cognitive domains are provided in Table 2.

Z-MEM	Z-ATT/IPS	Z-EF
SRT-LTS	SDMT	ST (average EIT and EIE)
SRT-CLTR	PASAT-3	Phonemic VF
SRT-D	PASAT-2	Alternate VF
SPART-I		MFPT-UDs
SPART-D		MFPT-CSs

Table 2. Neuropsychological tests considered for each Z-score domain index.
Z-MEM = Z-score – Memory; Z-ATT/IPS = Z-score – Attention/Information Processing Speed; Z-EF = Z-score – Executive Functions; SRT-LTS = Selective Reminding Test – Long Term Storage; SRT-CLTR = Selective Reminding Test – Consistent Long Term Retrieval; SRT-D = Selective Reminding Test – Delayed; SPART-I = Spatial Recall Test – Immediate; SPART-D = Spatial Recall Test – Delayed; SDMT = Symbol Digit Modalities Test; PASAT = Paced Auditory Serial Addition Test; ST - EIT = Stroop Test - Effect Interference Time; ST - EIE = Stroop Test - Effect Interference Error; VF = Verbal Fluency; MFPT-UDs = Modified Five Point Test – Unique Designs; MFPT-CSs = Modified Five Point Test – Cumulative Strategies

Statistical analysis

ANOVA models with Tukey post-hoc analysis and chi-square test were applied to evaluate the difference of demographic, clinical, and Z-index scores among CI-RRMS, CN-RRMS and HCs. Effects of physical disability (EDSS), disease duration, emotional state (DASS-21), and subjective fatigue (FSS) on the global cognitive functioning (Z-global) and on the three cognitive domains (Z-MEM, Z-ATT/IPS, Z-EF) were controlled for RRMS patients' data by using a stepwise multiple regression analysis.

Results

Among the 50 RRMS patients tested, 14 patients were classified as having CI, whereas 36 as being cognitive normal (CN). The majority (12/14: 86%) of the CI-RRMS patients were impaired in the domain of attention/IPS (64%) and EFs (71%). Demographic and clinical characteristics of RRMS and HCs are provided in Table 3.

	CI-RRMS (=14)	CN-RRMS (=36)	HCs (=36)	*p*
Gender (M/F)	3/11	10/26	13/23	.547
Age (years)	39.3±14.0	37.8±10.8	33.6±10.4	.170
Education (years)	13.8±4.0	14.4±2.0	15.1±2.6	.229
EDSS	2.0 (0–4)	1.0 (0-3)	/	/
Disease duration (years)	4.4±8.2	3.1±3.6	/	/
Time between diagnosis and neuropsychological assessment (months)	6 (±3)	6 (±2)	/	/

Table 3. Demographic and clinical characteristics of RRMS patients and HCs. Mean±SD was provided for continuous variables. Median (range) was provided for EDSS.
CI-RRMS = cognitive impaired – relapsing remitting multiple sclerosis patients; CN-RRMS cognitive normal – relapsing remitting multiple sclerosis patients; HCs = healthy controls; EDSS = Expanded Disability Status Scale

Group comparison results between CI (n=14), CN (n=36), and HCs (n=36) showed no significant difference among the 3 groups in terms of age (p=0.170), education (p=0.229), and gender (p=0.547).

The Z-score global index was significantly different among the three groups (p<0.001). Post-hoc analysis showed a significant difference between CI and HCs (p<0.001), between CI and CN (p<0.001), and also between CN and HCs (p=0.004) - Figure 3.

Significant difference has been found among the three groups also in the Z-MEM (p<0.001), Z-ATT/IPS (p<0.001), and Z-EF (p<0.001). Post-hoc analysis showed a significant difference between CI and HCs on Z- MEM (p<0.001), Z-ATT/IPS (p<0.001), and Z-EF (p<0.001); between CI and CN on Z-MEM (p=0.009), Z-ATT/IPS (p<0.001), and Z-EF (p<0.001); and between CN and HCs on Z-MEM (p=0.005) and Z-EF (p=0.006). No significant difference was found between CN and HCs on Z-ATT/IPS (p=0.087) – Figure 3.

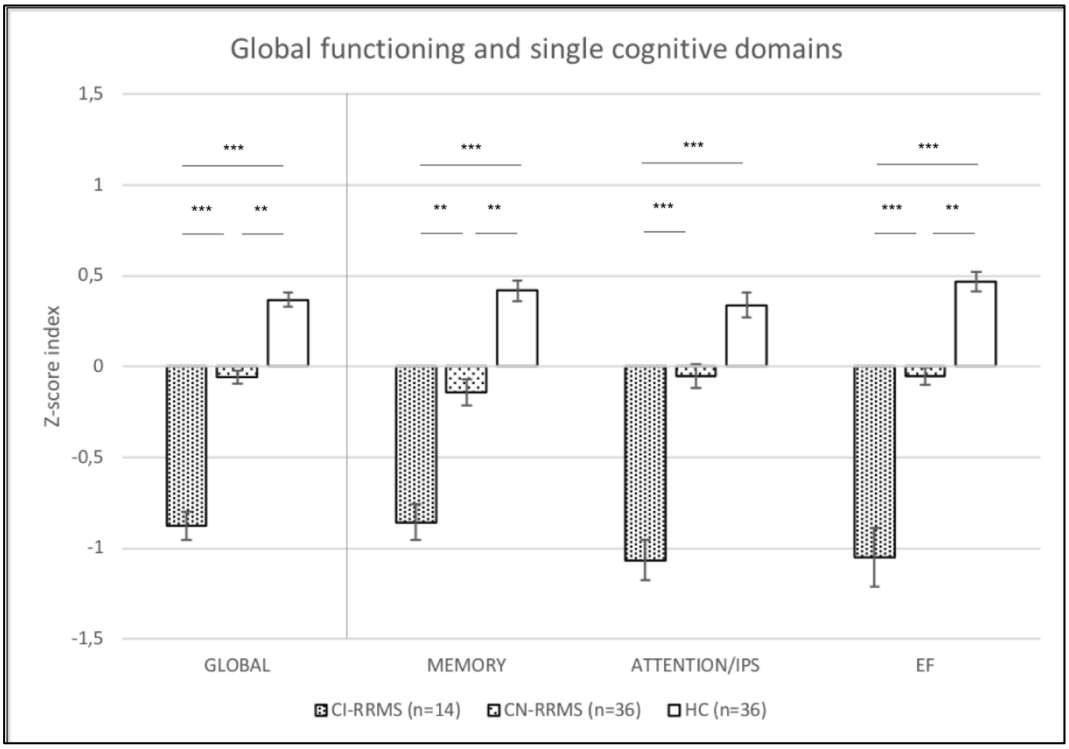

Figure 3. Comparison of the Z-global index and the Z-score indexes of memory (Z-MEM), attention/processing speed (Z-ATT/IPS), and executive functions (Z-EF) between CI-RRMS, CN-RRMS and HCs.
**: $p < 0.01$, ***: $p < 0.001$.
CI-RRMS = Cognitive Impaired Relapsing Remitting Multiple Sclerosis; CN-RRMS = Cognitive Normal Relapsing Remitting Multiple Sclerosis; HCs = Healthy Controls

Considering CI and CN patients together, the results of the multiple regression analysis (final model R^2=.254, p=0.170) showed no significant effects of age (β =-.255, p=0.155), education (β=.196, p=0.240), gender (β =.224, p=0.185), disability (β=-.125, p=0.437), disease duration (β =.085, p=0.591), emotional state (β=-.197, p=0.288), and fatigue (β=-.145, p=0.477) on the Z-global index. No significant effect was also found on Z-MEM (R^2=.289, p=0.099), Z-ATT/IPS (R^2=.208, p=0.311), and Z-EF (R^2=.252, p=0.173).

Considering each single neuropsychological test, a significant difference was found among the three groups (CI, CN, and HCs) in all the neuropsychological tests (all p<0.05), except for the WLG test (p=0.180). Post-hoc analysis showed a significant difference between CI and HCs in all

neuropsychological tests (all $p<0.05$). Moreover, a significant difference was found between CI and CN on the SRT-CLTR ($p=0.013$), SRT-D ($p=0.026$), SDMT ($p=0.040$), PASAT-3 ($p=0.001$), PASAT-2 ($p=0.043$), ST-EIT ($p=0.049$), ST-EIE ($p<0.001$), Phonemic Verbal Fluency ($p=0.023$), Semantic Verbal Fluency ($p=0.042$), MFPT-UDs ($p=0.001$), and MFPT-Error Index ($p=0.020$) – Table 4. Finally, comparing CN and HCs, a significant difference was found in the SRT-LTS ($p=0.012$), SRT-CLTR ($p=0.016$), SRT-D ($p=0.007$), SDMT ($p=0.014$), Phonemic Verbal Fluency ($p=0.034$), MFPT-UDs ($p=0.003$), and MFPT-CSs ($p=.019$) - Table 4.

		CI-RRMS (n=14)		**CN-RRMS (n=36)**		**HCs (n=36)**			
NP battery / NP test	**Subtest**	**Raw Scores**	***Z-score index***	**Raw Scores**	***Z-score index***	**Raw Scores**	***Z-score index***	***p* CI vs CN**	***p* CN vs HCs**
Brief Repeatable Battery of Neuropsychological Tests (BRB-NT)	SRT-LTS	38.5±15.7	-0.8±1.2	47.3±12.8	-0.2±0.9	55.7±10.2	0.5±0.8	0.08	**0.01***
	SRT-CLTR	26.2±13.1	-1.0±0.8	39.9±15.2	-0.1±0.9	49.7±14.3	0.5±0.9	**0.01***	**0.02***
	SRT-D	6.8±2.5	-0.9±1	8.7±2.5	-0.2±1	10.3±1.7	0.5±0.7	**0.03***	**0.007***
	SPART	20.3±4.4	-0.7±1	22.7±4.4	-0.1±1	25±4.1	0.4±0.9	0.2	0.05
	SPART-D	6.9±1.7	-0.6±0.9	7.9±2	-0.1±1	8.8±1.8	0.3±0.9	0.2	0.1
	SDMT	45.1±11.3	-0.8±0.9	53.9±9.7	-0.1±0.8	61.7±12.9	0.5±1	**0.04***	**0.01***
	PASAT-3	31.2±11.5	-1.1±1	43.6±10.6	0.009±0.9	47.8±9.6	0.4±0.8	**0.001***	0.2
	PASAT-2	26.1±10.8	-0.9±1.1	35.3±9.9	0.009±0.1	37.1±8.7	0.2±0.9	**0.04***	0.7
	WLG	23.8±7	-0.5±1.1	27.2±6.6	0.06±1	27.5±6	0.1±0.9	0.2	1.0
Stroop Test (ST)	ST-EIT	17.3±7.8	-0.7±1.4	13.4±5.4	-0.04±1	11.3±3.5	0.3±0.6	**0.05***	0.2
	ST-EIE	1.6±2.3	-1.1±2.1	0.2±0.5	0.2±0.4	0.1±0.4	0.3±0.3	**<0.001***	0.9
Verbal Fluency Test (VF)	Phonemic	34.1±11.9	-0.8±0.9	44.7±12.7	-0.01±0.9	53.9±10.6	0.7±0.8	**0.02***	**0.03***
	Semantic	47.5±11.5	-0.7±1	56.2±11.3	-0.005±0.9	63.5±9.1	0.6±0.8	**0.04***	0.07
	Alternate	37.8±12.3	-0.6±1.1	43.2±10.8	-0.09±0.9	50.5±9.9	0.5±0.9	0.4	0.07
	Shifting Index	0.9±0.3	0.06±1.5	0.9±0.1	-0.06±0.9	0.9±0.1	-0.02±0.9	0.9	1.0
Modified Five Point Test (MFPT)	MFPT-UDs	23.2±12.7	-1.1±1.2	34.3±7.7	-0.07±0.7	41.9±7.5	0.6±0.7	**0.001***	**.003***
	MFPT-CSs	8.7±11.9	-0.7±1	14.9±10.6	-0.1±0.9	22.8±11.1	0.5±0.9	0.2	**.02***
	MFPT-ErrInd	15.3±16.8	-0.7±1.7	6.7±6.9	0.2±0.7	7.1±8.4	0.1±0.8	**0.02***	1.0

Table 4. Characteristics of neuropsychological performances of cognitively impaired RRMS (CI-RRMS), cognitively normal RRMS (CN-RRMS) and healthy controls (HCs), and results

of the comparison between Z-score single subtests in CI-RRMS vs CN-RRMS and in CN-RRMS vs HCs. Mean±SD was provided for raw scores and Z-score indexes.
SRT-LTS = Selective Reminding Test – Long Term Storage; SRT-CLTR = Selective Reminding Test – Consistent Long Term Retrieval; SRT-D = Selective Reminding Test – Delayed; SPART-I = Spatial Recall Test – Immediate; SPART-D = Spatial Recall Test – Delayed; SDMT = Symbol Digit Modalities Test; PASAT = Paced Auditory Serial Addition Test; ST - EIT = Stroop Test - Effect Interference Time; ST - EIE = Stroop Test - Effect Interference Error; MFPT-UDs = Modified Five Point Test – Unique Designs; MFPT-CSs = Modified Five Point Test – Cumulative Strategies; MFPT – ErrInd = Modified Five Point Test – Error Index

Discussion

The aim of the present study was to investigate the cognitive performance of a group of newly-diagnosed RRMS patients classified as being "cognitive normal" (CN-RRMS) considering the most used classification criterion (Amato et al., 2006b), and compared their cognitive performance with that of a matched group of HCs by using a cognitive "functioning-based" approach instead of a cognitive "impairment-based" approach.

Considering this functioning-based approach (i.e., Z-score indexes), the results of the present study highlighted that newly-diagnosed RRMS patients can differ significantly from a group of HCs both on a global level and also with reference to single cognitive domains of attention/processing speed, memory, and executive functioning. However, the most interesting finding is related to the fact that this significant difference between RRMS patients and HCs persists even after isolating only those patients classified as CN, considering the classical categorization criteria (Amato et al., 2006b). Specifically, by using a functioning index based on Z-scores, the group of CN patients showed a significant decrease in cognitive performance as compared to HCs at global level as well as in the domains of memory and executive functions. The grading scores assigned on the basis of this cognitive "functioning-based" approach, as opposed to the classical "impairment-based" approach, highlight that also newly-diagnosed CN patients can show worse cognitive performance as compared to HCs since the early stages of the disease, independently from the effect of other clinical and

demographical variables such as age, education, physical disability, disease duration, fatigue or emotional state. The classic cognitive "impairment-based" approach is undoubtedly affected by different cut-offs threshold and by the different number of neuropsychological tests used, that can render uncertain the diagnosis of CI. Given though cognitive decline can occur as a result of gradual progression related to neurodegeneration, or more transient changes related to inflammatory (i.e., relapses) disease activity, by using a functioning-based approach (i.e., Z-score index) it was expected that also newly-diagnosed RRMS patients would perform worse with respect to HCs. In fact, it has been found that brain alterations due to GM and WM lesions and inflammatory phenomena can be observed since the time of diagnosis and are related to differences in the inflammatory profile (Frischer et al., 2009; Magliozzi et al., 2018). Considering previous studies that showed that early neurodegeneration phenomena affect mainly the frontal and the temporal lobes since the early stage of the disease (Calabrese et al., 2009), it is remarkable that a significant difference between newly-diagnosed CN and HCs was found specifically in the domains of memory and executive functions, that are mainly related to the activity of frontal and temporal brain areas, respectively. It is necessary to strongly highlight the alterations in executive functioning since this domain is often neglected and not included in the most used batteries of neuropsychological tests (i.e., the BRB-NT and the BICAMS).

The Z-score index, in which cognitive performance is considered as a continuum, seemed to effectively reflect the accumulation of cognitive alterations even in those RRMS patients that would be classified as "cognitively normal". As recently highlighted (Weinstock-Guttman, Eckert & Benedict, 2018), if is accepted that cognitive deficits in MS patients, or cognitive decline from baseline, reflect mainly cerebral dysfunctions related to MS disease, after excluding other confounding factors such as physical

disability, fatigue, and emotional state, then cognitive functioning merits clinical attention as would any other indication of disease activity.

With this perspective, it can be overcome the classic impairment-based approach, usually limited by outdated and less representative normative data, optimizing the identification of slight alterations in cognitive performance already evident in newly-diagnosed RRMS patients classified as being "cognitively normal" by using the traditional classification method. As underlined in previous studies, the early detection and monitoring of cognitive dysfunction may be crucial to identify MS patients with a probable worse prognosis and more severe disease progression (Moccia et al., 2015; Pitteri et al., 2017), enabling early pharmacological and non-pharmacological interventions aimed at preventing further cognitive decline and disability in the long term (Forn et al., 2012). According to this, a complete neuropsychological assessment in terms of level of performance, not just prone to classification criteria, seems of paramount importance not only in patients that show evident cognitive impairment (Kalb et al., 2018), but also in apparently "cognitively normal" patients, as highlighted in the present study.

As recently underlined by Weber et al. (2019), neuropsychological tests have shown a significant predictive value also regarding everyday life activity and can be used in the clinical setting as one of several measures to help the clinician understand the impact of MS disease on patients and on their families. This view of considering the patients' "cognitive performance" instead of patients' "cognitive impairment" might open an invaluable window on the real-life performance of MS patients since the time of diagnosis, given that early cognitive alterations can be considered as a signal of increased risk of disease progression (Kalb et al., 2018).

This study has some limitations. First, it is limited by the lack of a longitudinal neuropsychological assessment; this functioning-based approach should be tested more extensively with follow-up measures.

Second, this study focused only on patients with RR course; future studies should extend this approach by investigating different MS populations. However, this is a proof-of-concept study, which the aim is to raise the limitation of using the dichotomic approach derived from classic neuropsychological assessment frequently used in MS population.

Conclusions

The results of the present study suggest that cognitive dysfunction in RRMS is a phenomenon that can be detected also in newly-diagnosed patients. Extensive cognitive assessment since the early phase of the disease would be then of critical importance. This would support accurate judgement of decline in cognitive functioning and would be clinically meaningful to determine a baseline cognitive profile to be monitored in the follow-up. It is preferable to approach and interpret with extreme caution the traditional classification method of cognitive impairment because might fail in measuring actual cognitive performance. In this regard, preferring an approach based on the evaluation of cognitive functioning as a continuous variable should be therefore recommended, also considering the possibility to utilize computerized devices (De Meijer et al., 2018; Golan et al., 2019).

1.2 COGNITIVE FUNCTIONING OF MS PATIENTS AT THE TIME OF DIAGNOSIS IN ASSOCIATION WITH CSF BIOMARKERS OF BOTH INFLAMMATION AND NEURODEGENERATION

Inflammation and neurodegeneration have already been described as core processes in the pathogenesis of MS, both reported in association with a worse cognitive outcome due to neuronal cells transient suffering or death. For example, CSF level of Chitinase 3-like 1 (CHI3L1) and Chitinase 3-like 2 (CHI3L2) were described as effective CSF biomarkers of inflammation associated with cognitive impairment (Modvig et al., 2015; Møllgaard et al., 2016; Quintana et al., 2018). Significant associations with cognitive aspects in MS patients were also reported in favor of other CSF molecules, such as somatostatin (Roca et al., 1999), Bri2-23 (Harris et al., 2010), and amyloid-β (Mori et al., 2011). As mentioned in the previous sections, CSF level of neurofilament light chain (NF-L) was described as an effective biomarker of neurodegeneration and was also found to be associated with cognitive impairment (Quintana et al., 2018; Gaetani et al., 2019; Kalatha et al., 2019). However, other studies are needed to provide further evidence.

Cortical neuronal vulnerability and damage, including both dysfunction and loss, has been shown to highly contribute to the progression of MS lesions (Peterson et al., 2001; Klaver et al., 2015; Schirmer et al., 2019), even independently from cortical demyelination. A “surface-in” of neuronal loss was identified associated with a gradient of microglia activation in the motor cortex of post-mortem MS cases characterized by the presence of elevated inflammation and B-cell follicle like structure in the meninges. Dutta et al. (2006) showed a decrease GABA-ergic parvalbumin-positive (PVALB+) interneurons in the primary motor cortex (M1) of a group of MS patients, in particular within the normal appearing gray matter (NAGM): this change may lead to restructuration and perturbation in neuronal calcium levels causing harmful effects on neurons. Also Clements et al. (2008) highlighted a specific reduction in terms of PVALB-positive

interneurons within layer II of the primary motor cortex (M1) in a group of MS patients; furthermore, a significant correlation between numbers of PVALB staining cells and age of patients was found, suggesting a possible association between MS disease duration and loss of PVALB interneurons. Parvalbumin is a calcium binding protein expressed in a subset of GABA-ergic inhibitory interneurons, and PVALB staining cells have been classified as fast-spiking interneurons (Hu, Gan & Jonas, 2014). Role of PVALB was also investigated in neurological conditions, such as psychiatric disorders (Bitanihirwe et al., 2009), epilepsy (Jiang, Lachance & Rossignol, 2016), and Parkinson's disease (Lanoue, Blatt & Soghomonian, 2013). In line with these evidences, Magliozzi et al. (2019) found that reduction of PVALB gene expression in multiple sclerosis patients, in particular in the motor cortex of post-mortem SPMS cases, represents the most relevant pathological change observed in chronic active subpial gray matter lesion (GML) as well as in the nearby area of NAGM of MS patients, in particular in those with high level of meningeal inflammation and subpial GM demyelination.

As there is a need for specific surrogate biomarkers of both inflammation and GM neurodegeneration to identify those individuals at higher risk of a more severe and rapid disease progression, in the present study it has been investigated combined CSF protein and 3T MRI analysis in a cohort of 95 MS patients at the time of diagnosis and further integrated with cortical lesions number and both global and regional cortical thicknesses determined by 3-dimensional (3D) double inversion recovery (DIR) 3T MRI imaging, neurologic evaluation, and comprehensive neuropsychological assessment. The aim of the study was to provide evidence supporting that PVALB may represent a specific biomarker associated with cortical GM demyelination and neurodegeneration, which could correlate better than a well-established and widely used CSF biomarker in MS (NF-L levels). Furthermore, it was investigated the possible

association between cognitive functioning and 18 inflammatory CSF cytokines investigated in a previous study (Magliozzi et al., 2018).

Materials and methods

Ninety-five MS patients (16 clinically isolated syndromes, CIS; 72 relapsing-remitting MS, RRMS; 7 primary progressive MS, PPMS) were enrolled at the time of diagnosis. Inclusion criteria comprised diagnosis of MS (Polman et al., 2011), no psychiatric or other concomitant non-neurological pathologies, normal or corrected to normal vision, no substance abuse or other concomitant medications that can alter the CNS. All the MS patients enrolled in the study underwent at baseline a CSF withdrawal through lumbar puncture and a neurologic evaluation. Subgroups of MS patients also underwent two 3T MRI scans (T0 and T1), a baseline comprehensive neuropsychological (NP) assessment, and a follow-up neurological evaluation (T1). Flowchart reporting the number of MS patients' groups is provided in Figure 4.

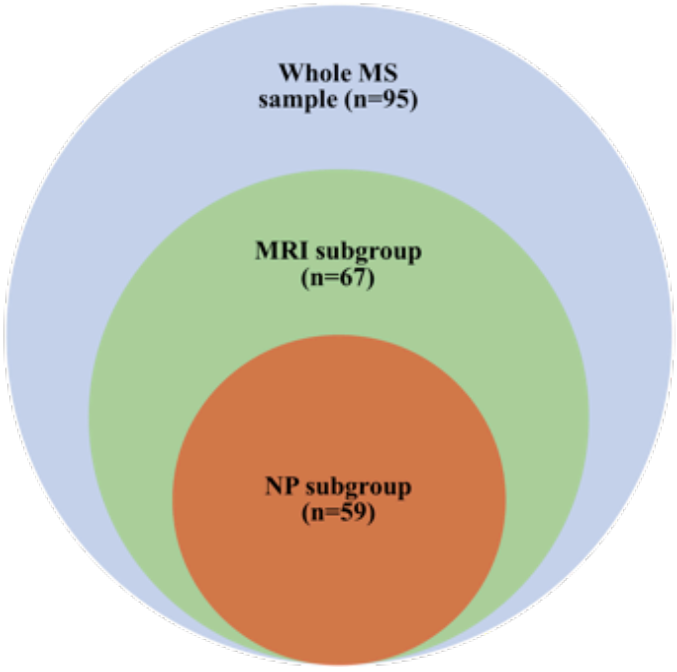

Figure 4. Study flowchart reporting the numbers of patients recruited for the study.

Twenty-seven control subjects including 13 individuals with non-inflammatory neurological diseases (NIND; 4 cerebrovascular diseases, 3 degenerative disease, 2 peripherical neuropathies, 2 headache/dizziness without any CNS abnormality, 2 ischemic white matter lesions), and 14 subjects with other inflammatory neurological diseases (OIND; 3 immune-mediated encephalitis, 2 neuromyelitis optica, 2 neurosarcoidosis, 2 leucoencephalopathies, 2 CNS paraneoplastic syndrome, 2 CNS vasculitis, 1 Behcet's disease), were also recruited at the time of diagnosis and underwent neurological evaluation, 3T MRI and CSF examination. Since no significant differences were found between NIND and OIND subjects in terms of PVALB and Nf-L levels in CSF ($p>0.05$), these patients were considered as a single control group.

The local Ethics Committee approved the study and informed consent was obtained from all the examined patients. All participants were recruited at the MS Center of the Verona University Hospital (Verona, Italy). Demographic, clinical, MRI and neuropsychological characteristics of the participants are provided in Table 5.

		Whole MS group (n=95)	MRI subgroup (n=67)	NP subgroup (n=59)
Demographic and clinical data	Age (years)	37.6 ± 12.1	36.4 ± 11.5	37.7 ± 11.9
	Gender (M/F)	33/62	21/46	15/44
	Disease duration (years)	2.1 ± 2.9	2.1 ± 2.5	2.5 ± 3.1
	MS type (CIS/RR/PP)	16/72/7	13/51/3	13/42/4
	EDSS T0	1.5 (0-6)	1.5 (0-6)	1.5 (0-6)
	Δ-EDSS T0-T1	0 (0-1)	0 (0-1)	0 (0-1)
MRI	CLs volume T0	/	392.5 ± 563.1	410.5 ± 608.7
	CLs number T0	/	4.2 ± 5.3	4.0 ± 5.4
	New CLs T0-T1	/	0.9 ± 1.5	0.8 ± 1.3
	Global CTh T0	/	2.4 ± 0.3	2.4 ± 0.3
	Global CTh T1	/	2.4 ± 0.3	2.4 ± 0.3
	T2WMLV T0	/	996.7 ± 565.1	992.7 ± 559.9
Neuropsychological assessment	Education (years)	/	/	14.0 ± 3.6
	Cognitive impairment classification (CN/mCI/sCI)	/	/	20/25/14
	Global cognitive functioning index (Z-global)	/	/	0.01 ± 0.6

Table 5. Demographic, clinical, MRI and neuropsychological variables of the study population. Means±SDs are provided.
CIS = clinically isolated syndrome; RR = relapsing-remitting; PP = primary-progressive; EDSS = Expanded Disability Status Scale; CLs = cortical lesions; CTh = cortical thickness; T2WMLV = T2 white matter lesion volume; CN = cognitive normal; mCI = mild cognitive impaired; sCI = severe cognitive impaired

MRI Acquisition Protocol and analysis

Three Tesla MRI was performed on all MS patients for diagnostic purpose. In a subgroup of 67 MS patients, MRI sequences were acquired at the Radiology unit of the University Hospital of Borgo Trento (Verona, Italy) by using a Philips Achieva 3T MR Scanner both at baseline (median time interval between baseline MRI and CSF withdrawal: 2 months; range 0-12 months) and also at T1.

The following image sets were acquired:

- 3D T1 weighted Turbo Field Echo (TFE) (Repetition Time (TR) / Echo Time (TE)= 8.4/3.7ms, voxel size of 1x1x1 mm), total acquisition time of 5:51 min;
- 3D Double Inversion Recovery (DIR) (TR/TE=5500/292ms, Inversion Times (TI) TI1/TI2=525ms/2530ms voxel size of 1x1x1mm), Turbo Spin Echo (TSE) read out with an optimal variable flip angle scheme, number of excitations 3, total acquisition time of 10:49 min;
- 3D Fluid Attenuated Inversion Recovery (FLAIR) (TR/TE=5500/292ms, TI=1650ms voxel size of 1x1x1mm), same TSE readout as the DIR sequence, number of excitations 1, total acquisition time of 4:50 min.

MRI analysis

White matter lesion detection and lesion load assessment.

A semiautomatic lesion segmentation technique included in MIPAV software (Medical Image Processing and Visualization. http://mipav.cit.nih.gov) was applied to FLAIR images to identify and segment WM lesions, thus obtaining a T2 hyperintense WM lesion volume (T2WMLV) at baseline. WM lesions were assessed on FLAIR images by an observer with a large experience on MS. The total number of WM lesions was evaluated at baseline while at the end of the study were evaluated the number of new WM lesions and the enlarging WM lesions.

Cortical lesion number and volume

The number of cortical lesions was assessed on DIR images following the recent recommendations for cortical lesions scoring in patients with MS (Fischl, 2012). The total number of CLs at baseline and the number of new CLs at the end of the study were evaluated. Owing to the suboptimal performance of the image-acquisition sequences on MRI in visualizing

subpial lesions, the present analysis has taken into account mainly the intracortical and leukocortical lesions.

The total CL volume at baseline was also calculated using a semiautomatic thresholding technique based on a Fuzzy C-mean algorithm (Pham & Prince, 1999) included in software developed at the National Institutes of Health, Medical Images Processing, Analysis and Visualization (http://mipav.cit.nih.gov).

Cortical thickness evaluation.

The mean cortical thickness was measured in each subject, at disease onset (T0) and at follow-up (T1) as previously detailed (Fischl and Dale, 2000). 3D T1 weighted Turbo Field Echo were analysed using the longitudinal stream of FreeSurfer image analysis suite (release v5.3.0), available online (http://surfer.nmr.mgh.harvard.edu). Topological defects in cortical surfaces due to WM and leukocortical lesions (LCL) were corrected using a semi-automated procedure, which includes WM lesions segmentation and lesions filling.

Neuropsychological assessment

A baseline comprehensive neuropsychological assessment was carried out in a subgroup of 59 patients, that were all tested within 12 months (range 0-12) from the time of the lumbar puncture (median: 1 month). The battery of neuropsychological tests comprised: the Brief Repeatable Battery of Neuropsychological Tests (BRB-NT; Amato et al., 2006a); the Stroop Test (ST; Caffarra et al., 2002); the forward and backward Digit Span Test (DST; Monaco et al., 2013); the Trail Making Test (TMT; Mondini et al., 2011); the Phonological, Semantic, and Alternate Verbal Fluency Test (VF; Costa et al., 2014); and the Modified Five Point Test (MFPT; Cattelani et al., 2011). Scores below the cut-off (5^{th} percentile) of the reference population of each test were classified as failed tests. With this method, MS patients were

classified as having normal cognition (CN, 0 failed subtest), mild cognitive impairment (mCI, up to 2 failed subtests), or severe cognitive impairment (sCI, at least 3 failed subtests) considering their performance on all the neuropsychological tests administered (for a similar procedure, Pitteri et al., 2017). Then, for each patient and for each neuropsychological test of the whole battery of tests, was calculated a cognitive functioning index by using the *Z*-scores formula: $Z = (\text{Raw score} - \text{Average})/\text{SD}$, considering the average and the SD of the whole MS sample (n=59). This procedure allowed us to calculate a global cognitive functioning index (Z-global) from the average of the indexes of each neuropsychological test that permit to compare the cognitive performance of MS patients among different neuropsychological tests.

Neurologic evaluation

All patients underwent a complete neurologic evaluation, including the disability score, as measured by the Expanded Disability Status Scale (EDSS; Kurtzke, 1983). For 74 of them, an EDSS score was available also at T1, allowing us to calculate the EDSS change (Δ-EDSS) between baseline and follow-up evaluations. Disability progression was defined as a sustained increase in EDSS score of at least 1 point when EDSS was 5.5 or less (or a sustained increase in EDSS of 0.5 points when EDSS was more than 5.5) independently from relapses and confirmed at least 12 months later (Kappos, 1998; Hommes et al., 2004).

CSF analysis

The PVALB levels were measured for both CSF of MS (n=95) and control patients with other neurological conditions (n=27) in duplicate in paired CSF post-mortem samples available for the training cohort samples using Human Parvalbumin ELISA kit (MBS2022353, MyBioSource, San Diego, CA, USA) according to manufacturer's instructions. Briefly, the pre-

coated wells were incubated with 100 µl of a 1:1 dilution of the CSF post-mortem for 2 hours at 37 °C and then incubated with reagent A for 1 hour at 37 °C. After washing the samples were incubated with B reagent at 37 °C for 30 minutes and incubated with TBM ELISA substrate at 37 °C after washing. The reaction was stopped after 20 minutes with a specific solution and the optical density was measured at 450 nm on a Model 680 Series microplate reader (Bio-Rad). Samples were analysed in random order and the staff was blinded to the treatment arms. The detection threshold was 0.55 ng/ml. Samples intra-assay variability (coefficients of variation) was below 10%.

In a subgroup of 69 patients and 15 controls, the levels of neurofilament protein light chain (NF-L) in CSF were measured using the Human NF-Light ELISA kit (MyBioSource, San Diego, CA, USA) according to the manufacturer's instructions and previously optimized procedure (Magliozzi et al., 2018), and the quantification was carried out on a VICTOR X3 2030 Multilabel Plate Reader (Perkin Elmer, Walluf, Germany). Intra-assay variability (coefficients of variation) of the samples was below 10%.

In the subgroup that underwent the neuropsychological assessment, concentrations (ng/ml/mg^{Prot}) of 18 CSF inflammatory mediators, identified in a previous cross-sectional study as associated with more severe meningeal inflammation and GM damage, both in post-mortem tissues and in living patients (Magliozzi et al., 2018) were assessed using a combination of immune-assay multiplex techniques based on the Luminex technology (40- and 37-Plex, Bio-Plex X200 System equipped with a magnetic workstation; BioRad, Hercules, CA). All samples were run induplicate in the same experiment and in 2 consecutive experiments, to verify the reproducibility and consistency of the results.

Statistical analysis

Statistical analyses were performed using GraphPad Prism 5 (GraphPad, La Jolla, CA, USA), R (https://www.r-project.org/, version 3.3)

and SPSS statistic (SPSS Inc, Chicago, Illinois, USA, version 24) software. R and Cytoscape (version 3.7.1) software were used to prepare graphs.

Normality distribution of the dependent variables was tested with the Shapiro-Wilk test: since all data were not normally distributed, group comparisons analysis was performed using Mann-Whitney U and Kruskal-Wallis H tests, while Spearman correlation analysis was used. Results were expressed as mean ± SD. A p-value lower than 0.05 was considered significant unless otherwise indicated.

Results

CSF analysis

Significant increased (fold change=6.5, $p<0.001$) PVALB levels were found in the CSF of MS patients respect to controls (Figure 5A).

When MS patients were stratified according to the level of CSF inflammation and cortical demyelination in *MShigh* and *MSlow*, as previously defined (Magliozzi et al., 2018), PVALB CSF level was significantly higher in *MShigh* group compared with the level in *MSlow* group (fold change=1.3, $p=0.015$) (Figure 5B).

No significant correlation was found between PVALB levels and age ($p=0.85$). No significant difference was found in PVALB levels between females and males ($p=0.23$).

Considering the MS disease subgroups, PVALB CSF levels were not significantly different in the 3 groups (CIS-RR-PP, $p=0.061$). However, post-hoc analyses showed significant differences between CIS and PPMS (fold change=1.6, $p=0.047$) and between RRMS and PPMS (fold change=1.6, $p=0.019$), while no significant differences were found between CIS and RRMS ($p=0.93$) (Figure 5C).

Similar significant increase of NF-L levels (fold change=1.8, $p<0.01$) was found in the CSF of MS patients respect to controls (Figure 5D). Then by performing correlation analysis between the two CSF biomarkers, a

significant positive correlation was indeed found between CSF levels of PVALB and NF-L between the two levels (R=0.29, p=0.018) (Figure 5E).

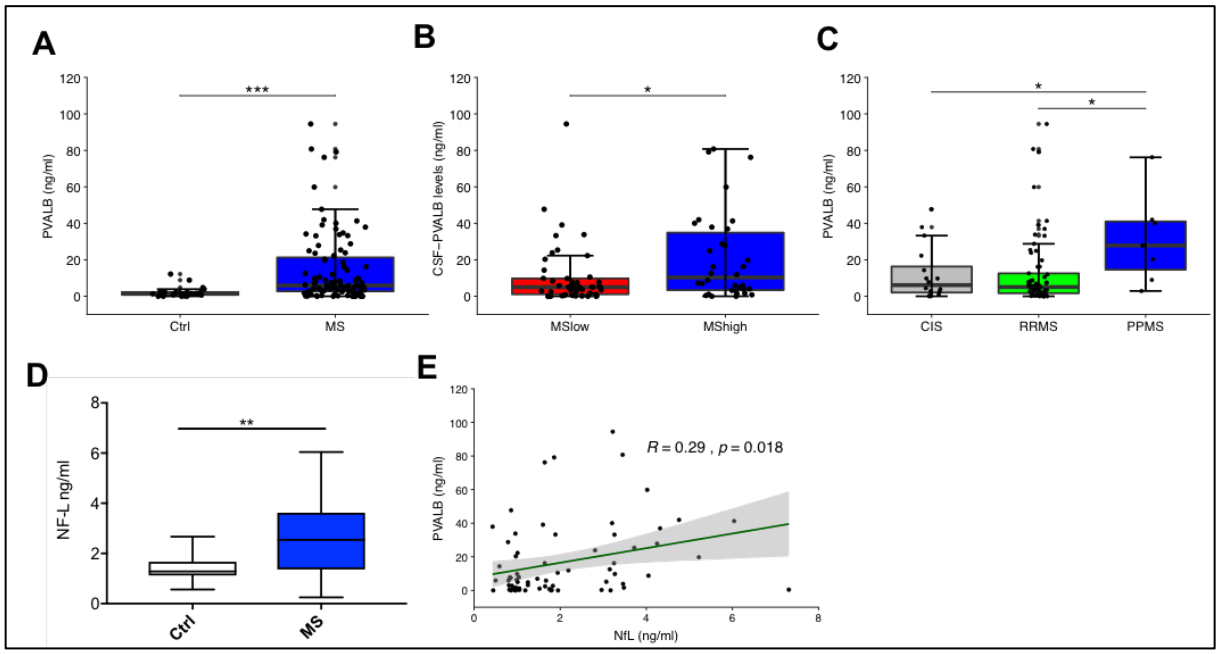

Figure 5. Parvalbumin (PVALB) and neurofilament light chain (NF-L) levels measured in the study.
(A) PVALB concentration in CSF derived from patients with other neurological conditions (Ctrl, gray) and patients with MS (blue).
(B) PVALB concentration in CSF derived from *MSlow* (red) and *MShigh* (blue) patients.
(C) Boxplot representing group comparison of PVALB concentration in CSF on the basis of the type of MS diagnosis (CIS, n=16; RRMS: n=72; PPMS: n=7).
(D) NF-L concentration in CSF derived from patients with other neurological conditions (Ctrl, white) and MS patients (blue).
(E) Spearman correlation revealed a significant association between PVALB and NF-L levels in our group of MS patients (R=0.29). The gray area shows the 95% CIs.
Shapiro test was used to assess the data distribution. Mann Whitney *U*-test was used. The dots represent the samples analysed. Means and SEMs were used to represent the data.
*: $p < 0.05$; **: $p < 0.01$, ***: $p < 0.001$.
CIS = Clinically Isolated Syndromes; RRMS = Relapsing Remitting Multiple Sclerosis; PPMS = Primary Progressive Multiple Sclerosis.

MRI outcomes

No correlation was found between the level of PVALB and the number of WM lesions (p=0.69) (Figure 6A), while significant, positive correlation was found between the level of PVALB and CLs number, both at baseline

(R=0.29, p=0.017) and new CLs at follow-up (R=0.30, p=0.015) (Figure 6B-6C). NF-L level was also significantly associated with CLs number at baseline (R=0.50, p<0.001), with a stronger correlation respect to PVALB level.

Moreover, a significant, negative correlation was found between the level of PVALB and baseline global CTh (R=-0.50, p<0.001) (Figure 6D). Considering specific brain regions (fig. 6E-6G), a significant, negative correlation was found between the level of PVALB and CTh of insula (R=-0.62, p<0.001) (Figure 6E), cingulate gyrus (R=-0.60, p<0.001) (Figure 6F), hippocampus (R=-0.56, p<0.001) (Figure 6G), postcentral gyrus (R=-0.50, p<0.001), calcarine cortex (R=-0.49, p<0.001), temporal gyrus (R=-0.48, p<0.001), fusiform gyrus (R=-0.48, p<0.001), parahippocampal gyrus (R=-0.45, p<0.001), precuneus (R=-0.42, p<0.001), cuneus (R=-0.41, p<0.001), frontal gyrus (R=-0.38, p=0.0013), and motor cortex (R=-0.36, p=0.0024). On the contrary, no significant correlation was found between the level of PVALB and CTh of the gyrus rectus, the occipital gyrus, and the precentral gyrus (all p>0.01). NF-L level was also significantly associated with global cortical atrophy at baseline (R=-0.37, p=0.004), albeit with weaker correlation respect to PVALB level. In Figure 6H, are showed also the association with regional CTh and NF-L, weaker compared with association with PVALB.

Even at T1, a significant negative correlation was found between PVALB CSF level and global CTh (R=-0.50, p<0.001) (Figure 6I).

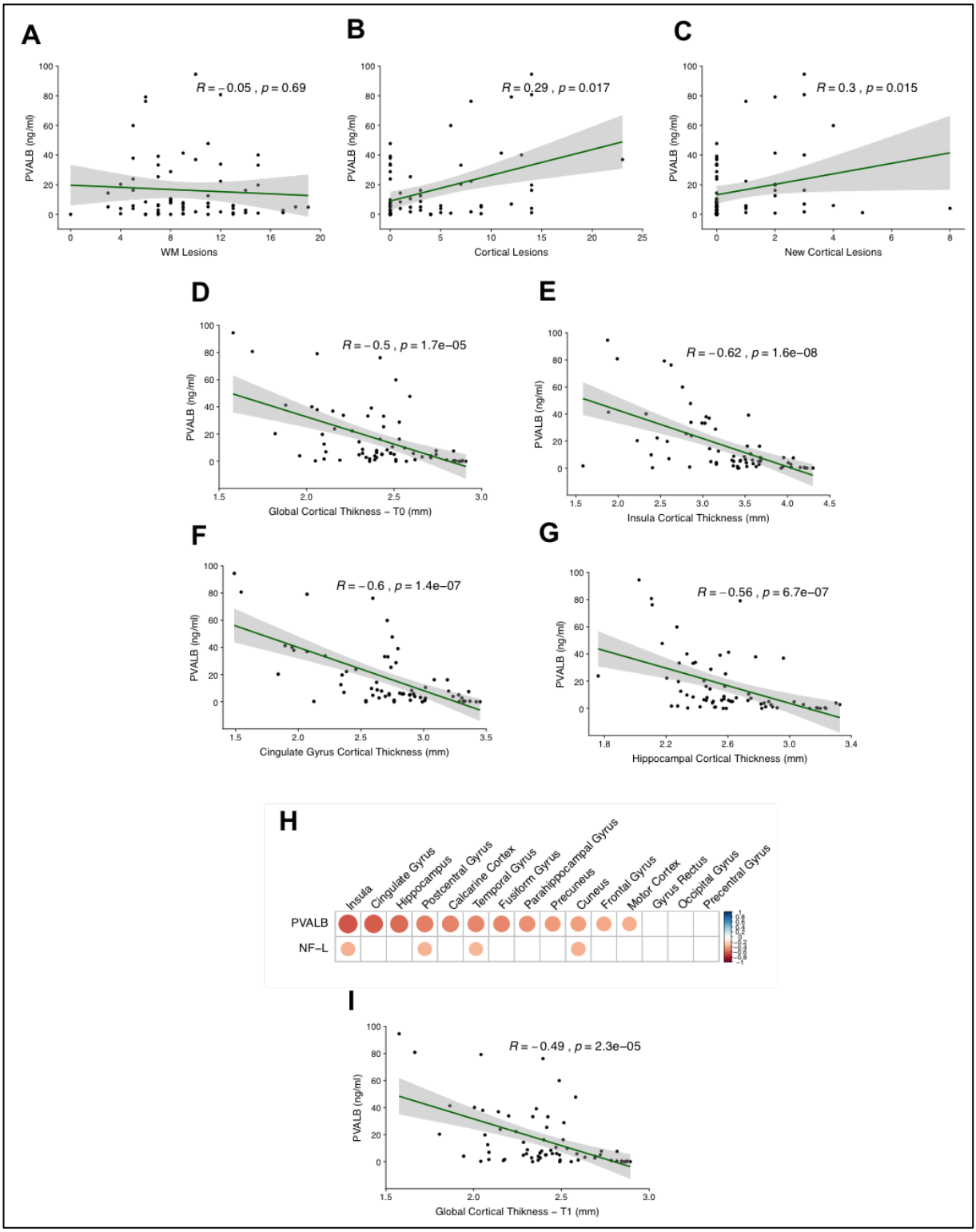

Figure 6. Baseline and follow-up 3T MRI outcome correlations with CSF PVALB levels in MS patients.

(A) Spearman correlation between PVALB level in CSF derived from MS patients and the number of white matter lesions.

(B-C) Spearman correlation analysis between PVALB levels in CSF derived from MS patients and both (B) cortical lesions number at baseline and (C) new cortical lesions at follow-up.
(D) Spearman correlation analysis between PVALB levels in CSF derived from MS patients and global cortical thickness at baseline (T0).
(E-F-G) Representative correlation between PVALB level, global cortical thickness and cortical thickness in the three strongest associated brain areas: insula (E), cingulate gyrus (F) and hippocampus (G).
(H) Bubble chart showing different strengths of correlation between PVALB levels and regional cortical thicknesses (Insula: R=-.62; Cingulated Gyrus: R=-.60; Hippocampus: R=-.56; Postcentral Gyrus: R=-.50; Calcarine Cortex: R=-.49; Temporal Gyrus: R=-.48, Fusiform Gyrus: R=-.48; Parahippocampal Gyrus: R=-.45; Precuneus: R=-.42; Cuneus: R=.41; Frontal Gyrus: R=-.38; Motor Cortex: R=-.36; Gyrus Rectus: R=.30; Occipital Gyrus: R=-.25 ; Precentral Gyrus: R=-.10). Significative threshold was considered when *p*-value > 0.01. A colour scale was used to determine the significance of the cluster enrichment. The bubble chart shows also the association, weaker than the one with PVALB levels, between regional cortical thicknesses and neurofilament light chain (NF-L) levels derived from CSF.
(I) Spearman correlation analysis between PVALB levels in CSF derived from MS patients and global cortical thickness at follow-up (T1).
The dots represent the samples analysed. The gray area shows the 95% CIs.

Neuropsychological outcomes

With respect to the 59 MS patients that underwent the neuropsychological assessment, 20 (33.9%) of them were classified as having no cognitive impairment (CN), 25 (42.4%) as having mild cognitive impairment (mCI), and 14 (23.7%) as having severe cognitive impairment (sCI). Among these three subgroups of CI classification (CN, mCI, sCI), PVALB levels were not significantly different (p=0.26) (Figure 7A). No significant difference was also found for NF-L levels between the three subgroups (p=0.22).

However, considering the global cognitive functioning index (Z-global) instead of cognitive impairment classification, we found a significant, negative correlation between the level of CSF PVALB and this index (R=-0.26, p=0.044) (Figure 7B). On the contrary, no significant correlation was found between NF-L levels and Z-global (p=0.99).

Positive correlation was also found between the baseline global CTh and the Z-global (R=0.39, p=0.004) (Figure 7C).

Furthermore, exploring the association between the level of PVALB and the *Z*-score of each subtest, significant correlations were found with the Spatial Recall Test – Delayed (SPART-D; R=-.26, p=0.45; Figure 7D), the strategy index of the Modified Five Point Test (MFPT-CSs; R=-.32, p=0.020; Figure 7E), the time in the Trail Making Test A (TMT-A time; R=-.61, p<0.001; Figure 7F) and B (TMT-B time; R=-.58, p=0.002; Figure 7G).

However, no significant correlations were found between the level of NF-L and the *Z*-scores of each neuropsychological test (all p>0.05).

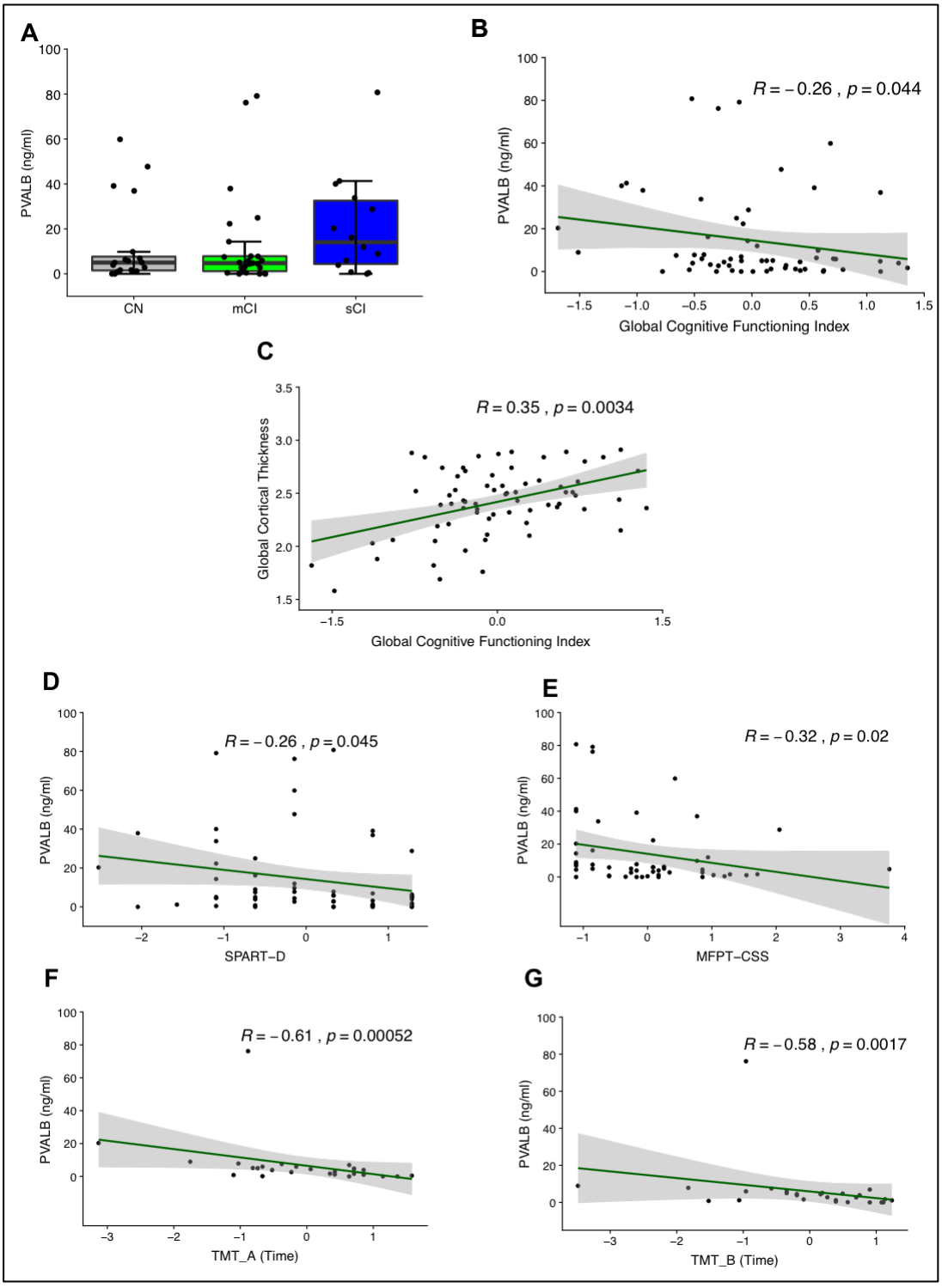

Figure 7. Cognitive functioning alterations in association with CSF PVALB levels and with global cortical thickness.

(A) Boxplot representing the comparison of PVALB level between MS groups of different cognitive impairment. Severely Cognitive Impaired (sCI, blue) MS patients were characterized by a higher, but not clinically significant, PVALB level compared to both Cognitive Normal (CN, gray) and to mildly Cognitive Impaired (mCI, green) MS patients.

(B) Spearman correlation revealed a significant negative association (R=-.26) between PVALB level and global cognitive functioning index.
(C) Spearman correlation revealed a significant negative association (R=.35) between global cognitive functioning index and global cortical thickness.
(D-E-F-G) Spearman correlation revealed a significant negative association (R=-.26) between PVALB level and cognitive functioning index in the Spatial Recall Test - Delayed (SPART-D; p=.045; Figure D), in the Modified Five Point Test – Cumulative Strategies (MFPT-CSs; p=.02; Figure E), in the Trail Making Test – A (TMT-A; p<.001; Figure F) and in the Trail Making Test – B (TMT-B; p=.002; Figure G).
The dots represent the samples analysed. The gray area shows the 95% CIs.

In the MS subgroup that underwent the neuropsychological assessment, the association between cognitive functioning and 18 CSF inflammatory mediators (Magliozzi et al., 2018) was also explored. Correlation analysis showed, after FDR correction, a significant negative association between global cognitive functioning index (Z-global) and LIGHT TNFSF14 (r=-0.26, p=0.044). No other associations were found for global cognitive functioning index with other CSF molecules (all p>.05). Comparison of the level of inflammatory cytokines in CN-MS patients vs. CI-MS patients showed a significant difference only for CSF level of LIGHT (p=0.034), while all the other molecules were not significant between the two groups (all p>0.05).

Physical disability

No significant correlation was found between the level of PVALB and the EDSS at baseline (p=0.16) (Figure 8A). On the contrary, at follow-up (T1), when 54 MS patients had no EDSS change (ΔEDSS=0), while 20 MS patients had EDSS change (n=7 with ΔEDSS=0.5; n=13 with ΔEDSS=1), significant difference was found for the levels of PVALB between the group with no EDSS change and the group with EDSS change (fold change=3.2, p<0.001) (Figure 8B).

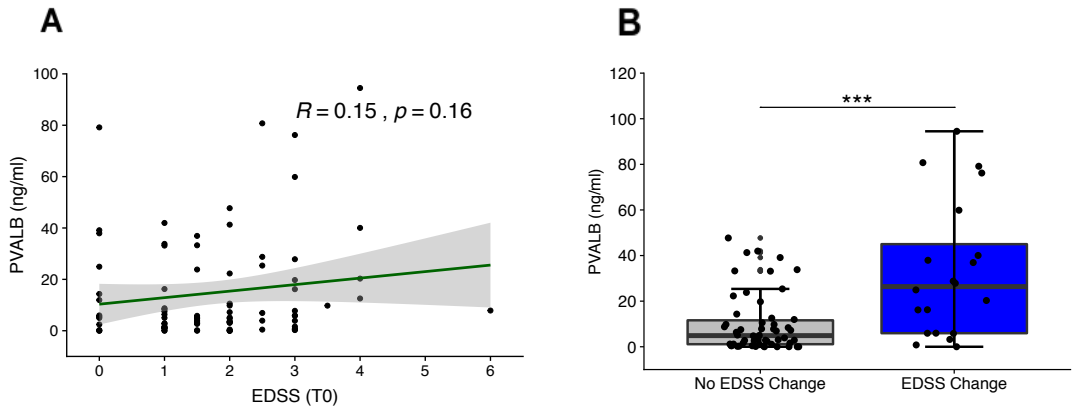

Figure 8. Physical disability in association with CSF PVALB levels.
(A) Spearman correlation revealed no significant negative association between baseline EDSS and PVALB level in CSF of MS patients.
(B) Boxplot representing the comparison of PVALB level between MS patients based on their disability progression at follow-up. MS patients with no disability progression (no EDSS change) were characterized by a significant lower PVALB level, compared to MS patients with an increment of physical disability.
The dots represent the samples analysed. The gray area shows the 95% CIs.
***: $p < 0.001$.
EDSS = Expanded Disability Status Scale

Discussion

In this study was reported for the first time that cerebrospinal levels of parvalbumin (PVALB) protein reflect cortical thickness reduction and worse physical and cognitive disability outcomes. These data suggest that PVALB may represent a new, potential, MS-specific biomarker of cortical neurodegeneration.

Parvalbumin (PVALB) is a calcium binding protein expressed by a subset of GABA-ergic inhibitory interneurons classified as fast-spiking interneurons (Hu, Gan & Jonas, 2014), known to buffer calcium and with a protective role for neurons from excess intracellular calcium (Beers et al., 2001; Dekkers et al., 2004). Clements et al. (2008) highlighted a specific reduction of PVALB-positive interneurons within cortical layer II in a group of MS patients; furthermore, a significant correlation between numbers of PVALB staining cells and age of patients was found, suggesting a possible association between MS disease duration and loss of PVALB interneurons,

that was not identified in the MS cases examined in this study as were selected age-matched. A detailed gene expression study of normal appearing gray matter has shown that decrease of GABA-ergic PVALB-positive interneurons in the primary motor cortex induces perturbation in neuronal calcium levels causing harmful effects on neurodegeneration in MS (Dutta et al., 2006).

Similar to other CSF biomarkers reflecting activity or damage associated with MS pathology, such as neurofilament light chains (NF-L), known to be released in the CSF derived from the damage of axons, PVALB is possibly released in the CSF as a consequence of interneuronal loss. Compared to NF-L, in this study PVALB CSF levels are able to better and early identify cortical thinning, EDSS exacerbation at follow-up and cognitive functioning alterations at the time of diagnosis. In particular, CSF NF-L levels were found mainly associated with cortical lesion number at baseline but not with the other examined MRI and clinical parameters, suggesting that NF-L may represent a good biomarker of focal cortical axonal damage. On the contrary, respect to CSF NF-L levels, the correlation between CSF concentration of PVALB and both baseline cortical lesion number and new cortical lesion at follow-up was still significant but lower, whereas the highest correlation was found with global cortical thickness, both at baseline and at follow-up. These findings may suggest that PVALB, by reflecting the interneuronal loss in the cortex, might represent primarily a biomarker of diffuse cortical neurodegeneration, detectable since the earliest disease stages. However, its measure at the time of diagnosis is able to predict both cortical thinning and EDSS change at follow-up, indicating an important predictive value of PVALB. The fact that high levels of PVALB may be detected in MS patients at the time of diagnosis and the correlation with cortical lesion load is lower respect to the correlation with cortical thinning, strongly support the hypothesis that neurodegeneration may be largely independent of cortical demyelination and may be better estimated by MRI

analysis of changes in the cortical thickness in combination with CSF assessment of PVALB levels.

A significant increment of PVALB levels was reported in the CSF of PPMS respect to both CIS and RRMS, suggesting that this marker might reflect a neuronal loss increasing with the progression. There is ample evidence implicating an active role for neurons accumulating in the pathophysiology of progressive MS including not only neuronal loss but also molecular and metabolic changes within neuronal cell bodies inducing to neuronal dysfunction (Mahad, Trapp & Lassmann, 2015). At the time of diagnosis, it is extremely important to early predict the progressive outcome, therefore PVALB CSF levels might represent a useful new predictive tool. The fact that were found in-vivo PVALB CSF levels early in the diseases supports the idea that PVALB is a good predictor of active cortical pathology, both in GM lesions and in normal appearing GM. This is an important aspect that needs to be validated in longer follow-up studies and independent MS cohorts, considering the actual lack of biomarkers able to early reveal active processes of neurodegeneration.

The results from this study seem to suggest a key role of inflammation in cortical damage and neurodegeneration. Then, when the CSF samples obtained from MS in-vivo patients were analysed and stratified according to the high or low degree of inflammation and high/low cortical lesion load (Magliozzi et al., 2018), was found that MS patients that at baseline have elevated inflammatory CSF profile have also increased levels of PVALB, supporting a strict link between neurodegeneration and intrathecal inflammation since the early disease stages. These data strongly encourage the use of combined CSF analysis, at the time of diagnosis, of a panel of inflammatory and neurodegenerative biomarkers for each individual, in order to early recognize specific MS endo-phenotype and consequently choose the most appropriate individualized therapy.

One of the most interesting results of this study was the correlation found between PVALB concentration in the CSF and both the baseline and follow-up cortical thickness levels, supporting a strict link between neurodegeneration, as detected by PVALB CSF level, and cortical thinning. This finding was particularly evident in some cortical regions, such as insula, cingulate gyrus, hippocampus, postcentral gyrus, calcarine cortex, temporal gyrus, fusiform gyrus, parahippocampal gyrus, precuneus, cuneus, frontal gyrus, and motor cortex, which were all previously identified by both neuropathological and neuroradiological studies to severely affected by cortical pathology (Kulzenigg et al., 2005; Wegner et al., 2006; Frischer et al., 2009; Calabrese et al., 2015a; Haider et al., 2016). These cortical regions, similar to the motor cortex, in which a reduction of PVALB gene expression and GABA-ergic interneurons have been observed, seem to be the most susceptible to neurogeneration, at least in the early stage of the disease. Previous studies hypothesized that the effect of meningeal inflammation was stronger in those regions of the cortex with deep foldings and low CSF flow, such as insula and cingulate gyrus (Haider et al., 2016; Eshaghi et al., 2018): in our study these regions seemed to be highly correlated with the CSF PVALB levels. The elevated association between PVALB CSF levels and cortical thinning in early MS stages strongly support a possible key role of PVALB as a prognostic biomarker of cortical pathology and neurodegeneration that, in combination with advance MRI analysis, may help clinicians to early identify MS patients at risk of a high degree of neurodegeneration and rapidly progressive MS.

The finding that high levels of PVALB in the CSF of patients are not reflected by high accumulation of physical disability at the time of diagnosis but significantly predicts increased EDSS change at follow-up suggests that PVALB may represent a good biomarker of physical disability evolution. These data are corroborated by the fact that the highest PVALB protein levels were detected in the CSF of MS patients with more severe and rapid disease

progression. It would be interesting to further clinically follow the examined MS patients in order to understand if this predictive value of PVALB could be supported.

Considering the global cognitive functioning index (Z-global), a significant correlation was found between this index and both level of CSF PVALB and global cortical thickness; these results corroborate previous studies in which it has been demonstrated that cognitive impairment is mainly associated with both focal and diffuse GM damage, more than WM damage (Sailer et al., 2003; Benedict et al., 2004; Fisniku et al., 2008; Calabrese et al., 2009, 2012; Hulst et al., 2014; Harrison et al., 2015; Daams et al., 2016; Tillema et al., 2016; Rimkus et al., 2019). Moreover, significant correlations were found between CSF PVALB levels and the Z-scores of the Spatial Recall Test – Delayed (SPART-D, a test of long-term verbal memory), both the two parts of the Trail Making Test (part A related to processing speed and part B related to task switching, a component of executive functions), as well as the strategy index of the Modified Five Point Test (a test figurative fluency, a component of executive functions). In summary, it seems that the cognitive domains that are more associated with the level of PVALB are verbal memory, processing speed and executive functions. These results are in line with the significant correlations that was found between level of PVALB and the level of CTh of several frontal and temporal regions, that are mainly related to memory and executive functions. This evidence suggests that PVALB might better reflect cortical interneuronal loss and be a valuable biomarker specific for neurodegeneration and cortical pathology in MS, also highlighted at a neuropsychological level in terms of cognitive performance instead of cognitive impairment. Since previous studies have also demonstrated that impairment of cognitive functions can occur from the early stage of the disease (Rocca et al., 2015) and is known to worsen over time in relation with GM damage (Pitteri et al., 2017; Eijlers et al., 2018), the early identification of RRMS patients at high risk of cognitive disability

progression is a critical problem that shall be addressed by MS clinical research in the near future. Although CI is a frequent and early phenomenon in MS, only few data are available about its role as a prognostic marker. In the present study, evidence was added on the importance of having sensitive cognitive functioning indexes to detect even slight worsening of cognitive performance since the early disease stages, as highlighted in the present study with the significant correlations between the global cognitive functioning index and both PVALB levels and MRI outcomes. Even though NF-L level in CSF (Quintana et al., 2018; Gaetani et al., 2019; Kalatha et al., 2019) and positive oligoclonal bands (OCB+) (Farina et al., 2017) have been proposed as potential prognostic biomarkers of neurodegeneration related to cognitive dysfunction in MS patients, results from this study showed that only the PVALB CSF level may be well related to cognitive functioning alterations; in fact, no significant association was found with global cognitive functioning index and NF-L level in the CSF. In this respect, neuropsychological results are in favor of considering the PVALB a more specific biomarker related to neurodegeneration and GM atrophy in MS.

Considering CSF inflammatory molecules analysed as potential biomarkers of alterations in cognitive functioning, LIGHT was the only one that showed a significant association with cognitive functioning and that was different between cognitive normal and cognitive impaired MS patients. These results are in line with biological literature that highlighted LIGHT as a member of tumor necrosis factor (TNF) family and its role with B-cell inflammatory activity (Hu et al., 2010): previous studies focused on physical disability highlighting the mechanisms between B-cell driven inflammation and cortical pathology (Magliozzi et al., 2007; Lucchinetti et al., 2011), while results from the present study suggest that B-cell activity in MS patients cause neuronal sufferance that could also lead to cognitive alterations.

Next chapter (1.3) will further discuss the role of LIGHT in MS, by highlighting its role as a potential biomarker for accumulation of disease progression, cortical damage and severe disease activity.

A larger MS sample and a follow-up study of the patients included in the current study could corroborate in the future the value of PVALB and LIGHT as potential predictive biomarkers of longitudinal cognitive decline.

Conclusions

By combining a clinical, MRI, and neuropsychological approach, for the first time this study highlighted that cerebrospinal levels of parvalbumin might represent a new potential biomarker of interneuron loss in the cortex of MS patients at time of diagnosis. In particular, this biomarker is able to reflect early alterations of cortical thinning, physical disability accumulation and cognitive functioning, suggesting the usefulness to recognize in advance MS patients that present early sign of neurodegeneration and related physical and cognitive decline to be treated with specific combination of neuroprotective and immunomodulatory treatments. Identifying biomarkers able to recognize those patients at higher risk of poorer disease evolution for whom highly efficacious (but with greater risk) treatments should be prescribed (Magliozzi & Cross, 2020) is nowadays mandatory. This is even more true for those patients that have to deal with the “silent” progression of the disease, which is mainly characterized by neurodegenerative phenomena.

1.3 COGNITIVE, CLINICAL, MRI AND CSF VARIABLES AT THE TIME OF MS DIAGNOSIS AND THEIR PROGNOSTIC VALUE AFTER 4 YEARS

Taking into account the large number of currently available DMTs characterized by different risk-benefit profiles, efforts are being made to implement, early in the disease course, a personalized therapeutic approach which can maximize the chances of achieving a sufficient control of the disease activity in the long term. However, early neurologic and neuroradiological features have limited prognostic value at an individual level: neuroimaging data have confirmed that the extent of GM damage is a good predictor of long-term disability (Calabrese et al., 2012). Therefore, the choice of the pharmacological treatment based on the patient's expected clinical outcome remains a challenging task and, consequently, there is a need for additional biomarkers that could be applied early in the disease course to identify patients destined to have a more severe disease outcome.

It has been demonstrated that leptomeningeal immune cell infiltration and compartmentalized inflammation within the subarachnoid space play a key role in the pathogenesis of the cortical pathology (Magliozzi et al., 2007; Lucchinetti et al., 2011). In particular, increased meningeal inflammation was associated with a gradient of neuronal, astrocyte and oligodendrocyte damage, and with microglial activation, especially in the superficial cortical layers (I-III) (Magliozzi et al., 2010). Furthermore, low levels of cerebrospinal fluid (CSF) β-amyloid were associated with higher risk of disease progression (Pietroboni et al., 2018), whereas CSF concentrations of neurofilament light chain (NF-L) were associated with the 2-years relapse risk but not with disease progression or clinical worsening in newly diagnosed CIS and RRMS patients (Sellebjerg et al., 2018). A recent comprehensive cross-sectional analysis of inflammatory and cytotoxic molecules expressed in the meninges and in the CSF from 27 post-mortem secondary progressive MS cases and in the CSF of 73 living MS patients at diagnosis (Magliozzi et al., 2018) resulted in the identification of increased

levels of 18 CSF inflammatory mediators which were associated with more severe meningeal inflammation and GM damage, both in post-mortem tissues and in-vivo. These data underlined the relationship between intrathecal inflammation and the GM damage, spanning from the early to the late disease stages, and the potential key role of B-cell and innate immune activity in characterizing a severe MS clinical phenotype. Although these inflammatory mediators could be useful early biomarkers of disease severity, the lack of longitudinal data on disease evolution has hitherto prevented the evaluation of their clinical prognostic relevance. Therefore, this study set out to investigate in a large, prospective, and multicenter cohort of MS patients, the association between CSF biomarkers evaluated at the time of diagnosis and the disease evolution over the following 4 years.

Materials and methods

Study population

Ninety-nine treatment naive Relapsing-Remitting (RR) MS patients (F/M=66/33, mean age=40.4±12.0 years, range 18-55, Table 6) from the MS Centre of Verona University Hospital and from the MS Centre of Montichiari (Brescia, Italy) were asked at the time of clinical onset (T0) to participate to this longitudinal 4-years prospective study. Inclusion criteria were MS diagnosis according to the 2010 revision of diagnostic criteria (Polman et al., 2011) and age between 18 and 55 years at baseline. Patients had to be free of any other inflammatory disease and had to have at least 1 ml of CSF obtained in the previous month. Following the enrollment, each patient underwent a 3T brain MRI (description below), was started on first line disease-modifying treatment therapy (Interferon beta1a, Glatiramer Acetate, Teriflunomide or Dimethylfumarate) and was monitored with clinical and radiological evaluation for at least four years after diagnosis. Both at baseline and at the end of the study (T1), a subgroup of patients underwent also a complete neuropsychological examination. The treatment

therapy was chosen by a neurologist with large experience of MS, blinded to the results of CSF profile (only the standard analysis of CSF and the presence/absence of CSF oligoclonal bands were available). The local Ethics Committee approved the study and informed consent was obtained from all the patients.

	Whole group n = 99	**Disease Activity**	
		EDA n= 41	*NEDA n = 58*
Age	40.4±12.0	34.7±12.5	39.5±11.3*
Gender (F/M)	66/33	28/13	38/20
Disease duration (years)	1.8±2.2	1.4±2.1	2.1±2.3*
EDSS median (range)	2.0 (0-6.0)	2.0 (0-4.0)	1.5 (0-6.0)
OCBs (yes/no)	87/12	38/3	49/9
CLs number	4.3 ±5.3	7.8±6.0	1.8±3.0***
CTh T0 (mm)	2.4 ±0.4	2.3±0.4	2.5±0.3**
T2WMLL T0 (mm^3)	960.7±352.7	1000.1±357.6	932.9±349.6
Number of Gad+ lesions◊	0.2±0.6	0.3±0.5	0.2±0.6
Number of relapses◊	1.0±0.3	1.0±0.4	1.0±0.2
Median time to event (IQR)	4.0 (1.1)	3.0 (1.8)	4.4 (1.0)

Table 6. Characteristics of patients at the beginning of the study and after having divided them on the base of their clinical evolution. Data are reported as mean and standard deviation if not differently reported. Comparison analyses results between EDA and NEDA patients are provided.

* for $p<0.05$, ** for $p<0.01$, *** for $p<0.001$.

◊ = in the year before lumbar puncture;

CLs = cortical lesions; CTh = cortical thickness; EDA = evidence of disease activity; NEDA = no evidence of disease activity; F = female; M = male; EDSS = Expanded Disability Status Scale; OCB = oligoclonal band; T2WMLL = T2 white matter lesion load; Gad = gadolinium; IQR = interquartile range

CSF protein analysis

CSF samples were obtained at the time of diagnosis, at least 2 months after the last relapse and within one month of the MRI, according to Consensus Guidelines for CSF and Blood Biobanking (Teunissen et al., 2009). After centrifugation, the supernatant and the cell pellet were stored separately at -80°C. The CSF analysis was optimized and performed by two independent investigators, blinded with respect to the clinical and MRI features.

The concentrations (ng/ml/mg^{Prot}) of 18 inflammatory mediators identified in a previous cross-sectional study (Magliozzi et al., 2018) were assessed using immune-assay multiplex techniques based on the Luminex technology (Bio-Plex-X200 System equipped with a magnetic workstation, BioRad, Hercules, CA, USA) according to procedures previously optimized (Magliozzi et al., 2018).

Combined CSF protein analysis was previously performed in a control group (Magliozzi et al., 2018), in order to define standard CSF levels of the inflammatory mediators.

To verify the reproducibility and consistency of the results, all samples were run in duplicate in the same and in two consecutive experiments. According to previous published procedures (Opsahl et al., 2016), for each examined patient, the CSF level of each protein was normalized to the total protein concentration of each CSF sample (measured by the Bradford protocol). However, the use of absolute concentrations of CSF protein levels did not affect the correlations nor the multivariate models.

MRI Acquisition Protocol

Three Tesla MRI was performed on each patient at the time of diagnosis and then yearly for at least 4 years. All the scans were acquired at the Neuroradiology Unit of the University Hospital of Verona, at least 2

months after the last relapse. MRI sequences were acquired using a Philips Achieva 3T MR Scanner. The following image sets were acquired:

- 3D-T1 weighted Turbo Field Echo (TFE) (Repetition Time (TR) / Echo Time (TE)= 8.4/3.7ms, voxel size of 1x1x1mm), acquisition time of 5:51 min;
- 3D-Double Inversion Recovery (DIR) (TR/TE=5500/275ms, Inversion Times (TI) TI1/TI2=450ms/2550ms voxel size of 1x1x1mm), Turbo Spin Echo (TSE) readout with an optimal variable flip angle scheme, number of excitations 3, acquisition time of 10:49 min;
- 3D-Fluid Attenuated Inversion Recovery (FLAIR) (TR/TE=8000/288ms, TI=2356ms voxel size of 1x1x1mm), same TSE readout as the DIR sequence, number of excitations 1, acquisition time of 4:48 min;
- 3D-T1 weighted TFE post contrast with the same parameters of the pre-contrast sequence (TR/TE= 8.4/3.7ms, voxel size of 1x1x1 mm), acquisition time of 5:51 min.

MRI Analysis

Lesion detection

The number of WM lesions at baseline and the number of Gad positive, new and enlarging WM lesions at the end of the study were assessed on FLAIR images by a neuroradiologist with extensive experience of MS. The number of total and new cortical lesions were assessed on DIR images following the recent recommendations (Geurts et al., 2011). Owing to the suboptimal performance of the MRI in visualizing subpial lesions, the present analysis has taken into account mainly the intracortical and leukocortical lesions. These can be considered as a surrogate marker for the total GM demyelination as well as demonstrated by several neuropathological/MRI studies.

Cortical thickness evaluation.

The annualized cortical thinning was calculated using the longitudinal stream of FreeSurfer (Fischl, 2012) image analysis suite (version 6.0; http://surfer.nmr.mgh.harvard.edu). Topological defects in cortical surfaces due to WM and leukocortical lesions were corrected using a semi-automated procedure, which includes WM lesion segmentation and lesion filling. The software was applied to the 3D-T1 weighted Turbo Field Echo sequences. Five-time points, approximately equally spaced, were used for each subject to create an unbiased, subject-specific template (Reuter & Fischl, 2011) which is generated by iteratively aligning all input images to a median image using a symmetric robust registration method. Several steps in the processing of the serial MRI scans (e.g., skull stripping, atlas registration, etc.) are then initialized with common information from the subject-specific template. The software calculated the mean cortical thickness change between the first and the last scan; the result was then divided for the exact numbers of years between the two scans in order to obtain the annualized cortical thickness change.

Compared with cross-sectional studies, this longitudinal design significantly reduces the confounding effect of inter-individual morphological variability by using each subject as his/her own control.

Neurologic evaluation

A neurologic evaluation, including the EDSS (Kurtzke, 1983) assessment, was performed every six months. The number of relapses and the treatments administered was also recorded.

Patients with (EDA) or without (NEDA) evidence of disease activity were identified. EDA/NEDA parameter was defined as a composite score obtained from three related measures of disease activity: (i) evidence of relapses; (ii) confirmed disability progression as assessed by an increase of

the EDSS score by at least 1 point sustained over 6 months; and (iii) evidence of new or newly enlarging WM T2 lesions (Havrdova & Galetta, 2010).

Neuropsychological assessment

A subgroup of 52 MS patients was assessed at baseline and at the end of the study with the Brief Repeatable Battery of Neuropsychological Tests (BRB-NT; Amato et al., 2006a) and the Stroop Test (Caffarra et al., 2002). MS patients were classified in 2 groups considering a conservative approach (Pitteri et al., 2017): "cognitively normal" if patients performed above the cut-off in all neuropsychological tests, and "cognitively impaired" if patients performed below the cut-off in 1 or more tests.

Statistical analysis

The Mann-Whitney *U*-test was used to compare patients with different clinical and MRI characteristics both at baseline and at the end of the study. Pairwise univariate Spearman rank correlation index was used to evaluate correlations between CSF proteins levels at diagnosis and both clinical and MRI variables at follow-up. A false discovery rate (FDR) correction was applied.

To estimate the relative contributions of CSF proteins levels at the time of diagnosis to the clinical profile at the end of the study, logistic regression models with backward stepwise model selection (using Akaike Information Criteria, AIC) were applied. In these models, CSF levels were log transformed to approximate normal distribution as well as to obtain reliable odds ratio (OR) estimates. The log base 2 ($\log_2$) was preferred to a more intuitive interpretation of ORs: each unit in $\log_2$ (protein level) corresponds to a doubling in protein level. Specifically, EDSS progression (binary: worsened/stable), relapse appearance (binary: yes/no), NEDA status (binary: yes/no), new WM lesions (binary: yes/no), new Gad+ lesions (binary yes/no) and cognitive status (binary: impaired/normal), all at the

end of the study, were used as outcome variables while all the 18 CSF profiles were covariates. Logistic regressions were adjusted for age at lumbar puncture. The multivariate model was also performed including treatments variable as categorical covariate: one for each first line DMT. Five-fold cross-validation was performed to estimate the goodness of the models in order to increase the generalization: for the first fold, four-fifth of data are used as the training set while the remaining data are used as the testing set to measure performance. The average over the five-fold of accuracy, specificity and sensitivity were used as performance measures.

The receiver operating characteristic (ROC) curve analysis (applying the Youden approach) was performed to assess specificity and sensitivity of each CSF protein in discrimination between EDA and NEDA patients. The whole sample was randomly split 70%-30% into training and testing set datasets, respectively. A stratified 5-fold-cross-validation was applied on the training set in order to detect the optimal cut-off which represents the minimal threshold for the CSF biomarker to classify EDA patients (mean for each molecules' thresholds over the 5-test fold) and their effectiveness was checked on the test data.

In order to further evaluate the CSF cytokines predictiveness, a separate backward stepwise logistic regression was performed (stepwise approach), assessing the risk of EDA/NEDA based on the disease duration, the number of relapses in the year before diagnosis, the EDSS, the WM lesion number, the T2WMLL and the number of Gad+ lesions at LP. The performance of this model was then compared to the performance of the model including CSF protein levels.

To estimate the relative contributions of CSF protein levels to the annualized cortical thinning, stepwise regression analysis was performed. Finally, to differentiate the specific contribution of CSF protein levels in patients with and without evidence of disease activity, two separate stepwise

regression analyses were performed in both NEDA and EDA patients. All the abovementioned models were adjusted for age at lumbar puncture.

Results

Neurologic and neuroradiological data at the end of the study

Each patient completed the 4-years neurologic and neuroradiological follow-up (median= 4.3 years, range= 4-5 years). Demographic and clinical characteristics of the study population are reported in Table 6. Following the diagnosis, patients were started on disease modifying treatment: Dimethylfumarate (n= 51), Teriflunomide (n=21), Glatiramer acetate (n=16), and interferon beta1a (n=11).

By the end of the study, 58 patients (58.6%) did not show any evidence of disease activity (NEDA), while 41 (41.4%) patients experienced some sort of disease activity: 29 patients had one or more clinical relapses (mean number of relapse per patient =1.4±0.6, range=1-3), 43 patients showed one or more new or enlarged WM lesions (3.9±2.1, range=1-9), which in 22 patients were active (new Gad+ lesions= 1.7±0.7, range=1-3). In addition, 36 patients experienced EDSS increase (mean increase 1.0±0.6, range=0.5-3.0). Twenty-three patients (23.2%) were switched to a second line treatment (Natalizumab, Ocrelizumab or Fingolimod), after the occurrence of disease activity while on first line treatment. The distribution of first line treatments was not significantly different among patients with or without disease activity (Table 6 for more details).

CSF profiles identify patients with disease activity during the follow-up period

Compared to NEDA patients, those who experienced disease activity during the observation period had at diagnosis, among others, higher CSF levels of CXCL13 (27.9±39.8 in EDA vs 5.4±6.7 in NEDA), CXCL12 (3.760± 3.187 in EDA vs 1.564±0.954 in NEDA), IFNγ (33.8±34.7 in EDA vs 7.3±8.8 in NEDA), TNF (59.8±45.4 in EDA vs 19.1±15.7 in NEDA) and

sCD163 (54460±21801 in EDA vs 42419±22012 in NEDA) (adjusted p <0.001 for all tests; Table 7 and Figure 9). An example of a patient's CSF profile with severe clinic and radiological activity is showed in Figure 10.

	Relapse		New WM lesions		EDSS change		New CLs		Disease Activity	
	Yes (n=29)	*No (n=70)*	*Yes (n=43)*	*No (n=56)*	*Yes (n=36)*	*No (n=63)*	*Yes (n=46)*	*No (n=53)*	*NEDA (n=58)*	*EDA (n=41)*
CXCL13	34.8±45.3 [0-213]***	6.4±7.7 [0-33]	26.6±39.3 [0-213]***	5.6±6.6 [0-25]	30.1±42.0 [0-213]***	6.0±7.1 [0-33]	27.1±37.7 [0-213] ***	4.0±4.8 [0-17]	5.4±6.7 [0-25]	27.9±39.8 [0-213]***
CXCL12 (x10^3)	3.8±3.7 [0.4-17.9]**	1.9±1.3 [0.08-6.0]	3.5±3.2 [0.3-17.9]***	1.7±1.0 [0.1-5.3]	3.9±3.3 [0.7-17.9]***	1.7±1.1 [0.1-5.3]	2.9±2.2 [0.1-17.9]	2.1±2.6 [0.1-12.8]	1.6±1.0 [0.1-5.3]	3.8±3.2 [0.5-17.9]***
CCL25 (x10^2)	1.3±1.0 [0-4.0]	1.1±0.7 [0-3.1]	1.3±0.9 [0-4.0]	1.0±0.7 [0-3.1]	1.3±1.0 [0-4.0]	1.0±0.6 [0-3.1]	1.1±0.7 [0-3.0]	1.1±0.8 [0-4.0]	0.9±0.6 [0-3.1]	1.4±0.9 [0-4.0]
TNFα	62.3±49.1 [6-223]***	25.0±24.1 [0-133]	54.7±44.1 [6-223]***	21.5±22.4 [0-133]	60.9±47.7 [5-223]***	21.7±18.5 [0-76]	42.8±41.4 [2-223]	30.0±32.5 [0-133]	19.1±15.7 [0-60]	59.8±45.4 [6-223]***
sTNFR1 (x10^3)	7.5±5.3 [0.3-24.0]**	3.9±1.9 [0.8-10.9]	6.6±4.6 [0.3-24.0]***	3.6±1.8 [0.8 -9.9]	7.1±4.8 [0.3-24.0]***	3.7±1.9 [0.8-9.9]	5.7±4.5 [1.0-24.0]	4.3±2.6 [0.3-12.1]	3.6±1.7 [0.8-9.9]	6.9±4.6 [0.3-24.0]***
TWEAK (x10^3)	2.5±2.7 [0.06-10.7]	1.6±1.6 [0.1-8.2]	2.0±2.4 [0.06-10.7]	1.7±1.6 [0.1-8.2]	2.4±2.6 [0.06-10.7]	1.6±1.5 [0.1-8.2]	2.2±2.3 [0.01-10.7]	1.6±1.6 [0.06-8.2]	1.6±1.5 [0.1-8.2]	2.3±2.5 [0.06-10.7]
APRIL (x10^3)	36.4±32.5 [7.2-145.8]*	28.4±28.8 [0.9-125.2]	33.2±30.5 [7.0-145.8]	28.8±29.7 [0.9-125.2]	34.0±32.2 [6.8-145.8]	28.9±28.8 [0.9-125.2]	34.1±32.1 [0.9-125.2]	27.7±27.9 [2.6-145.8]	26.3±27.6 [0.9-125.2]	36.9±32.4 [7.9-145.8] **
BAFF (x10^3)	12.6±10.0 [3.4-45.0]*	8.1±4.3 [1.5-21.6]	11.5±8.6 [2.4-45.0]*	7.8±4.4 [1.5-21.6]	11.1±8.3 [3.8-45.0]	8.5±5.7 [1.5-36.6]	9.3±6.6 [2.4-45.0]	9.5±7.0 [1.5-36.6]	7.5±4.3 [1.5-21.6]	12.1±8.6 [2.4-45.0] ***
LIGHT (x10^3)	0.9±1.1 [0-6]***	0.2±0.3 [0-1]	0.7±1.0 [0-6]**	0.2±0.3 [0-1]	0.8±1.0 [0-6]***	0.2±0.3 [0-1]	0.6±1.0 [0-6]	0.3±0.3 [0-2]	0.2±0.3 [0-1]	0.7±1.0 [0-6] ***
IFN-γ	35.3±39.2 [2-144]***	11.2±14.4 [0-64]	33.6±34.0 [2-144]***	6.5±7.5 [0-38]	32.8±36.5 [0-144]***	10.0±13.2 [0-66]	18.3±20.6 [0-78]	18.2±31.1 [0-144]	7.3±8.8 [0-38]	33.8±34.7 [2-144] ***
IFN-α2	31.2±28.2 [0-99]**	13.1±13.0 [0-49]	30.2±24.2 [0-99]***	9.4±9.9 [0-30]	28.2±26.1 [0-99]**	12.8±13.6 [0-58]	18.4±17.4 [0-59]	18.4±22.8 [0.0-99.0]	10.8±10.7 [0.0-32.5]	29.1±25.5 [0.0-99.0] ***
IFNλ2 (x10^3)	0.5±1.0 [0-5]*	0.3±0.7 [0-4]	0.4±0.9 [0-5]	0.4±0.8 [0-4]	0.5±1.0 [0-5]**	0.3±0.8 [0-4]	0.4±0.9 [0-5]	0.4±0.8 [0-4]	0.4±0.8 [0-4]	0.4±0.9 [0-5]
IL-6	46.2±68.3 [2-286]**	18.0±38.1 [1-249]	50.0±69.2 [2-286]***	8.0±7.8 [1-38]	49.6±73.3 [3-286]***	12.9±21.3 [1-140]	33.9±62.6 [1-286]	19.6±35.6 [1-229]	8.9±8.4 [1-38]	50.7±70.9 [2-286]***
IL-8 (x10^2)	1.1±1.1 [0.1-4.5]**	0.4±0.5 [0.02-2.8]	1.05±1.0 [0.1-4.5]***	0.3±0.2 [0.02-1.0]	1.0±1.1 [0.1-4.5]**	0.4±0.4 [0.02-2.1]	0.7±0.9 [0.02-4.5]	0.5±0.7 [0.02-3.1]	0.3±0.2 [0.02-1.1]	1.0±1.0 [0.1-4.5]***
IL-10	25.7±25.7 [3-93]	19.4±18.8 [0-79]	24.9±22.8 [3-93]	18.5±19.5 [0-79]	22.9±23.8 [3-93]	20.3±19.6 [0-79]	19.2±16.9 [1-73]	23.1±24.2 [0-93]	18.6±19.3 [0-79]	25.0±23.3 [3-93]
MMP2 (x10^3)	1.9±6.7 [0.09-36.5]	2.0±4.9 [0.07-27.9]	1.4±5.5 [0.09-36.5]	2.4±5.4 [0.07-27.9]	0.6±0.5 [0.07-3.0]	2.8±6.7 [0.09-36.5]	2.0±6.0 [0.09-36.5]	2.0±5.0 [0.07-27.9]	2.3±5.4 [0.07-27.9]	1.5±5.6 [0.09-36.5]
Pentraxin 3 (x10^3)	0.7±0.9 [0-4] ***	0.3±0.4 [0-3]	0.5±0.8 [0-3]	0.3±0.5 [0-3]	0.5±0.8 [0-3]	0.3±0.5 [0-3]	0.5±0.7 [0-3]	0.3±0.4 [0-3]	0.3±0.4 [0-3]	0.5±0.8 [0-4]
sCD163 (x10^3)	58.0±22.9 [0-109]**	43.0±21.1 [1-126]	51.4±21.8 [0-109]	44.4±23.0 [1-126]	52.7±21.9 [0-109]	44.4±22.6 [0-126]	50.5±18.8 [10-86]	44.7±25.4 [0-126]	42.4±22.0 [1-126]	54.5±21.8 [0-109]**

Table 7. Comparison of the mean values of CSF cytokines between neurologic and MRI patients' subgroups. Data are reported as mean ± standard deviation [minimum-maximum]. Concentration of cytokines is expressed in ng/ml/mg^{Prot}
*: $p < 0.05$; **: $p < 0.01$, ***: $p < 0.001$.

CLs = cortical lesions; EDSS = Expanded Disability Status Scale; EDA = evidence of disease activity; NEDA = no evidence of disease activity; WM = white matter

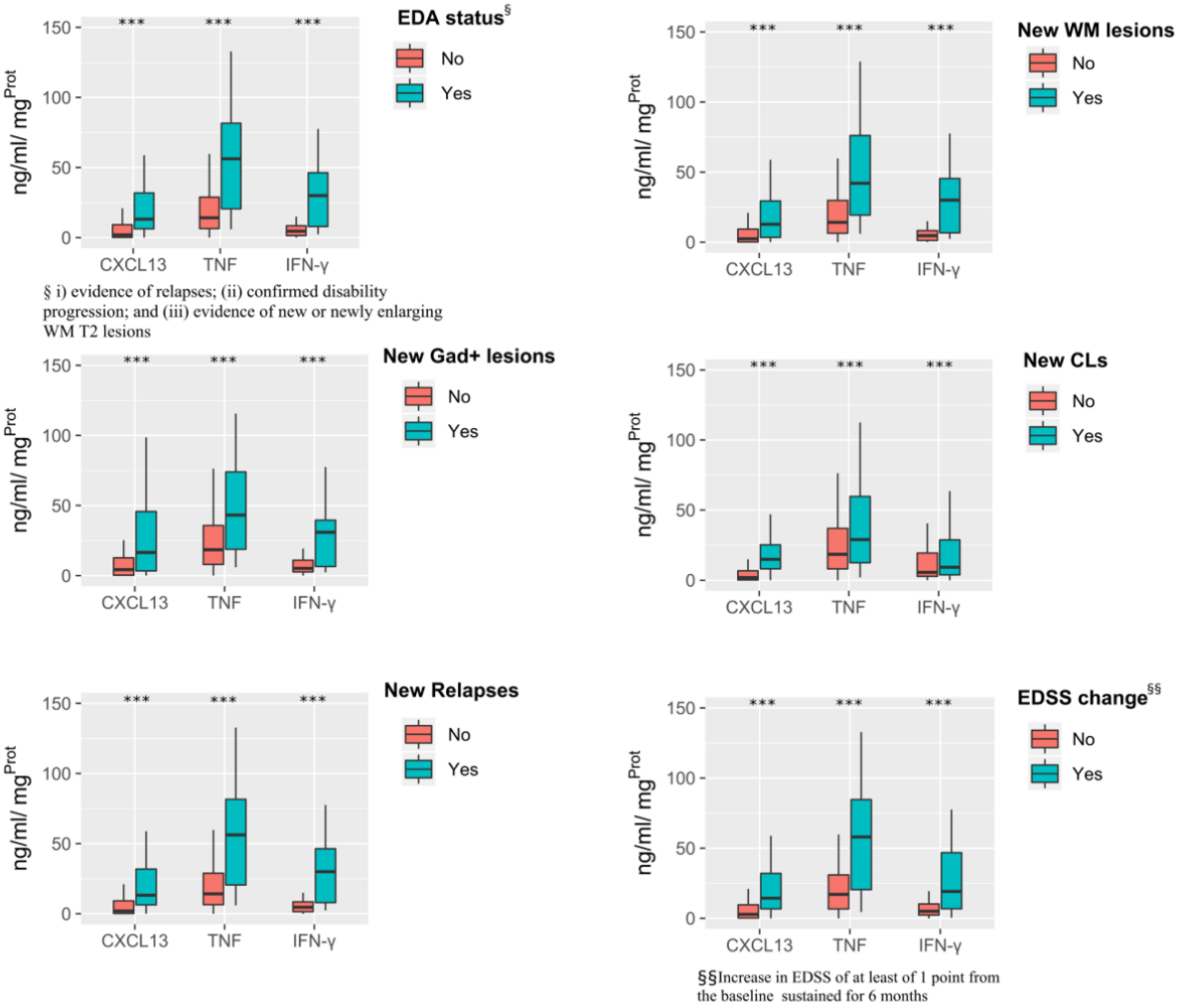

Figure 9. Boxplots of the concentrations (ng/ml/mg^{Prot}) of CXCL13, TNF and IFNγ in patients with or without disease activity, with or without new WM lesions, with or without new Gad+ lesions, with or without new CLs, with or without relapses, and with or without EDSS change. Boxplots show the medians and the two hinges.
***: $p < 0.001$.
NEDA = no evidence of disease activity; WM = white matter; Gad = gadolinium; CLs = cortical lesions; EDSS = Expanded Disability Status Scale

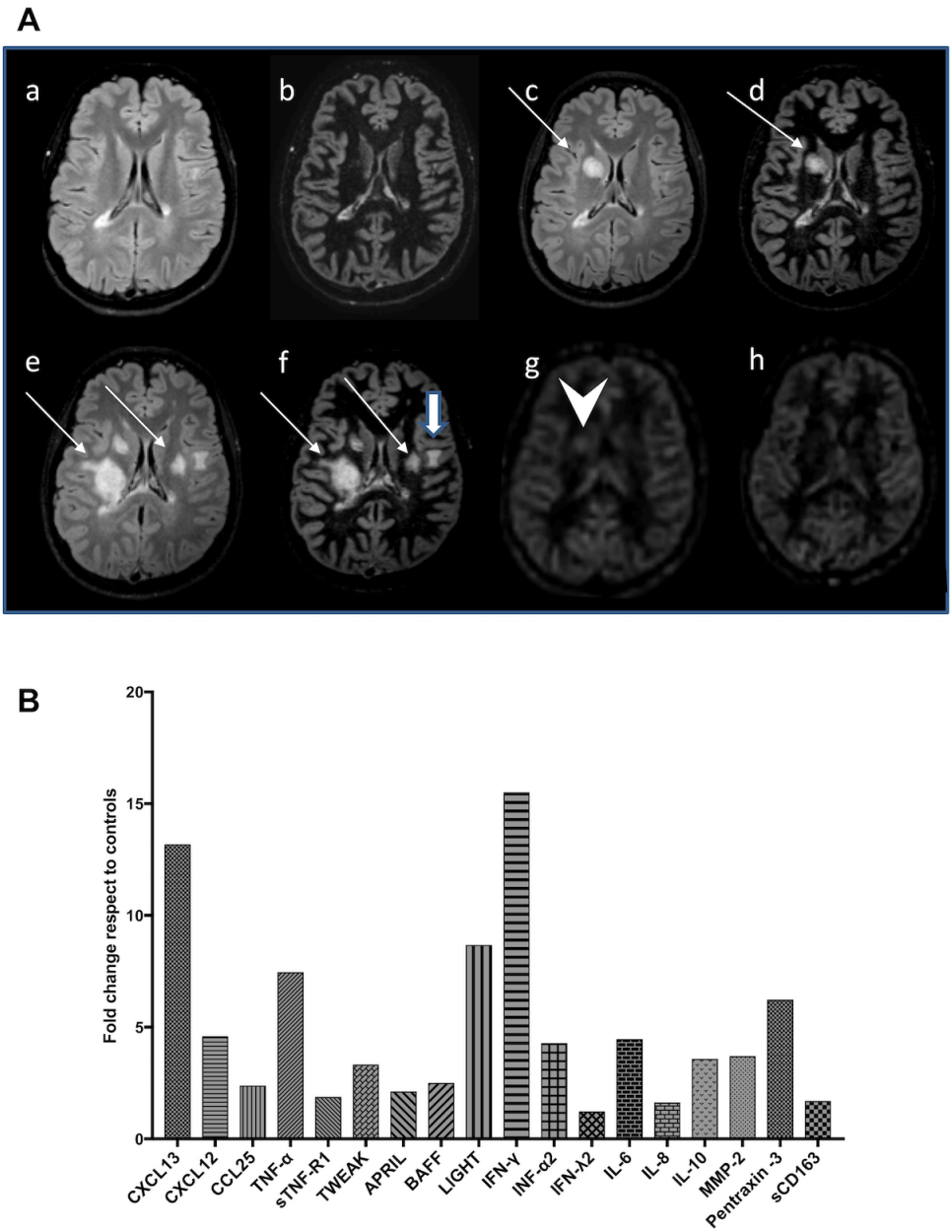

Figure 10. A: Male patient 25 years old. Disease onset in February 2015. FLAIR and DIR sequence obtained at diagnosis (a and b), in 2017 (c and d) and in 2018 (e and f). The significant increase of the lesion load in the white matter (thin arrows) and in the gray matter (thick arrow) parallels several severe clinical relapses with the appearance of a severe left hemiparesis. Perfusion sequence showed increased perfusion in one of the lesions, adjacent to the head of the R caudate nucleus (arrowhead); this hyperperfused lesion disappeared at the following MRI control (h).
B: Combined CSF inflammatory profile of the above MS patient shown in A. The diagram shows fold changes of the 18 examined molecules with respect to the group of controls (other neurological diseases).

Several associations between increased cytokine level and neurologic, and neuroradiological outcomes were identified: CXCL13 correlated (adjusted $p<0.001$) with EDSS worsening (rho=0.52), with the occurrence of new WM lesions (rho=0.46) and with number of new relapses (rho=0.45); TNF correlated (adjusted $p<0.001$) with the occurrence of new WM lesions (rho=0.50), EDSS change (rho=0.47) and with the number of new relapses (rho=0.42); IFNγ correlated (adjusted $p<0.001$) with the occurrence of new WM lesions (rho=0.57), EDSS change (rho=0.38) and with the number of new relapses (rho=0.38) (Figure 11 and Table 8 for other univariate correlations).

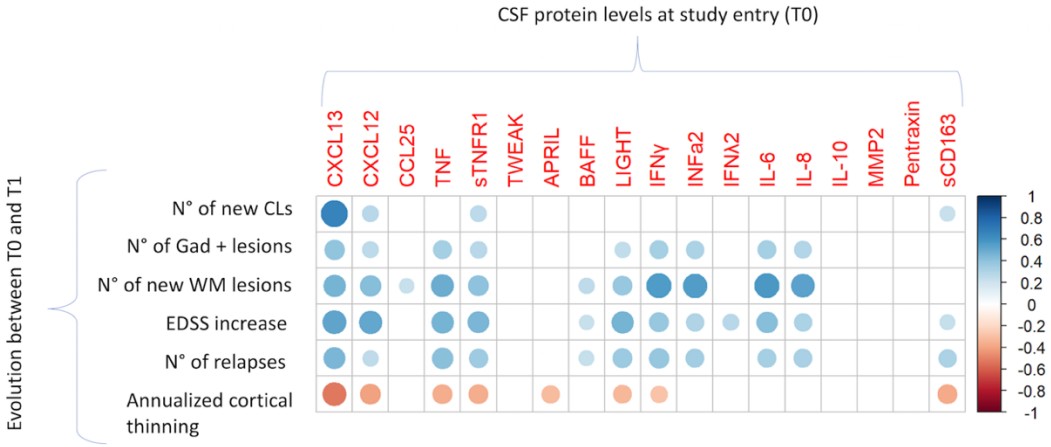

Figure 11. Correlation matrix showing Pairwise Univariate Spearman rank correlation index between CSF protein levels at diagnosis and both neurologic and MRI variables at the end of follow-up. A false discovery rate (FDR) correction was applied. The Spearman rank is proportional to the dimension/intensity of the bubbles.
CLs = cortical lesions; Gad = gadolinium; WM = white matter; EDSS = Expanded Disability Status Scale

	New CLs	New Gad+ lesions	New WM lesions	EDSS change	Number of Relapses	Annualized cortical thinning
CXCL13	0.67***	0.39***	0.46***	0.52***	0.45***	-0.58***
CXCL12	0.37**	0.27*	0.42***	0.51***	0.26*	-0.45***
CCL25	0.03	0.07	0.22*	0.14	0.06	-0.08
TNF	0.19	0.34**	0.50***	0.47***	0.42***	-0.36***
sTNFR1	0.26*	0.28*	0.40***	0.45***	0.35**	-0.37***
TWEAK	0.20	0.04	-0.01	0.16	0.15	-0.37***
APRIL	0.21	0.08	0.16	0.20	0.21	-0.31**
BAFF	0.03	0.07	0.25*	0.23*	0.23*	-0.16
LIGHT	0.20	0.25*	0.37***	0.47***	0.37***	-0.33**
IFNγ	0.13	0.33**	0.57***	0.38***	0.38***	-0.39***
INFα2	0.00	0.31**	0.55***	0.31**	0.34**	-0.14
IFNλ2	0.10	0.18	0.19	0.27*	0.22	-0.16
IL-6	0.15	0.33**	0.57***	0.42***	0.34**	-0.21
IL-8	0.16	0.29**	0.54***	0.32**	0.33**	-0.22
IL-10	-0.03	0.05	0.21	0.05	0.13	0.00
MMP2	0.06	0.05	0.06	0.01	0.12	-0.11
Pentraxin3	0.18	0.05	-0.05	0.10	0.13	-0.21
sCD163	0.33**	0.19	0.20	0.24*	0.32**	-0.47***

Table 8. Pairwise Univariate Spearman rank correlation index between CSF profile at T0 and clinical and radiological data at T1
*: $p < 0.05$; **: $p < 0.01$, ***: $p < 0.001$.
CLs = cortical lesions; EDSS = Expanded Disability Status Scale; Gad = gadolinium; EDA = evidence of disease activity; WM = white matter

The logistic regression analysis showed that IFNγ (OR=3.09, CI=1.09-8.79) and CXCL13 (OR=2.36, CI=1.11-5.22) were independently associated with the risk of disease activity (Table 9 for other multivariate analyses). The generalized linear model correctly classified 35 of 41 patients who showed EDA and 51 of 58 patients who retained a NEDA status (sensitivity=88%, specificity=85% and accuracy=87%). On the contrary, the logistic regression including only neurologic and conventional MRI parameters showed a very low sensitivity (19%), high specificity (86%) and an overall accuracy of 52% which was much lower than the model including the MRI, neurologic and CSF variables which correctly classified 32 of the

41 patients who showed EDA and 48 of the 58 patients who retained the NEDA status (Sensitivity 83%, Specificity 84% and Accuracy 81%).

Statistically significant variables	Odds Ratio	Confidence Interval (95%)
N° of Relapses T0-T1		
CXCL13	3.10	1.26-4.48
Pentraxin3	1.65	1.11-2.46
New Gad lesions T0-T1		
IFNγ	4.06	1.56-10.60
TNF	3.91	2.11-8.67
CXCL13	1.69	1.02-2.80
EDSS change T0-T1		
TNF	5,61	1.06-29.65
IFNα2	2.95	1.42-6.12
CXCL13	2.57	1.41-4.66
New WM lesions T0-T1		
CXCL13	10.59	2.88-38.80
IFNγ	8.06	2.51-25.82
INFα2	7.54	2.12-78.00
EDA at T1		
IFNγ	3.09	1.69-8.79
CXCL13	2.36	1.11-5.22
New CLs T0-T1		
CXCL13	2.27	1.64-3.14

Table 9. Logistic Regression Analysis between CSF profile at the beginning of the study (T0), and neurologic and MRI variables at the end of the study (T1).
CLs = cortical lesions; EDSS = Expanded Disability Status Scale; Gad = gadolinium; EDA = evidence of disease activity; WM = white matter

CSF profile predicts the severity of GM damage at 4 years

Forty-six patients experienced one or more new CLs (3.6±2.7, range=1-8) during the study. This subgroup had more than fivefold increased CSF levels of CXCL13 (27.1±37.7 vs 4.0±4.8, adjusted $p<0.001$), compared to patients without new CLs. A strong positive correlation was observed between the number of new CLs and the CXCL13 CSF level (rho=0.67, adjusted $p<0.001$) (Figure 12), which was also confirmed by the logistic regression analysis (OR=2.27, CI=1.64-3.14).

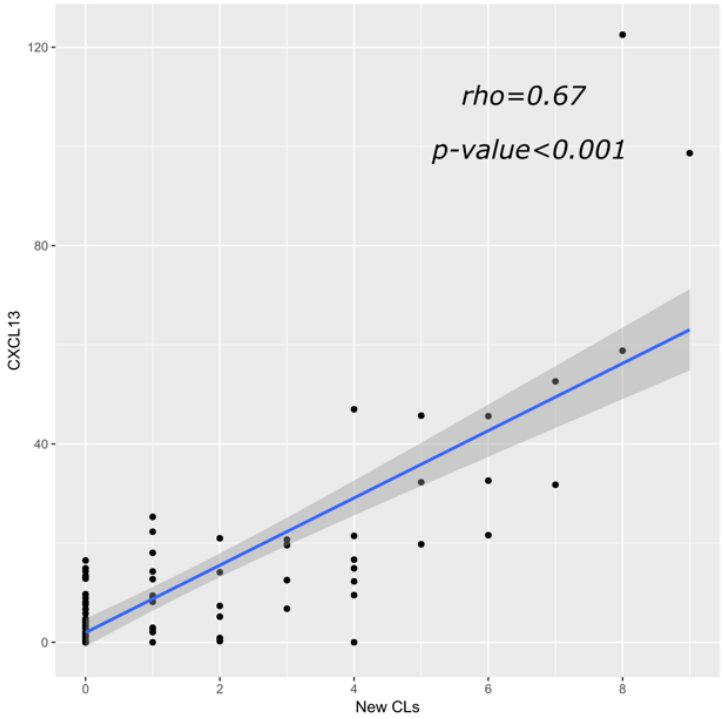

Figure 12. Scatter plot of the relationship between the CXCL13 and the new cortical lesions (new CLs) at T1 and its corresponding correlation coefficient (rho) and p-value.

During the 4-years observation period, in the whole group, the mean annualized cortical thinning was -0.69% (standard deviation=0.27%, range=-0.19%/-1.6%). Among others (Table 8), the rate of cortical atrophy correlated with CSF levels of CXCL13 (rho=-0.58), CXCL12 (rho=-0.45), and sCD163 (rho=-0.47) (adjusted $p<0.001$). The stepwise regression analysis revealed that CXCL13 ($\beta=4.7*10^{-4}$, $p<0.001$), TNF ($\beta=3.1*10^{-3}$, $p=0.004$), LIGHT ($\beta=2.6*10^{-4}$, $p=0.003$), sCD163 ($\beta=4.3*10^{-3}$, $p=0.009$) and TWEAK ($\beta=3.4*10^{-3}$, $p=0.024$) were independent predictors of the annualized cortical thinning.

The rate of cortical atrophy was much higher in the EDA group (-0.88%±0.23%) compared to the NEDA group (-0.56%±0.20%, $p<0.001$). In each group, was assessed the correlation between CSF cytokine levels and the cortical thickness changes. The multiple regression analysis demonstrated that, among patients with EDA, CSF levels of CXCL13 ($\beta=6.1*10^{-4}$, $p<0.001$), TNF ($\beta=5.1*10^{3}$, $p<0.001$), LIGHT ($\beta=4.6*10^{-4}$, $p<0.001$) and sCD163 ($\beta=3.3*10^{-4}$, $p<0.001$) were independently

associated with annualized cortical thinning. Interestingly, in the NEDA group, CXCL13 was the only CSF cytokine found to be associated with the rate of global cortical loss ($\beta=3.0^{*}\ 10^{-3}$, $p=0.009$).

CSF profile is not associated with cognitive dysfunction at 4 years

At baseline, 28 patients had no cognitive abnormalities and 24 patients were cognitively impaired, whereas, at the end of the study, 1 cognitive normal patient worsened in the cognitive performance and was reclassified as cognitive impaired. No significant differences in the CSF profile were observed between cognitive normal and impaired patients both at study entry and at the end of the study.

ROC curve analysis

Table 10 shows the ROC curve analysis assessing area under the curve (AUC), accuracy, sensitivity and specificity for each CSF cytokines in predicting EDA and NEDA patients, in the testing set.

The CSF proteins were sorted by accuracy and the first 8 proteins (CXCL13, IFN-γ, TNF, sTNFR, sCD163, LIGHT, CXCL12, IFN-λ2) can be considered good biomarkers for identification of patients with disability progression. All these molecules, in fact, show an accuracy of ≥ 70%.

	Optimal cut-off	AUC [95%]	Accuracy (%)	Specificity (%)	Sensitivity (%)
LIGHT (x10^3)	0.16	0.78 [0.62-0.95]	80	84	72
sTNFR1 (x10^3)	5.69	0.77 [0.60-0.93]	80	89	63
CXCL13	7.85	0.80 [0.62-0.93]	77	73	81
TNFα	37.73	0.76 [0.59-0.93]	77	79	73
CXCL12 (x10^3)	1.91	0.80 [0.62-0.93]	77	74	82
IFN-γ	10.54	0.67 [0.49-0.85]	70	79	55
sCD163 (x10^3)	46.07	0.67 [0.49-0.85]	70	79	55
IFNλ2	37.78	0.67 [0.49-0.85]	70	79	55
INF-α2	27.61	0.58 [0.43-0.74]	67	89	27
IL-6	11.07	0.67 [0.50-0.86]	67	63	72
BAFF (x10^3)	7.94	0.68 [0.50-0.87]	67	63	73
APRIL (x10^3)	18.786	0.61 [0.43-0.80]	63	68	54
IL-8	44.34	0.52 [0.34-0.71]	57	68	36
MMP2	449.80	0.56 [0.37-0.75]	57	57	55
CCL25	114.40	0.54 [0.35-0.73]	57	63	45
TWEAK (x10^3)	1.01	0.55 [0.37-0.74]	47	53	36
IL-10	17.18	0.56 [0.37-0.75]	43	42	45
Pentraxin3 (x10^3)	0.23	0.49 [0.31-0.68]	33	15	63

Table 10. Results of the ROC curve analysis in the testing set. For each CSF protein was reported the optimal threshold, AUC [95%], Accuracy, Specificity and Sensitivity. Concentration of the cytokines is expressed ng/ml/mg^{Prot}.
AUC= area under the curve

Discussion

In this 4-year, longitudinal, prospective, multicenter study was observed that specific CSF inflammatory cytokines at diagnosis are able to differentiate patients destined to have a more severe disease course and worsening cortical pathology over time. The logistic regression multivariate model highlighted the CSF levels of CXCL13 and IFNγ as the most accurate early predictors of clinical activity, reflected by the occurrence of new relapses, new white matter lesions, and disability accumulation among patients on first line DMTs. These results are in agreement with previous studies demonstrating an association between elevated CSF CXCL13 levels and intrathecal inflammatory markers (i.e. CSF restricted IgG oligoclonal bands) with a high number of early relapses (Ferraro et al., 2015).

With the large number of currently available DMTs, characterized by different efficacy and safety profiles, efforts are being made to implement, early in the disease course, a personalized therapeutic approach which can maximize the chances of achieving a potential disease control in the long term. However, early neurologic and neuroradiological features have limited prognostic value and optimizing the treatment selection, based on the patient's expected clinical outcome, remains a significant challenge (Ruggieri et al., 2018). These results demonstrate that specific CSF profiles at diagnosis characterize patients that are likely to have a suboptimal response to first line treatments. The study demonstrated that, compared to the evaluation of only neurologic and MRI parameters, the characterization of CSF biomarkers allows to achieve much higher accuracy in predicting the disease severity. Therefore, the CSF analysis can more accurately identify, early in the disease course, patients at high risk of experiencing disease activity, which should be treated more aggressively before the occurrence of irreversible damage. The logistic regression model including only clinical and radiological variables showed less accuracy (52%) in identifying patients at risk of disease activity, compared to the model incorporating CSF variables (87%).

The NEDA status places emphasizes on the inflammatory clinical and subclinical WM activity and mostly overlooks ongoing GM damage (Giovannoni et al., 2017). This imbalance was tried to be addressed by also assessing the correlation between CSF cytokines and the accumulation of cortical pathology (new cortical lesions and cortical thickness change) and on the cognitive outcome. A strong association was observed between high levels of specific CSF chemokines, mainly related to lymphoid neogenesis and B-cells, such as CXCL13, TNF, TWEAK and LIGHT, not only with more severe GM damage at diagnosis but also with the cortical damage accumulation over 4 years. This is in line with a previous study that showed a correlation between early inflammation and mechanisms underlying the

focal and global GM pathology in the long term (Scalfari et al., 2018). Besides, it supports that B-cells play a key role in the biological mechanisms underlying the cortical pathology (Magliozzi et al., 2007). Importantly, CSF levels of CXCL13 and TNF were also identified as independent predictors of EDSS increase, supporting the previously demonstrated link (Magliozzi et al., 2007; Lucchinetti et al., 2011) between B-cell driven inflammation and long-term disability progression, secondary to the worsening cortical pathology.

Several studies have provided evidence that soluble demyelinating and cytotoxic factors produced by B-cells, both in the perivascular and meningeal spaces, are likely to play a major role in the development of GM damage, including cortical demyelination and synapse and neuronal loss (Lisak et al., 2012; Gardner et al., 2013). In-vivo, such molecules could originate both from cells resident in the meninges or from cells within the CNS parenchyma (Rainey-Barger et al., 2011; Esen et al., 2014; Stilund et al., 2015), but little is known about the occurrence of this expression in the MS brain. Interestingly, both CXCL13 and CXCL12 are indispensable chemokines for the initial recruitment and organization of B cells in the CNS. Therefore, together with the increased presence of TNF and lymphotoxin-α/β in the subarachnoid space, raised CXCL13 is likely to contribute to the formation of tertiary lymphoid tissues in the meninges (Aloisi & Pujol-Borrell, 2006). Noteworthy is also the observation that the CSF levels of sCD163, a marker of myeloid lineage, cells and expressed by activated monocyte/macrophage in MS lesions and infiltrates (Stilund et al., 2014), was found to be predictive of the rate of cortical thinning. Activated monocytes/macrophages in the meningeal infiltrates and circulating in the CSF could, therefore, mediate phagocytosis, inflammatory cell recruitment/activation and antigen presentation, via sCD163 and CXCL13 related mechanisms (Aloisi & Pujol-Borrell, 2006). This agrees with the

notion that innate immunity significantly contributes to the cortical pathology since the onset of the disease (Calabrese et al., 2015a).

This study demonstrates that while some molecules are associated with lesion activity, such as new T2 WM lesions and Gad+ lesions, other CSF biomarkers correlated with non-conventional MRI parameters (such as cortical thinning) not strictly linked to acute blood-brain-barrier damage and WM lesion activity but, possibly, to intrathecal inflammation (such as CXCL13) and microglial activity (such as sCD163) which persists in the progressive stages (Komori et al., 2015; Magliozzi et al., 2018). While some CSF molecules (CXCL13) were found to correlate with mechanisms driving both WM and GM damage, others were shown to be associated selectively with only one. A potential link was found between B-cell intrathecal activity (CXCL12, LIGHT and TWEAK) and the cortical pathology. In contrast, parameters of WM inflammatory activity, such as new WM lesions and number of relapses, correlated with high levels of different CSF biomarkers (TNF, IFNγ) that indicate increased intrathecal T-cell activity.

All the pro-inflammatory molecules, produced by both lymphocytes and/or macrophages compartmentalized within CSF and meningeal/perivascular infiltrates, may play a direct role in determining demyelination and pathological cell alterations, but may also induce tissue damage indirectly by mediating glial (microglia and astrocytes) activation, as suggested by the relationship between sCD163 and both GM and WM pathology. Indeed, the interaction of complex inflammatory mechanisms, leading to perivascular infiltration by both T- and B-lymphocytes and monocytes, is likely to result in blood brain barrier leakage and the formation of active WM plaques (Lassmann, 2018).

This study is not free from limitations. The low number of patients enrolled from only two MS centers and the lack of an independent validation cohort might limit the general applicability of the model. Nevertheless, the characteristics of this cohort, enrolled at the time of diagnosis when

treatment had yet to begin, the comprehensive CSF evaluation and the 4-year neurologic and neuroradiological follow-up by using high field 3T MRI make results robust. Furthermore, patients with or without cognitive impairment at the beginning and at the end of the study did not show any significant difference in the CSF profile. As shown in the previous chapter (1.2), it seems probable that at the time of diagnosis a more specific analysis is required to highlight significant associations between CSF and cognitive aspects. Instead, for the follow-up part is important to highlight that, since patients were included in the study very early considering their disease course, the non-significant differences in the CSF composition between cognitive normal and cognitive impaired patients could be due to the fact that patients are still able to compensate the low level of accumulated neuronal/axonal degeneration and, consequently, to postpone the more severe cognitive deterioration. In fact, four years of follow-up could be a period of time excessively short for severe cognitive outcomes to occur: probably a longer period of time would cause a greater accumulation of cognitive deterioration that, at baseline, could be reflected by different CSF profiles. The study presented in the previous chapter (1.2) described cross-sectional results about the association between CSF and cognition, while future studies should probably include a longer follow-up time in order to obtain consistent longitudinal results. Moreover, treatments might have affected the models results. Nevertheless, all the patients were started on first-line DMTs with an almost similar effect on the disease activity (Mikol et al., 2008; Fox et al., 2012) and with no proven effect on the accumulation of cortical damage. Furthermore, the choice of the pharmacological treatment was blinded with respect to the CSF cytokines profile. Therefore, it can be assumed that results are not impacted by the use of DMTs. Finally, the low sensitivity of DIR for cortical lesions, especially for subpial lesions, might affect the correlation with CSF inflammation; for this reason, a

separate analysis was run to assess also the relationship between CSF cytokines and the annual cortical thickness change.

Conclusions

In conclusion, by prospectively following a large cohort of MS patients from diagnosis to the 4-year time point, we demonstrated that high CSF levels of the B-cell related cytokines, CXCL13, CXCL12, TWEAK, and LIGHT, together with proinflammatory molecules, IFNγ and TNF, and the monocyte activity biomarker, sCD163, represent a potential molecular signature that could be used to distinguish patients at high risk of experiencing increased more severe disease activity, increased GM damage and disability progression in the early phase of the disease. A more detailed CSF analysis may allow an early stratification of patients, based on their expected outcome, which can assist the optimization of the therapeutic approach. Further studies on larger cohorts are required to support the potential application of CSF profiling as a reliable prognostic tool which could be used in clinical practice with a high level of reproducibility.

PROJECT 2 – NON-CONVENTIONAL COGNITIVE ASSESSMENT IN MS

2.1 FALSE MEMORIES IN RRMS: A PRELIMINARY INVESTIGATION WITH THE DRM PARADIGM

As previously highlighted, cognitive impairment (CI) is a common consequence of MS (Amato, Zipoli & Portaccio, 2006; Grzegorski & Losy, 2017), mainly associated with GM damage (Calabrese et al., 2009; Rocca et al., 2015), and also visible in the long term (Pitteri, et al., 2017; Eijlers et al., 2018). Among cognitive functions, impairment in learning and memory is one of the most frequently and early detected in MS patients (Amato et al., 2001; Bobholz & Rao, 2003; Sepulcre et al., 2006; Chiaravalloti & DeLuca, 2008). Although extensively debated, the pathogenesis of memory dysfunction in MS is still unclear. A key challenge that research in this field is still facing is whether memory impairments arise from deficits in *encoding*, *retrieval*, or *both*. To this regard, some authors argued that the impaired *retrieval* process is responsible for long-term memory impairment (Beatty et al., 1993; Rao et al., 1993; Bobholz et al., 2006). Conversely, other studies suggested that dysfunctions of encoding and information storing processes lead to memory deficits (DeLuca et al., 1998; Thornton, Raz & Tucker, 2002). Finally, other studies suggested different mechanisms depending on the stage of the disease, with impairments in information retrieval in the early stage of MS and additional impairments in the encoding phase as the disease progresses (e.g., Brissart, et al., 2012).

A recent behavioral study by Abad et al. (2015) pointed out that when a failure occurs in the retrieval of lexical information, this might be because of a reduction of the lexical pool, related to semantic memory. By studying a large cohort of MS patients, the authors found that their content of semantic networks was noticeably inferior to that of the matched control networks, both in terms of nodes and edges (i.e., relationships between

nodes), despite the network complexity and cluster hierarchy of related words seemed to be preserved. The authors concluded that semantic networks from MS patients were smaller, with fewer nodes and edges scattered in a narrower distribution than that of the healthy controls.

These behavioral findings are in line with the results of neuroimaging studies which have shown that neurologic diseases are often associated with brain damage that may differentially affect short-term connections in the GM (due to direct damage to neurons or synapses) and long-term connections in the WM (Rocca et al., 2015). Specifically, for MS patients, Rimkus et al. (2019) have shown that single-subject GM networks have fewer connections compared to controls, with the most pronounced differences shown by patients with CI. Moreover, authors observed GM network disruptions related to specific cognitive functions: they identified the most significant loss of GM connections in the bilateral fusiform gyrus and the mesial temporal-occipital regions, implicated in visual, categorical, and also semantic memory. Other imaging studies have also reported that networks become increasingly disorganized in MS patients as related to disease progression and that these measures might be related to CI (Riccitelli et al., 2011).

Semantic verbal fluency tests are the most common methods to quantify semantic memory; however, these tests, in addition to their lexical–semantic component, are affected by executive function components since the task involves strategy generation, monitoring of strategy success, and inhibition of previous responses. As a consequence, the use of semantic verbal fluency tests cannot exclude the reflection of other levels of cognitive impairment outside the semantic domain (Abad et al., 2015), leaving unsolved the hypothesis of a primary semantic memory impairment in MS patients.

In the present study, on the basis of these findings, memory alterations in MS patients were further investigated, by focusing on memory

distortions (i.e., false memories) for semantically-related material. To this aim, the Deese-Roediger-McDermott paradigm was used (hereafter called DRM; Deese, 1959; Roediger & McDermott, 1995), a robust paradigm to investigate false memories for associatively/semantically-related neutral words. In the task, participants are required to learn lists of words constructed so that all the words in the list are associates of a word not presented in the list (i.e., critical word or critical lure; e.g., bed, rest, awake…all related to sleep). After each list is presented, participants are given a free recall task and they are instructed not to guess. After participants have studied and recalled several lists, they are given a final overall recognition test in which studied items from the lists are mixed with the critical words (e.g., sleep) and other unrelated distractors. Studies on healthy young adults have found that the non-presented critical word is falsely recalled with relatively high probability (i.e., from 0.40 to 0.55 in different studies). At the recognition task, the false recognition rates of the critical words approximate the hit rate (i.e., rate of correct recognition) for studied items (i.e., from 0.80 to 0.85 in different studies). When asked to do a further "Remember/Know" judgment, in most cases (0.80 or more) participants claim to remember the presentation of the critical words as frequently as they do for studied (old) words (for a review, Gallo, 2010). According to the activation-monitoring account (Roediger, Balota, & Watson, 2001), false memories for critical words arises because of processes occurring at both encoding and retrieval: while participants are listening to the DRM list, the non-presented critical word may be mentally activated (either consciously or unconsciously), due to a spread of activation across the items within a semantic network, which converges on the non-presented critical representation. If the spreading activation process occurred for the critical item, at retrieval participants would be faced with a classic reality-monitoring problem (e.g., discrimination between sources of activation).

In this preliminary study, conducted in collaboration with the University of Florence and the University of Padua, possible semantic networks alteration in MS patients would be investigated by using the DRM paradigm. If alterations in the semantic network of MS patients are present, one might hypothesize that the activation of the critical word during the encoding phase was less likely, thereby resulting in a lower amount of false memories for critical words compared to healthy controls.

Materials and methods

Participants

Forty consecutive relapsing-remitting MS patients (RRMS; 28 F) and a matched control group of 40 healthy controls (HCs; 21 F) took part in the present study. Inclusion criteria for the experimental group were diagnosis of RRMS (Polman et al., 2011), no concomitant neurological or other pathological health conditions, no relapse in the last six months before the testing phase, no substance abuse or other concomitant medications. Inclusion criteria for HCs were absence of cognitive impairment, assessed through the Montreal Cognitive Assessment (MoCA) test (Santangelo et al., 2015), absence of neurological or psychiatric conditions, and no substance abuse or medications. RRMS and HCs were recruited at the Multiple Sclerosis Center of Verona University Hospital (Verona, Italy). Demographic and clinical characteristics of the participants are provided in Table 11.

	RRMS (=40)	**HCs (=40)**	***p***
Gender (M/F)	12/28	19/21	*0.17*
MoCA*	26.4±2.3	26.0±2.0	*0.43*
Age (years)	41.6±9.2	42.3±8.9	*0.74*
Education (years)	13.4±2.7	13.9±3.9	*0.52*
EDSS	2 (0-7)	/	
Disease duration (years)	9.0±8.4	/	

Table 11. Demographic and clinical characteristics of RRMS patients and HCs. Mean±SD was provided for continuous variables. Median±range was given for EDSS. (*) Raw scores. *P*-value with significant level of 0.05 was used.
MoCA = Montreal Cognitive Assessment; EDSS = Expanded Disability Status Scale

Each MS patient underwent a neurologic examination (including the evaluation of the expanded disability status scale, EDSS; Kurtzke, 1983) and a neuropsychological assessment consisting in the Italian version of the Montreal Cognitive Assessment test (MoCA; Santangelo et al., 2015), the Brief Repeatable Battery of Neuropsychological Tests (BRB-NT; Amato et al., 2006a) and the Stroop Test (ST; Caffarra et al., 2002). Tests' scores were classified as failed according to the cut-off scores (fifth percentile) derived from the Italian normative sample. Depression, anxiety, and stress status were evaluated by means of the Depression Anxiety Stress Scale (DASS-21; Bottesi et al., 2015).

The study was approved by the Verona University Ethics Committee and informed consent was obtained from all participants prior to their participation in the study.

Stimuli

The learning phase consisted of 4 lists, each composed of 12 semantically related words. The words of each list were strongly associated with a critical item not included in the list (the *critical lure*). For example, for the words "hot", "snow", "winter", "ice" etc., the critical item is "cold". The recognition phase consisted of other 4 lists, composed of 12 words each.

The 4 lists used for the learning phase showed a probability to recall the critical item of 52% (first list), 44% (second list), 34% (third list), and 42% (fourth list). Besides, the 4 lists used for the recognition task showed a probability to recognize the critical item of 60% (first list), 53% (second list), 35% (third list), and 42% (fourth list). The total number of words used during the recognition phase was 32: 12 from the lists previously presented (3 from each of the 4 lists, in particular in position 1, 5, 8 of the serial order of presentation), 4 critical items associated to the lists previously presented (1 from each of the 4 lists), 4 critical items associated to the lists not previously presented (1 from each of 4 lists), 12 new words belonging to the

lists not previously presented (4 weakly associated to the lists presented and 8 completely not related to the lists presented; all of them were presented in position 1, 5, 8 of the serial order of presentation of each list).

Procedure

Participants were seated in comfortable position and tested individually in a quiet room. During the learning phase, participants were asked to listen to an electronic male voice that pronounced the 12 words of each list (lapse of 2 seconds between each word). Immediately after, they were asked to recall orally as many words as possible within 60 seconds. The study-recall cycle was repeated for all the 4 lists of words. No feedback was given to the participants regarding the correctness of responses.

After the 4 cycles, there was a brief time interval (about 2 minutes) and then patients performed the recognition task: during this phase, the examiner read the list of 32 words, listed in a pseudorandomized way, and for each word participants had to say whether the pronounced word was already presented earlier or not. For each response, participants had also to express their confidence level on a scale ranging from 1 (low confidence) to 5 (high confidence).

The experimental session was audio recorded to allow an accurate off-line analysis of the patients' responses.

Statistical Analysis

Mann-Whitney *U*-tests were performed to investigate the difference between RRMS and HCs. In the recall task, the total number of critical false recalls, the total number of correct recalls, the total number of intrusions (non-critical false recalls), and the accuracy (proportion of correct recalls on the total number of recalls) served as dependent variables. In the recognition task, the number of correct recognitions (hits), the number of false recognition of critical words (false alarms to critical words), and the number

of false recognition of new words (false alarms to non-critical words, which includes: false recognition of new-non related words and false recognition of new-weakly related words) served as dependent variables. Confidence ratings for each of these variables also served as dependent variables in the recognition task.

Spearman correlation analyses were carried out between disease duration and significant dependent variables derived from the Mann-Whitney *U*-tests.

Results

MS patients and HCs did not differ in gender, age, education, and global cognitive performance (Table 11). No patients were considered as having emotional distress as assessed by the DASS-21 (all scores were not clinically relevant). Detailed results about neuropsychological performance of RRMS patients are shown in Table 12.

	RRMS (=40)	**Mean±SD**	**Patients scores below the cut-off**
Montreal Cognitive Assessment (MoCA)	MoCA total score	26.4±2.3	/
Brief Repeatable Battery of Neuropsychological Tests (BRB-NT)	SRT-LTS	45.2±12.9	3
	SRT-CLTR	34.6±15.4	5
	SRT-D	8.2±2.9	8
	SPART	21.5±5.2	0
	SPART-D	7.1±2.2	1
	SDMT	48.9±12.8	6
	PASAT-3	42.2±11.4	5
	PASAT-2	30.5±10.0	4
	WLG	25.3±5.6	2
Stroop Test (ST)	ST-EIT	14.3±7.5	2
	ST-EIE	0.4±0.6	0

Table 12. RRMS patients' neuropsychological performances. Mean±SD was provided for each test.
RRMS = Relapsing Remitting Multiple Sclerosis; SRT-LTS = Selective Reminding Test – Long Term Storage; SRT-CLTR = Selective Reminding Test – Consistent Long Term Retrieval; SRT-D = Selective Reminding Test – Delayed; SPART-I = Spatial Recall Test – Immediate; SPART-D = Spatial Recall Test – Delayed; SDMT = Symbol Digit Modalities Test; PASAT = Paced

Auditory Serial Addition Test; WLG = Word List Generation; ST - EIT = Stroop Test - Effect Interference Time; ST - EIE = Stroop Test - Effect Interference Error

Recall task

Results showed a significant difference between RRMS and HCs in the total number of critical false recalls ($p=0.009$, $r=0.29$): RRMS reported fewer critical false recalls (median±range=1±3) than HCs (median±range=2±4). No significant differences between the two groups were found in the other measures, i.e. the number of correct recalls ($p=0.67$), the number of intrusions (non-critical false recalls) ($p=0.99$), and accuracy ($p=0.21$) (Figure 13).

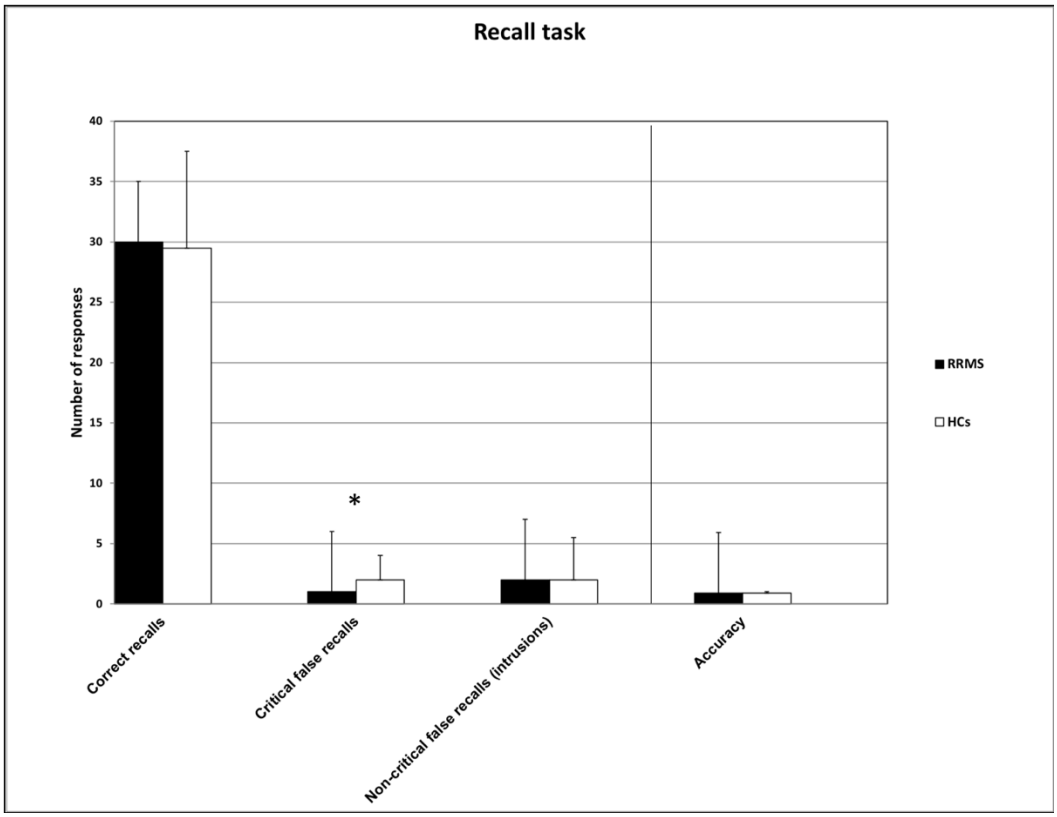

Figure 13. Median scores of responses in the recall phase of the DRM paradigm. Error bars represent range of scores. Black bars: RRMS patients; white bars: HCs.
*: $p < 0.05$
RRMS = Relapsing Remitting Multiple Sclerosis; HCs = Healthy Controls

No significant correlation between the total number of critical false recalls and the years of disease duration (p=0.28) was found.

Recognition task

No significant differences between the two groups were found in the number of correct recognitions (hits; p=0.22), the number of false recognitions of critical words (p=0.21), and the number of false recognitions of new words (p=0.96).

With respect to confidence ratings, results showed a significant difference between RRMS and HCs in the confidence ratings of false recognition of critical words (p=0.019, r=0.26): RRMS were less confident about the critical false recognitions (median±range=4.5±2) than HCs (median±range=4.75±4). No significant differences between the two groups were found in the other confidence ratings, that is of correct recognitions (hits; p=0.63) and false recognition of new words (p=0.81) (Figure 14).

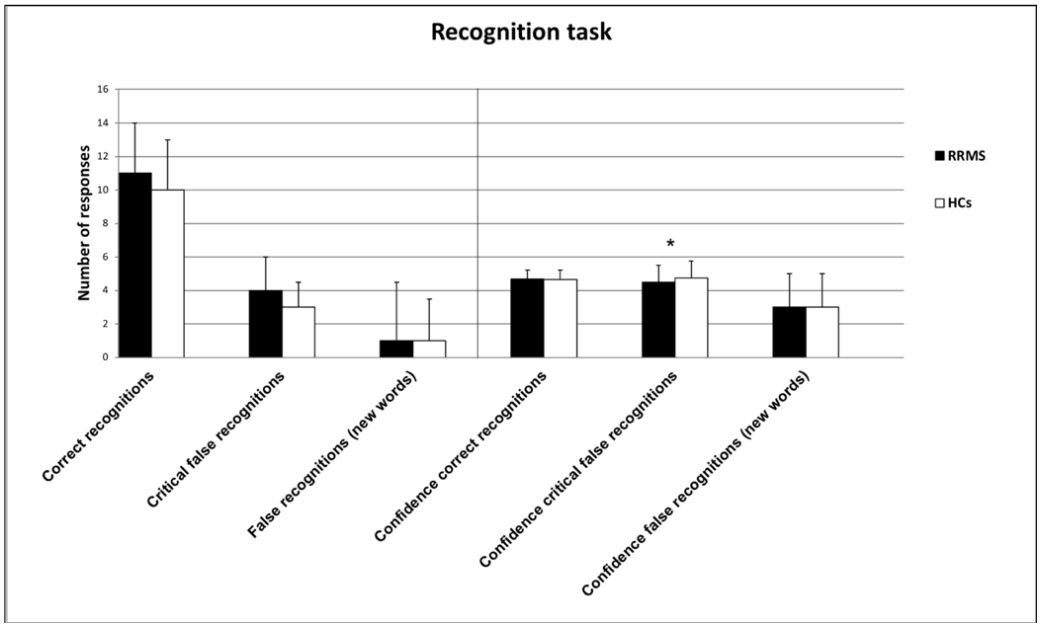

Figure 14. Median scores of responses in the recognition phase of the DRM paradigm. Error bars represent range of scores. Black bars: RRMS patients; white bars: HCs.
*: $p < 0.05$
RRMS = Relapsing Remitting Multiple Sclerosis; HCs = Healthy Controls

No significant correlation between the confidence ratings of false recognition and the years of disease duration ($p=0.90$) was found.

Subgroup analysis

The group of RRMS patients included some patients with verbal memory or semantic memory deficits. As a matter of safety, since 10 RRMS patients performed below the cut-off in the test of verbal memory (i.e., the SRT) and 2 RRMS performed below the cut-off in the test of semantic fluency (i.e., the WLG), the same analyses were ran excluding these 12 patients (new n=28). The two groups (RRMS vs. HCs) did not differ in age ($p=0.54$) and education ($p=0.91$). At the recall task, the results confirmed a significant difference in the total number of critical false recalls between the two groups ($p=0.024$, $r=0.27$). Even at the recognition task, the results confirmed a significant difference in the confidence ratings of false recognition of critical words ($p=0.029$, $r=0.26$) between the two groups. No other comparisons reached the statistical significance (all $p>0.05$).

Discussion

The present study aimed at investigating veridical and false memory for associatively/semantically related material in MS patients, by using the Deese-Roediger–McDermott false memory paradigm (DRM; Deese, 1959; Roediger & McDermott, 1995). The results showed that the group of RRMS patients were less likely to falsely recall highly related but non-presented items (critical false recalls) than the group of HCs were. Moreover, at the recognition task, RRMS patients showed a reduced level of confidence for false recognitions of critical items.

This is the first report describing memory distortions, and specifically false memories for semantically-related words, in MS patients. The most interesting finding of this study is the reduced susceptibility to this kind of false memories in RRMS patients compared to HCs.

Such an effect could be due to alterations in the semantic memory networks in MS patients. To this regard, a recent study by Abad et al. (2015) has shown that MS patients are characterized by an alteration in the connectivity of semantic networks that involves reduced number of nodes (words) and links (significant associative relationship between words), besides reduced cohesive force of the network. Based on the functional efficacy of semantic networks (Goñi et al., 2011), a higher number of nodes means more information stored in the semantic memory, while a higher number of connections between nodes represents an indicator of the ability to retrieve information.

In the light of these findings, it might be suggested that the reduced number of critical false recalls found in the group of RRMS patients might arise from alterations in the structure of semantic memory. In RRMS patients, the alteration of semantic networks might interfere with the activation of semantic associates from the list words during the encoding phase, thereby reducing the likelihood of critical false recalls.

In the recognition task, the two groups did not differ in the number of false recognitions for critical items but in their level of confidence: RRMS patients were less confident in their false recognitions for critical items than HCs. Nevertheless, the two groups exhibited similar levels of confidence in their correct recognition responses to studied items. Previous studies have shown that normal individuals experience false recognitions for critical items accompanied by *strong* feelings of recollection, as indicated by a high level of confidence (for a review, Roediger & McDermott, 2000). The reduced level of confidence for false recognitions reported by RRMS patients provided further support of a less robust false remembering in patients compared to controls.

The reduced susceptibility to false memories for critical words shown by RRMS patients could not be ascribed to an impairment in verbal memory, since the same pattern of results arose when the analyses were limited to the

subgroup of patients without deficits in verbal memory and semantic fluency as assessed with the SRT and the WLG, respectively.

However, some limitations of the present study should be considered when interpreting the results. First, the sample size was relatively small, considering the high heterogeneity of clinical characteristics of MS patients: therefore, studies with higher number of MS patients, stratified considering MS type and other neurological characteristics, e.g., disease duration and disease type, are needed. Second, this study did not include analysis of neuroimaging data: future research will incorporate advanced structural and functional MRI outcomes which could shed light on the neuroanatomical organization that promotes semantic networks alteration associated with false memories. Third, both MS patients and HCs underwent, in addition to the false memories protocol, also a screening cognitive evaluation by means of MoCA test: however, only MS patients underwent a comprehensive neuropsychological evaluation by means of BRB-NT and the Stroop Test. In future studies also HCs will need to be administered with a complete neuropsychological battery, in order to have the possibility to compare between the two groups not only the false memories phenomenon but also the cognitive impairment level. Some evidence about the comparison of cognitive functioning between MS and HCs have been described in the previous chapter (1.1).

Despite these limitations, the present study provides a first exploratory contribution to understanding of semantic memory mechanisms and distortions shown by RRMS patients, a topic that has been neglected for a long time and that might have important implications also in clinical setting. In this regard, although it is well known that MS is characterized by neurodegeneration and demyelination, subtle cognitive changes and initial alterations in semantic networks might not be evident in the classic neuropsychological examination. Including the DRM paradigm in the cognitive assessment of MS patients might provide more sensitive

information about the functioning of semantic memory networks that might be associated with the course of degenerative processes underlying the disease.

2.2 SOCIAL COGNITION AND THE ROLE OF AMYGDALA IN RRMS PATIENTS WITHOUT COGNITIVE IMPAIRMENT

Despite a wealth of research investigating the role of traditional domains of CI in MS patients, fewer studies have been performed related to deficits in social cognition, i.e., mental operations underlying social interactions and processes required to establish/sustain interpersonal relationships (Dulau et al., 2017). Social cognition (SC) is a multidimensional construct that involves, among other aspects, theory of mind (ToM), facial emotion recognition, and empathy (Chalah & Ayache, 2017). Emerging evidence suggests that individuals with MS have significant deficits in this domain (Genova et al., 2016; Bora et al., 2016; Cotter et al., 2016). However, an inhomogeneity in social cognitive performance has been observed across MS studies, which may be due to confounding factors, such as fatigue, mood, alexithymia, and cognitive impairment (Chalah & Ayache, 2017). Among these factors, one variable that is often overlooked when studying SC is controlling for CI. Several studies found significant correlations between cognitive measures and ToM tasks (Henry et al., 2009; Kraemer et al., 2013; Dulau et al., 2017; Raimo et al., 2017; Ciampi et al., 2018), as well as facial emotion recognition (Ouellet et al., 2010; Berneiser et al., 2014; Cecchetto et al., 2014;) and empathy (Kraemer et al., 2013). In contrast, other studies showed no correlation between CI and SC deficits in MS patients (Banati et al., 2010; Pottgen et al., 2013; Phillips et al., 2014; Batista et al., 2017c; Neuhaus et al., 2018). Among the different cognitive domains that can be involved in SC deficits, executive functions (EF) can be frequently associated with social behavior problems but, at the same time, can also occur independently of the level of EF performance (e.g., Batista et al., 2017c). Previous studies in patients with neurological disorders have shown variable results, either supporting an association or a dissociation between EF and ToM (Aboulafia-Brakha et al., 2011); also in MS patients,

the nature of the relationship between EF and ToM is still debated (Chalah & Ayache, 2017).

One way to clarify the abovementioned issues would be to study individuals with MS who are not experiencing CI, in order to examine if SC deficits can occur even in the absence of CI. Although brain MRI measures (i.e., global and regional GM volume, and cortical thickness) have been widely reported as being correlated with cognitive performance (Rocca et al., 2015; Rimkus et al., 2019), only a few recent studies have begun to explore this behavioral-neuroradiological relationship in SC, including the involvement of bilateral amygdala and associated frontal, temporal, and parietal brain areas (Batista et al., 2017b; Chalah & Ayache, 2017) as well as diffuse WM damage (Batista et al., 2017a). Moreover, the role of the amygdala in SC in MS has not been well established and needs further investigation.

The aim of the present study, conducted in collaboration with the Kessler Foundation (West Orange, NJ, USA), was to examine whether deficits in SC are present in a group of relapsing-remitting MS patients (RRMS) without CI, with a specific interest in the role of bilateral amygdala, assessed with non-conventional imaging techniques.

Materials and methods

Participants

The present study enrolled 31 RRMS patients recruited at the MS Center of Verona University Hospital (Verona, Italy) and a healthy controls (HCs) group of 38 matched on age, gender, and education. Demographic and clinical characteristics of the participants are provided in Table 14.

The patients were under immunomodulatory therapy: 19 were treated with dimethylfumarate, 5 were treated with high doses of interferon beta-1a, one patient was treated with glatiramer acetate and 6 were treated with fingolimod. Inclusion criteria for the MS group were diagnosis of RRMS

according to the most recent diagnostic criteria (Polman et al., 2011), no concomitant neurological or other pathological health conditions, no relapse in the last six months before the testing phase, no substance abuse or other MS concomitant medications (as benzodiazepines or antidepressant drugs), no visual impairment, and absence of cognitive impairment (CI). Cognitive status was evaluated by means of the Brief Repeatable Battery of Neuropsychological Tests (BRB-NT) (Amato et al., 2006a) and the Stroop test (Caffarra et al., 2002). Depression, anxiety, and stress disorders were evaluated by means of the Depression Anxiety Stress Scale (DASS-21; Bottesi et al., 2015). To be considered cognitively intact, each MS patient had to achieve scores above the cut-off score (5th percentile) on each subtest of the BRB-NT and on the Stroop test. Table 13 shows RRMS patients' performance on neuropsychological tests (*Z*-scores) on each BRB-NT subtest and the Stroop test.

BRB-NT *Z*-scores range

Cognitive test	less than 2 SD	between -2 and -1 SD	between -1 and 0 SD	more than 0 SD	Mean (SD) Z-scores
SRT-LTS	0	2	8	21	0.33 (0.84)
SRT-CLTR	0	2	11	18	0.22 (0.95)
SRT-D	0	1	6	24	0.37 (0.73)
SPART	0	1	4	26	0.66 (0.78)
SPART-D	0	0	8	23	0.53 (0.66)
SDMT	0	0	9	22	0.54 (0.79)
PASAT-3	0	1	11	19	0.14 (0.73)
PASAT-2*	0	3	14	13	-0.04 (0.83)
WLG	0	3	16	12	-0.08 (0.74)

Stroop test Equivalent Scores (ES)°

	ES 0	ES 1	ES 2	ES 3	ES 4
Time	0	0	1	15	15
Errors	0	3	4	2	22

Table 13. Number of patients according to the classification of *Z*-scores on each subtest of the BRB-NT. *=one missed value. °Equivalent Scores (ES) is a 5-point (range: 0-4) scale that offers a solution to the problem of standardizing neuropsychological scores after adjustment for age and education when the common *Z*-standardization is not applicable. The cut-off score is *zero*.

SD = Standard Deviation; SRT-LTS = Selective Reminding Test – Long Term Storage; SRT-CLTR = Selective Reminding Test – Consistent Long Term Retrieval; SRT-D = Selective Reminding Test – Delayed; SPART-I = Spatial Recall Test – Immediate; SPART-D = Spatial Recall Test – Delayed; SDMT = Symbol Digit Modalities Test; PASAT = Paced Auditory Serial Addition Test; WLG = Word List Generation

The inclusion criteria for HCs consisted in absence of CI assessed through the Montreal Cognitive Assessment (MoCA; Santangelo et al., 2015), absence of neurological or psychiatric conditions, substance abuse or medication, and normal or corrected to normal vision.

Social cognition was examined with the Reading the Mind in the Eyes test (RME; Baron-Cohen et al., 2001), a task of facial affect recognition (FAR), and the Empathy Quotient questionnaire (EQ; Baron-Cohen & Wheelwright, 2004).

The study was approved by the local Ethics Committee and the informed consent was obtained from all individual participants included in the study.

Reading the Mind in the Eyes test (RME) – Stimuli and procedure

The Italian version of the RME test (Vellante et al., 2013) was used to assess ToM. The RME test consists of 36 black-and-white images representing only the horizontal portion of faces showing the eyes of different persons (17 female faces and 19 male faces). The participants are presented with an image to view for 5 seconds; afterwards, 4 emotion adjectives appeared on each corner of the image. Participants were asked to choose which of the 4 emotion adjectives best described the emotion transmitted by the eyes in the image. As a control task, participants were also asked to indicate whether each pair of eyes was that of a female or a male face. The emotion adjectives provided the meaning of a belief or an intention with the same valence, e.g., “serious”, “ashamed”, “alarmed”, and “bewildered”. For each image, a correct response received a score of 1 point and no points for incorrect responses.

Facial Affect Recognition (FAR) - Stimuli and procedure

In the FAR task, the participant is presented with a series of 36 black and white photos of different actors displaying one of six emotions: happiness, sadness, anger, surprise, fear, and disgust. The stimuli were selected from the Karolinska Directed Emotional Faces (KDEF) database (Lundqvist, Flykt & Öhman, 1998; Goeleven et al., 2008), which has been well validated. Each emotion is presented randomly six times. Participants then decided which emotion each face is expressing, and their oral response was recorded. The total test score on this measure was calculated as total items correctly identified.

Empathy Quotient (EQ) - Stimuli and procedure

The EQ questionnaire (Baron-Cohen & Wheelwright, 2004) was administered to the participants to assess their level of empathy. The EQ is a self-report questionnaire consisting of 60 items (40 about empathy and 20 about other domains) (Lawrence et al., 2004). For each item, a 4-point Likert-scale is provided to measure the negative or positive empathic valence of the sentences; for the 40 empathy questions, 2 points were assigned for the extreme answers (i.e., “totally disagree” [score 1] and “totally agree” [score 4]) and 1 point for intermediate answers (i.e., “partially disagree” [score 2] and “partially agree” [score 3]). No points were assigned to the 20 non-empathic control questions. The maximum score is 80: higher scores indicate higher empathy. This measure has been largely used in MS and it has been shown to be sensitive to measuring reduced empathy in MS (Kraemer et al., 2013).

MRI acquisition and imaging analysis

Each patient was scanned at enrolment using a 3.0 Tesla Philips Achieva MRI (Philips Medical Systems, Best, The Netherlands). No major hardware upgrades were carried out on the MRI scanner during the study

and weekly quality assurance sessions took place to guarantee measurement stability. The following sets of images were acquired:

- 3D Fluid Attenuated Inversion Recovery (FLAIR) TR/TE=5500/292 ms, TI=1650 ms, voxel dimension of 1x1x1 mm, matrix=256x256;
- 3D Double Inversion Recovery (DIR) TR/TE=5500/292 ms, TI1/TI2=525/2530 ms voxel dimension of 1x1x1 mm;
- 3D T1 weighted Fast Field Echo (FFE) TR/TE=8.4/3.7 ms, voxel dimension of 1x1x1 mm, matrix=256x256;
- 2D T1w SE GD: TR/TE=550/10 ms, 50 contiguous axial slices with a thickness=3.0 mm, matrix=256x256.

Regional cortical thickness and volume evaluation

Global cortical thickness and cortical volume of amygdalae were calculated on the volumetric T1 weighted data set, using the FreeSurfer image analysis suite (release v5.3.0), available online (http://surfer.nmr.mgh.harvard.edu). Topological defects in cortical surfaces due to white matter lesions were corrected using a semi-automated procedure with lesions filling. Each MRI has been evaluated by a neurologist well trained and experienced in MS, blinded to patients' behavioral results.

Lesions number and volume

The total number of CLs and the number of CLs located in the amygdala were assessed on DIR images and confirmed on T1 weighted images, following the recent recommendations for CLs scoring in patients with MS (Geurts et al., 2011) – Figure 15.

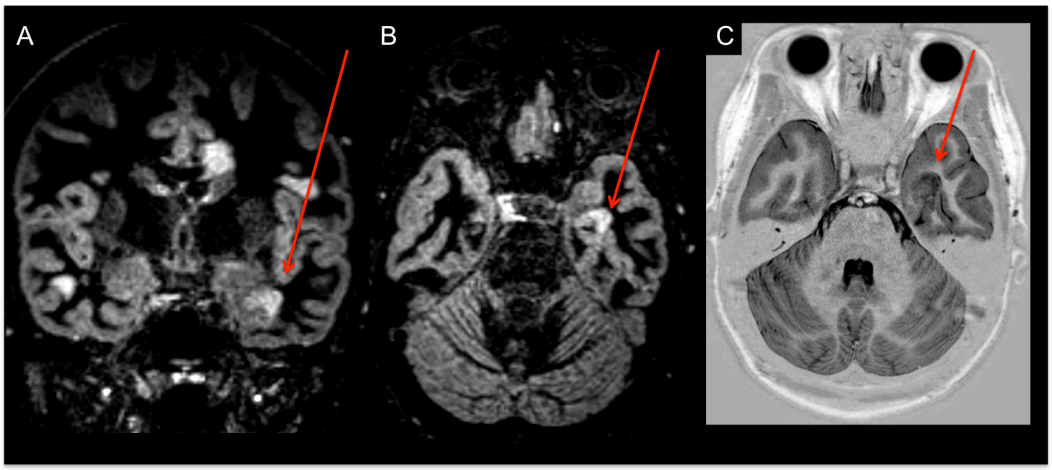

Figure 15. Coronal (A) and axial (B) double inversion recovery (DIR) sequence of a RRMS patient showing left amygdala lesion (red arrows). The same lesion is also depicted as hypointensity in the phase sensitive inversion recovery (PSIR) sequence (C).

Owing to the suboptimal performance of the image-acquisition sequences on MRI in visualizing subpial lesions, the present analysis has considered mainly the intracortical and leukocortical lesions.

Cortical lesion volume (CLV) was calculated using a semiautomatic thresholding technique based on Fuzzy C-mean algorithm (Pham & Prince, 2000) included in software developed at the National Institutes of Health (NIH), Medical Images Processing, Analysis and Visualization (MIPAV) (http://mipav.cit.nih.gov). The same procedure was applied to FLAIR images to calculate the white matter lesion load.

Statistical analyses

Mann-Whitney *U*-tests were applied to test the differences between RRMS patients and HCs on the dependent variables (i.e., RME, FAR task, EQ). Chi-square test was applied to test female/male ratio between groups. Partial correlation analysis with Bonferroni correction, with age as covariate, was applied among imaging parameters and the dependent variables. Multiple linear regressions, weighted for age, were applied to assess the predictive role of imaging parameters on each of the dependent variables. Effect sizes for group comparisons were also calculated.

Results

Clinical and demographic characteristics of RRMS patients and HCs are provided in Table 14.

	RRMS (n=31)	**HCs (n=38)**
Gender (M/F)	7/24	10/28
Age (years)	36.3±7.6	37.1±8.9
Education (years)	13.4±3.4	14.6±3.4
EDSS	1±3.5	/
Disease duration (years)	7.0±4.5	/

Table 14. Clinical and demographic characteristics of experimental (RRMS) and control (HCs) groups. Mean±SD was provided for continuous variables; Median±range was provided for the Expanded Disability Status Scale (EDSS). Range is expressed as the largest value minus the smallest value.

There were no significant differences between the two groups on age (U=562, p=0.74), education (U=463, p=0.12), and gender ($\chi^2(1)=0.09$, p=0.78). No significant correlations were found among depression, anxiety, and stress disorders (DASS-21 scale) and the RME, the FAR task and the EQ (all $p>0.05$). Cognitive performance (Z-scores) in the MS group did not correlate significantly with the RME, the FAR task, and the EQ (all $p>0.05$). There were no significant correlations between the EDSS and any of the neuropsychological variables examined (all $p>0.05$).

RME

The number of correct responses on the RME served as the dependent variable. The RRMS group had significantly fewer correct responses on the RME (median±range=26±16) than the HCs group (median±range=28±11) ($U=293.5$, $p<0.001$, $\eta^2=0.19$) - Figure 16.

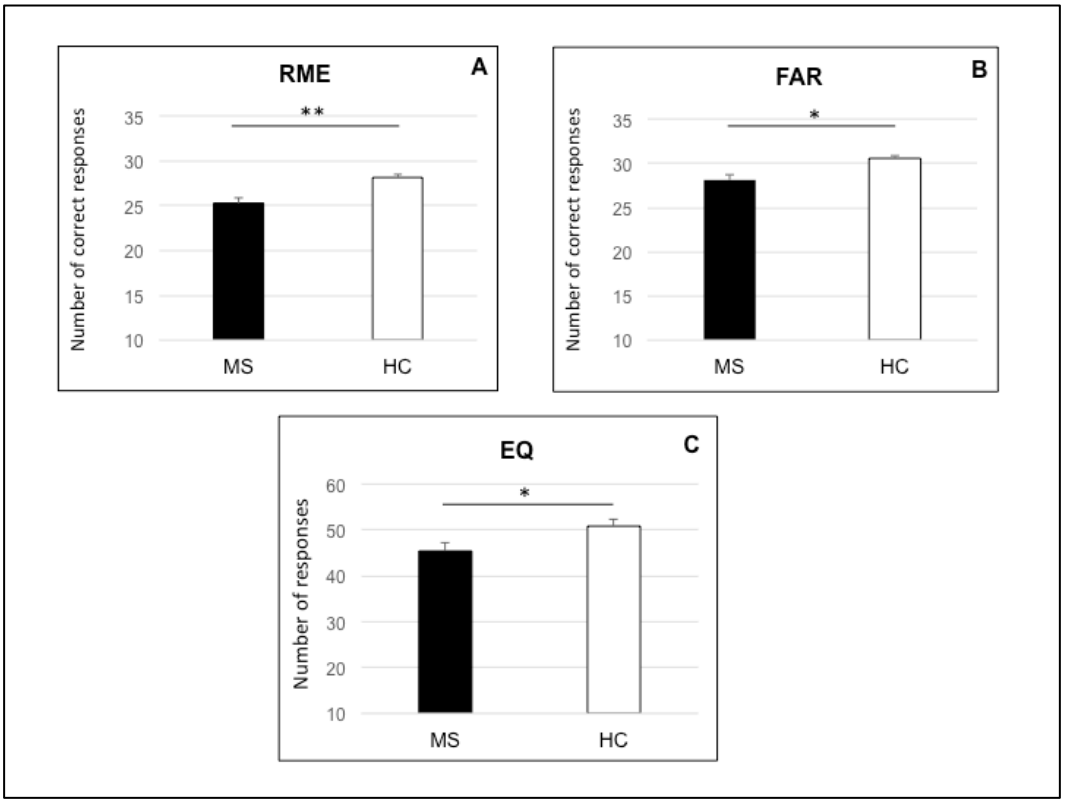

Figure 16. Mean scores of correct responses on the RME (box A), FAR (box B), and EQ (box C) tasks. Error bars represent the standard error of the mean (SEM). Black bars: MS patients; white bars: healthy controls.
*: $p < 0.05$; **: $p < 0.001$

FAR Task

The number of correct responses on the FAR task served as the dependent variable. In the FAR task, there was a significant difference between RRMS and HCs groups, (U=311.5, p=0.001, η^2=0.17); the RRMS group had fewer correct responses (median±range=29±13) than the HCs group (median±range=31±8) – Figure 16. Specifically, the RRMS group had significantly fewer correct responses compared to the HC group in recognition of fear (RRMS: median±range=2±5; HCs: median±range=4±5; U=275.5, p<0.001, η^2=0.22), and anger (RRMS: median±range=5±4; HCs: median±range=5±4; U=423.5, p=0.036, η^2=0.06) - Figure 17.

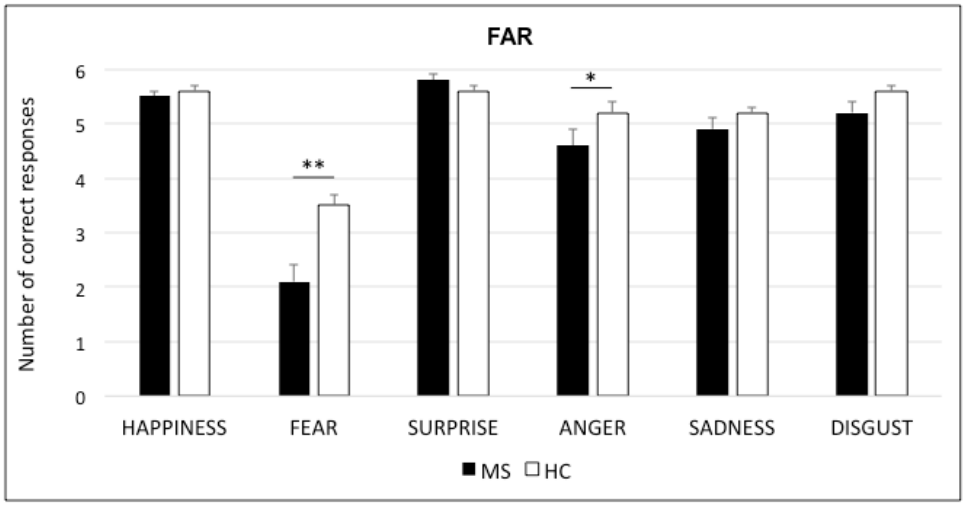

Figure 17. Mean scores of correct responses by single emotions on the FAR task. Error bars represent the standard error of the mean (SEM). Black bars: MS patients; white bars: healthy controls.
*: $p < 0.05$; **: $p < 0.001$

EQ

The EQ total served as the dependent variable. The RRMS group scored significantly worse than the HCs group on the EQ score (U=399, p=0.022, η^2=.08) with the RRMS group scoring lower (median±range=45±36) than the HCs group (median±range=51±47) - Figure 16.

Lesion volume of bilateral amygdala

The cortical lesion volume (CLV) of amygdala (left and right), the number of cortical lesions (CLs) in the whole brain, the brain cortical thickness (CTh), and the brain white matter lesion volume (WMLV) were considered as predictors of RME, FAR task, and EQ in the RRMS subjects. Partial correlation analysis showed a significant correlation between amygdala CLV and the RME, $r(28)$=-0.611, p<0.001 (the greater CLV, the lower the correct responses on the RME), and between amygdala CLV and the FAR task, $r(28)$=-0.462, p=0.010 (the greater CLV, the lower the correct responses on the FAR task). There was no significant correlation between

amygdala CLV and the EQ ($r(28)=0.018$, $p=0.924$). Given the extensive literature linking fear and anger with amygdala, fear and anger results were combined and compared them with the combination of the results from the other emotions (happiness, surprise, sadness, disgust) in relation to amygdala CLV. A significant correlation was found between amygdala CLV and fear/anger ($r(28)=-0.435$, $p=0.016$), but not between amygdala CLV and the other emotions ($r(28)=-0.349$, $p=0.059$). A multiple regression analysis including all six emotions as predictors showed that fear was the only significant predictor of bilateral amygdala CLV ($R^2=377$, $t(30)=-2.207$, $p=0.037$).

To exclude a possible confounding role of global brain impairment on the social cognition tasks, separate multiple regression analyses were run with the amygdala CLV, the brain CLs, the brain CTh, and the brain WMLV as predictors of RME, FAR task, and EQ. The results showed that bilateral amygdala CLV was the only significant predictor of RME ($R^2=476$, $t(30)=-4.154$, $p<0.001$) and FAR task ($R^2=287$, $t(30)=-3.04$, $p=0.005$). Brain CLs was the only significant predictor of EQ ($R^2=157$, $t(30)=-2.163$, $p=0.040$).

Separate multiple regression analyses were also run with cognitive measures as predictors of RME, FAR task, and EQ. The results showed that none of the cognitive measures predicted significantly the social cognition variables (all $p>0.05$).

Discussion

The present results showed that RRMS patients manifested specific impairment in SC tasks compared to matched HCs and that SC impairment can occur even in the absence of CI. These results are consistent with previous studies that did not find any association between CI and SC deficits in MS (Banati et al., 2010; Pottgen et al., 2013; Phillips et al., 2014; Batista et al., 2017c; Neuhaus et al., 2018). The RRMS group performed significantly worse on the RME test compared to the control group,

corroborating the hypothesis that RRMS have a specific impairment in the domain of ToM, which can occur independently of the level of global cognitive performance. In the current cohort of RRMS patients, performance on the RME test correlated significantly with the bilateral amygdala CLV and the CLV was the only significant predictor of RME scores; specifically, the higher the total score on the RME, the lower the CLV value. It has been recently reported that amygdala volume was significantly lower also in pediatric onset MS (POMS) compared with controls and that the left amygdala volume was a strong predictor of parent-reported social skills in POMS (Green et al., 2018). These findings suggest that amygdala damage might play a central role in SC deficits in both adults with MS and POMS.

In the domain of facial affect recognition, the RRMS group scored significantly lower than the control group, specifically in recognition of anger and fear, as has been shown in other studies assessing MS patients (Henry et al., 2009; Ciampi et al., 2018). Also, facial affect recognition correlated significantly with bilateral amygdala CLV and that the amygdala CLV was the only significant predictor of FAR task scores (the higher the total score on the FAR task, the lower the CLV value). This is the first study in which it has been shown a correlation between bilateral amygdala CLV and deficits in FAR, specifically related to anger and fear, in a group of RRMS patients without CI, adding evidence on the role of bilateral amygdala in recognizing basic emotional states. This result underlines the key role of amygdala in the domain of SC also in MS (Batista et al., 2017b; Green et al., 2018).

Finally, empathy, as assessed with the EQ, was significantly lower in the RRMS group compared with HCs. RRMS patients defined themselves as less empathic and less prone toward sharing emotional states of other persons and in sympathizing towards the other's situations, in agreement with previous studies (Banati et al., 2010; Kraemer et al., 2013). The amygdala may play a lesser role regarding empathy than its role in fear and anger, as several studies have related empathy more with the anterior insula,

and that empathic behavior can be modulated by various factors, such as those related to individual traits and situational contexts (Lenne et al., 2014). However, brain CLs were a predictor of EQ score, which is consistent with recent studies suggesting that a distributed network involving multiple cortical and subcortical regions may underlie the experience of empathy (Nomi et al., 2008; Bernhardt & Singer, 2012).

There are several limitations in the current study. First, the control group of HCs did not underwent an MRI, so comparisons with the RRMS group on brain parameters could not be conducted. Further studies are required to collect also MRI data of a matched control group. Second, HCs did not undergo neuropsychological assessment except with MoCA, so CI classification of MS patients, although conservative, was based on Italian normative data and not with respect to the HCs group. Another limitation is that SC deficits in MS have only recently been investigated, and thus there are no measures that have been validated in this population. A recent meta-analysis (Cotter et al., 2016) suggests that investigations of SC in MS have used a wide number of SC tests, most validated in other populations such as Autism Spectrum Disorder, but validation and standardization of measures to clinically investigate SC aspects in individuals with MS should be the target of future research. In the current study, was utilized a database of facial affect stimuli that has been well validated in the general population (Lundqvist, Flykt & Öhman, 1998; Goeleven et al., 2008). Moreover, future studies will be conducted by also considering the social functioning in everyday life of patients, in order to evaluate the reflection of SC deficits in daily living of MS patients and in their quality of life, even in the absence of formal CI, and strengthen the validity of the results presented. Furthermore, future research should be conducted by implementing "dual-task" conditions: this is an emerging paradigm which has been increasingly studied over the past few years, that could represent a more ecological method to identify issues related to social cognitive processes in MS.

Conclusions

In summary, the results of the present study showed that SC can be impaired in several domains in RRMS patients even in the absence of CI. The lower performance on RME and FAR task was related specifically to bilateral amygdala damage as measured by CLV, and not to global brain impairment. This is the first study investigating three different aspects of SC with both behavioral and structural imaging data in the same group of RRMS patients without CI. These results point to the prominent role of amygdala in SC deficits in persons with MS, in particular fear and anger. It could be argued that bilateral amygdala, as a primitive subcortical area and as a fundamental region for the "social brain" (Bickart, Dickerson & Barrett, 2014), in RRMS might be related to SC deficits also in the absence of CI.

The results of the current study reinforce the need of further investigating SC earlier in the course of the disease in both adults with MS and POMS (Charvet et al., 2018; Green et al., 2018), not only for clinically related aspect of MS, but also because of its influence on social aspects of quality of life of MS patients and their families, such as the ability to regulate emotions (Phillips et al., 2011). As such, SC should be assessed independently of the presence of CI.

PROJECT 3 – REHABILITATION OF COGNITIVE FUNCTIONING IN MS

CLINICAL, CONVENTIONAL MRI AND ADVANCED MRI PREDICTORS OF EFFECTIVENESS OF COGNITIVE REHABILITATION IN MS PATIENTS

Considering the fact that cognitive alterations are highly frequent in MS patients since the very early stages of the disease, it is of paramount importance to intervene with specific rehabilitation approaches in order to improve cognitive functioning. Cognitive rehabilitation (hereafter CR) refers to a specific process based on behavioral interventions aimed at improving cognitive functioning, awareness and quality of life in individuals with cognitive deficits due to brain injuries or neurodegenerative conditions (Mateer, 2005; Mitolo et al., 2015). CR should be considered and offered to all patients suffering from long-term problems, and not only to patients that are expected to recover completely (Wilson & Betteridge, 2019). However, individuals respond in various ways and at different times to various types of cognitive interventions: premorbid functioning, personality, social support, and environmental demands are some of the factors that can profoundly influence rehabilitative outcomes (Sohlberg & Mateer, 2001 & 2017).

CR can be divided into two major categories: compensatory and restorative. Compensatory approaches aim to adjust and modify behavioral patterns and styles of thinking to compensate for cognitive deficits. Otherwise, restorative approaches are designed to strengthen impaired cognitive functions through repeated exercises and similar forms of training (Sandry et al., 2016). Restorative interventions may be suitable for patients at the early stages of the disease, while compensatory interventions may be more appropriate for patients with severe cognitive impairment. Besides the type of rehabilitation, it is important to offer treatments that are suitable for

patients with particular needs and capacities to benefit for a specified approach.

The first studies describing CR programs for patients with MS were used to focus on learning and memory problems, but recently most of the interventions were centered on problems regarding attention, executive functions and information processing speed domains (Goverover et al., 2018). In particular, slowed cognitive processing speed, the most frequent cognitive deficit in patients with MS (Benedict et al., 2017a), is associated with decrements in quality of life, employment, and independence of daily activities (Benedict et al., 2017b; Sumowski et al., 2018). Fortunately, early data suggest that CR interventions are efficacious in enhancing processing speed in patients with MS, relative to control interventions (Goverover et al., 2018). One recent study demonstrated the efficacy of a restorative telerehabilitation approach in a randomized controlled trial (Charvet et al., 2017). An adaptive software was used (BrainHQ, Posit Science Corporation), that is affordable, easy to use, and can be implemented on home computers, tablets, and laptops, being therefore accessible to people with limited mobility or other constraints (Charvet et al., 2017). The study showed a significant effect of the cognitive training measured by a composite score derived from neuropsychological tests, as compared with the active control group who underwent ordinary computer games. Despite the moderate success, the change in group means overlay considerable variability in treatment response across participants. For instance, the mean improvement in the treatment group ($z=0.25$) was smaller than the standard deviation (0.45).

In order to better investigate which patient is best suited for different treatments of CR, a research project was carried out at the Buffalo Neuroimaging Analysis Center (NY, USA) between 2017 and 2018, aimed at identifying baseline variables that presage favorable outcome with this same approach. Since providers and patients alike would benefit from data that

would inform the goodness of fit between patients and symptomatic therapies (Sumowski et al., 2018), in this pilot, single-arm repeated measures study, demographic, clinical, neuropsychological, both conventional and non-conventional brain MRI, and brain functional MRI variables were measured, in order to explain variability in individual patient response to this form of restorative rehabilitation in a real-world outpatient setting.

Materials and methods

Participants

Fifty-four MS patients were recruited for the present study in Buffalo, (NY, USA) between 1 November 2017 and 30 April 2018 from a sub-sample of eligible subjects previously enrolled in a larger Cardiovascular, Environmental, and Genetics (CEG) study (Zivadinov et al., 2016a). A sub-group of 25 MS patients also underwent resting-state functional MRI (rs-fMRI) at baseline. Of the total sample, 3 subjects were lost to follow-up. Subjects were provided free access to a CR software and financial compensation for participation and completion of the program. Sample size was determined with a power analysis relating to effect size reported in a previous cognitive training study (Studer-Luethi et al., 2012).

Inclusion criteria were diagnosis of clinically definite MS (Polman et al., 2011), age 18+, English fluency, able to provide informed consent to all procedures, access to internet and home computer, computer competency, and high-quality structural MRI available from the CEG study. Participants were excluded based on the following criteria: other neurological/psychiatric disorder, alcohol/substance abuse, relapse/acute corticosteroid treatment within 1 month of testing, color-blindness, and motor or sensory defect that might interfere with cognitive test performance (e.g. corrected near vision of at least 20/70).

Cognitive impairment was not an inclusion criterion for this study for two main reasons: firstly, baseline cognitive status was intended as a candidate predictor of pre- to post-intervention change and so heterogeneity in patients' cognitive functioning was one key aim of the study. Secondly, individuals may have cognitive impairment relative to their own pre-disease status but not relative to average healthy populations. These individuals might be bothered by their cognitive disturbances and still seek out CR. Individuals were therefore considered for the study if interested in the service. Additional information about eligibility, enrollment, and loss to follow-up, is provided in the study flowchart (Figure 18).

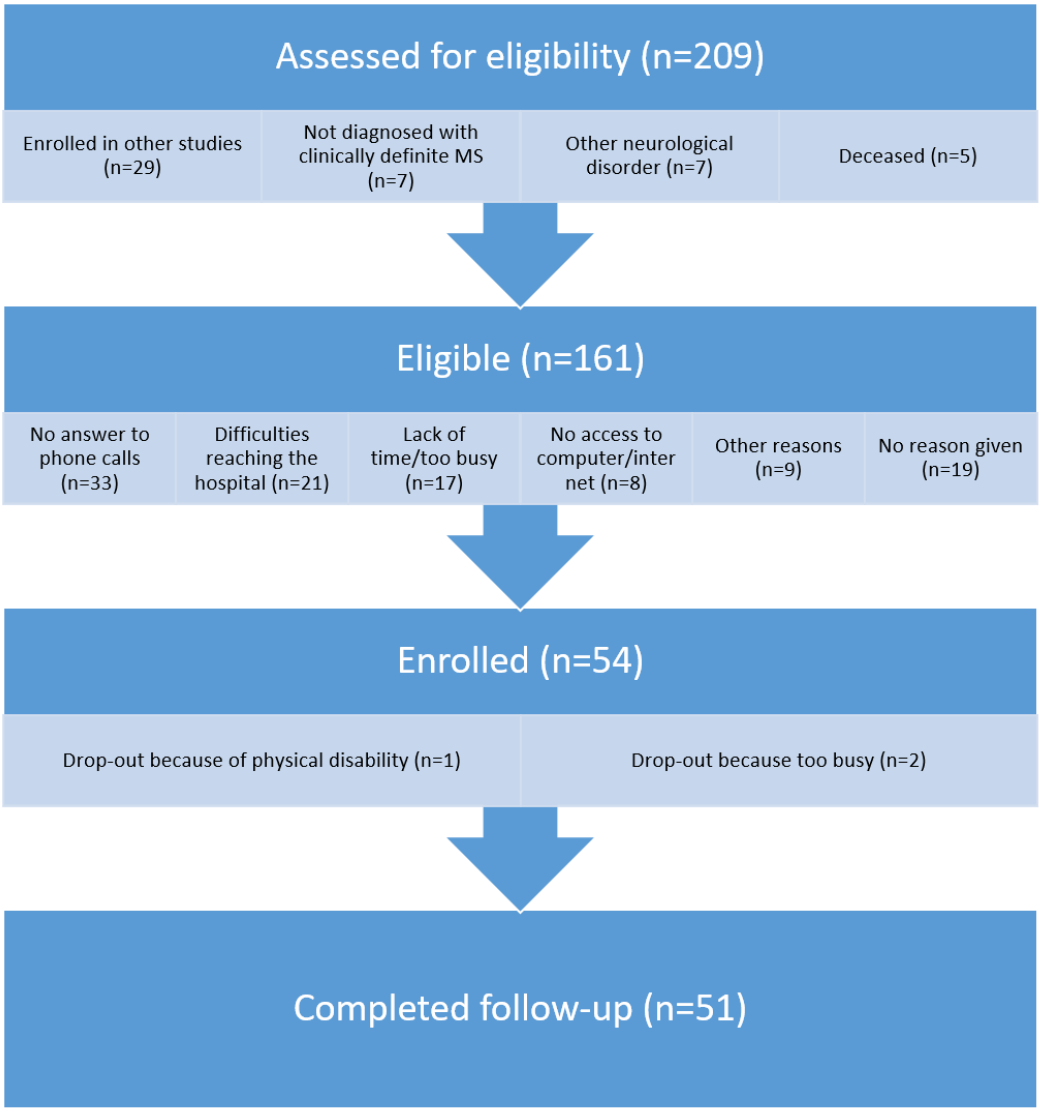

Figure 18. Study flowchart: eligibility, enrollment, and loss to follow-up.

Study design

The primary aim of the study was to explore the baseline factors associated with individual patient variability in response to treatment rather than to replicate a treatment effect relative to control (Charvet et al., 2017). For this reason, a single-arm repeated measures study design was applied using the same intervention structure as in the previous randomized-controlled trial (Charvet et al., 2017). Figure 19 shows a schematic of the study timeline.

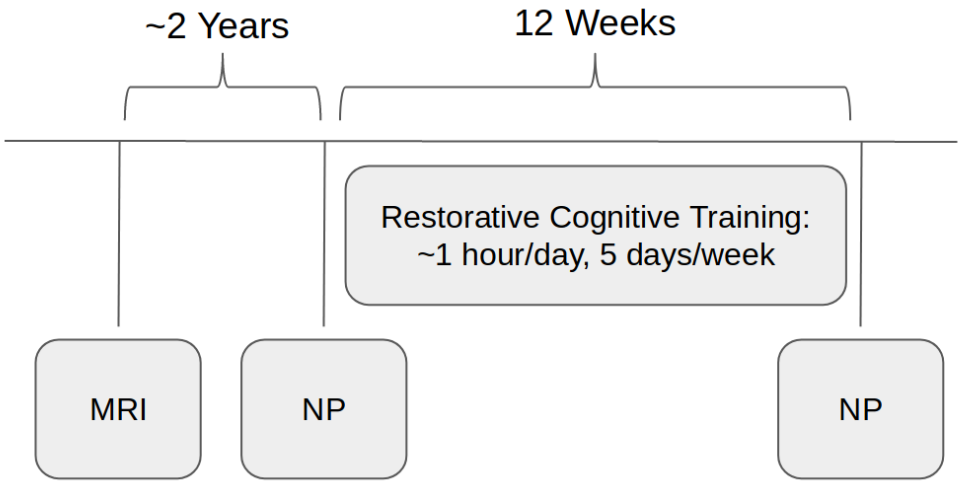

Figure 19. Study timeline. MS patients underwent a 3T MRI at baseline. Then neuropsychological assessment was conducted before and after the restorative cognitive training period.

Participants underwent a baseline neuropsychological assessment and were trained in the use of an online restorative CR software (BrainHQ, Posit Science Corporation). Since cognitive processing speed is the most common cognitive deficit in MS patients (Benedict et al., 2017a), were selected exercises within this program which were expected to improve cognitive processing speed. The initial assessment was completed an average of 2.2 years (median=2.2, interquartile range IQR=1.6-2.8) following the baseline MRI scans available from the CEG study. Notably, study participants only experienced an average of 0.18 (median=0, IQR=0-0, range=0-2)

relapses between the time of MRI and the initial neuropsychological evaluation.

Participants were asked to complete 12 weeks of training on their own home computers. For each week, participants were asked to complete 1 training session/day (~45-60 minutes), for 5 days each week. The rehabilitation software increases difficulty as users improve by adapting parameters such as speed of processing and distractor stimuli. Study participants were only contacted once each week by research assistants to be provided with technical support and mild encouragement. Research assistants had expertise in using the BrainHQ software. Under the supervision of neuropsychologists, they completed phone calls in accordance with the following protocol:

- participants were asked to verify the accuracy of usage data collected by the software;
- participants were provided with one compliment specific to their training performance and were reminded of the study protocol and goals;
- participants were provided with technical support and questions & answers as needed.

After completion of the training period, participants underwent a post-treatment neuropsychological assessment.

Pre- to post-intervention change, exposure, and compliance

Pre- to post-intervention change: The post-rehabilitation improvement on the primary outcome measure, cognitive processing speed (SDMT).

Treatment exposure: The number of exercises completed across all 12 weeks of the study period. Exposure was measured to determine whether pre- to post-intervention change was positively associated with total exposure to treatment; that is whether there was a dose-effect. It is expected

that an increased exposure would correlate with greater pre- to post-intervention change.

Treatment compliance: The percent exercises completed each week out of those scheduled was measured. Then we considered the average weekly compliance across all 12 weeks as the final value.

Neuropsychological assessment and neurologic examination

Neuropsychological testing was conducted under the supervision of a board-certified neuropsychologist according to consensus recommendation, using the Brief International Cognitive Assessment for MS (BICAMS) (Langdon et al., 2012). This included assessment of cognitive processing speed using the Symbol Digit Modalities Test (SDMT) (Smith, 2013), visual/spatial memory using the Brief Visuospatial Memory Test-Revised (BVMTR) (Benedict, 1997), and verbal memory using the California Verbal Learning Test-II (CVLT-II) (Delis, 2000). Performance on the SDMT was the primary outcome measure for this study because of its previously stated as a sensitive and reliable measure of cognitive processing speed, and gold-standard cognitive assessment for MS clinical trials (Strober et al., 2018), in addition to the relevance to the expected primary impact of the rehabilitation software. Total learnings across trials during the BVMTR and CVLT-II were utilized for final analyses.

Executive functions were also assessed using three subtests of the Delis-Kaplan Executive Functions System (DKEFS) (Delis, Kaplan & Kramer, 2001): Tower, Sorting, and Verbal Fluency. Scores from these tests used in the final analyses were achievement score from the Tower test, description score from the Sorting test, and phonemic total correct from the Verbal Fluency test.

The NEO five-factor inventory (NEO-FFI) was applied to measure informant-report trait Conscientiousness (McCrae & Costa, 1987). The sub-scales of Conscientiousness for trait Orderliness, Goal-striving, and

Dependability were considered. Self-report questionnaires were used to assess depression (Beck Depression Inventory Fast Screen, BDI-FS; Benedict et al., 2003), fatigue (Fatigue Severity Scale, FSS; Krupp et al., 1989), and subjective cognitive impairment (Multiple Sclerosis Neuropsychological Questionnaire, MSNQ; Benedict et al., 2004).

Participants were also asked about their neurologic disease course and then these responses were confirmed by trained MS neurologists via inspection of electronic medical records. For the purposes of final analyses, they were categorized as having either relapsing-remitting MS (RRMS) or progressive MS (PMS). Employment status was assessed using the Occupational Functioning Questionnaire (Strober et al., 2012). Clinical disability was assessed using the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983).

At follow-up, the same neuropsychological tests and questionnaires were applied. Alternate testing forms were used when available. Follow-up assessment also included measures of physical disability: Timed 25-Foot Walk (T25FW; Goldman et al., 2013), 9-Hole Peg Test (9HPT; Mathiowetz et al., 1985), and Low Contrast Letter Acuity (LCLA; Regan, Silver & Murray, 1977).

MRI

MRI acquisition

Brain MRI were collected at baseline using a 3T GE scanner. Imaging included a 3D T1-weighted inversion recovery fast spoiled gradient echo, repetition time (TR) 5.9ms, echo time (TE) 2.8 ms, inversion time (TI) 900 ms, flip angle (FA) 10°, field-of-view (FOV) 25.6 ×19.2 cm^2 (256 × 256 matrix with phase FOV 0.75), 180 slices of 1 mm for a final voxel size of 1 × 1 × 1 mm^3.

2D T2-FLAIR scans were also collected with TR 8500 ms, TE 120 ms, TI 2100 ms, FA 90º, echo train length 24, FOV 25.6x19.2cm^2 (256x256

matrix with phase FOV 0.75), 48 slices of 3mm without gap for a final voxel size of $1 \times 1 \times 1$ mm^3.

For resting-state fMRI, 240 volumes of gradient-echo echo-planar images were acquired (TR 2500ms, TE 35ms, FA 90°), 3.75x3.75x4.00-mm^3 voxels, and 4mm slice thickness. During the scan, subjects were instructed to lie awake with their eyes closed and to think of nothing in particular.

MRI processing: conventional volumetry and image parcellation

N4 bias field correction and lesion-filling to T1w images were applied to minimize the impact of inhomogeneity and hypointensity on subsequent processing. FSL tools was used for brain extraction of these images and then the SIENAX cross-sectional software (Smith et al., 2002) (version 2.6) to obtain global volumetric measures, including lateral ventricular volume (LVV), gray matter volume (GMV), white matter volume (WMV), and neocortical gray matter volume (NCV).

Then, the FMRIB Integrated Registration and Segmentation Tool (FIRST; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FIRST) (Smith et al., 2004) were applied to measure deep gray matter volume (dGMV). All volumetric data were normalized for head size.

MRI processing: white matter tract disruption

White matter (WM) lesion masks and T2 lesion volume (T2LV) data were generated using a previously described semi-automated edge detection technique (Zivadinov et al., 2012). Lesion masks were non-linearly registered to the Montreal Neurological Institute (MNI) space using ANTs (Avants et al., 2011). To measure WM tract disruption between pairs of GM regions, the lesion masks were processed using the Network Modification tool (NeMo; Kuceyeski et al., 2013). The use of this tool has been previously described (Fuchs et al., 2018). In short, this tool probabilistically infers alterations to the structural connectivity between pairs of GM regions by

comparing the lesion masks with a database of 73 healthy control tractograms. WM tract disruption between each pair of GM regions is calculated as the number of WM tract streamlines connecting those regions which pass through a lesion and are therefore disrupted. This was done for all pairs of structurally connected GM region-pairs in the 86-region FreeSurfer brain atlas (Dale, Fischl & Sereno, 1999). Figure 20 shows procedure details.

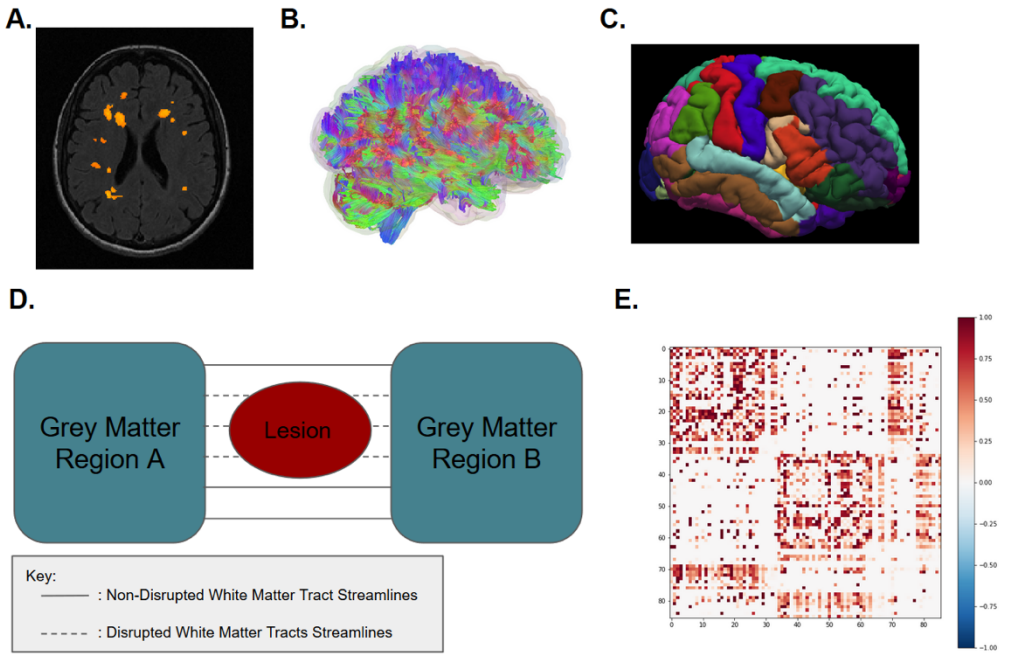

Figure 20. White matter (WM) tract disruption in people with multiple sclerosis. WM tract disruption was measured using the network modification (NeMo) tool. First, lesion masks (A) were registered to Montreal Neurological Institute standard space. Then, a database of healthy control tractograms (B) was applied to probabilistically determine the impact of lesions on white matter tracts which connect gray matter regions in an 86-region FreeSurfer atlas (C). This was done for all pairs of gray matter regions (D), resulting in an 86x86 connectivity/dysconnectivity matrix (E) that illustrates the proportion of WM tracts streamlines, normally connecting each pair of regions, that are disrupted by lesions (range -1 to 1).

MRI processing, preservation/deviation of brain functional connectivity

fMRI data were pre-processed using FSL tools and included the following steps: (a) removal of volumes 1 and 2, (b) slice time correction,

(c) motion correction, (d) intensity normalization, (e) high pass temporal filtering (2000s), (f) brain extraction, (g) fieldmap unwarping, and (h) 4mm spatial smoothing. These data were then non-linearly registered to their corresponding T1w structural images using ANTs (Avants et al., 2011). FSL MELODIC was then applied to perform independent component analysis, classification, and removal of noise components according to published guidelines (Griffanti et al., 2011). Motion parameters, mean cerebrospinal fluid, and mean WM signal were then regressed from the data.

Next, the aim was to produce the summary measure of deviated functional connectivity, as consistent with previous reports (Fuchs et al., 2019a). First, 86x86 functional connectivity matrices were generated using the inverse covariance matrix function of Nilearn (akin to partial correlations, controls for global signal) (Abraham et al., 2014). These matrices were then compared with the average and the standard deviation of a cohort of 29 healthy controls of similar age and sex-distribution (non significantly different) to derive z-score matrices. The absolute value of these was taken, such that lower values reflect functional connectivity which is more similar to HCs and higher values reflect greater deviation.

Statistical Analysis

Sample characteristics, exposure, compliance, and pre- to post-intervention change

All scores from the BICAMS tests (SDMT, BVMTR, CVLT-II) were converted to age- and sex-corrected z-scores according to norms reported in a previous publication (Parmenter et al., 2010).

Similarly, performance on executive function (DKEFS) tasks were converted to age-corrected z-scores, according to abovementioned published manual. These scores were averaged to calculate an executive function composite score.

Z-scores were also calculated for the FSS, BDI-FS, trait Conscientiousness domains, MSNQ, 9HPT, and T25FW according to normative data available in abovementioned publications.

To determine whether study participants exhibited significant cognitive improvement following treatment, paired t-tests were applied to compare cognitive processing speed (SDMT) at the pre- and post-rehabilitation time-points. This was repeated for the secondary outcome measures, BVMTR and CVLT-II. Then, a partial correlation analysis was applied to assess the relationship between treatment exposure and pre- to post-intervention change (improvement on SDMT), controlling for age and disease course. As well, subjects were divided into either cognitive-impaired or cognitively-intact depending on whether their baseline SDMT performance was below or above a z-score of -1. Paired t-test were repeated on each group separately to determine whether post-rehabilitation improvement on SDMT varied between them.

Baseline demographic, clinical and neuropsychological factors associated with response to treatment

Next, attempts have been made to address the primary aim of exploring baseline factors that help explain individual patient variability in response to the rehabilitation. These factors were intentionally explored separately from quantitative MRI factors because clinical/neuropsychological factors are often more accessible in clinical environments and so their predictive capacity were thought to be more clinically translatable. To identify baseline demographic, clinical, and neuropsychological factors associated with pre- to post-intervention change (improvement on SDMT), a two-stage approach was followed. First, two-tailed Spearman correlation analysis, Pearson's correlation analysis, or independent sample t-tests were applied when appropriate, with pre- to post-intervention change as the dependent variable. Next, baseline predictors

from the first stage of analysis with a p-value ≤ 0.15 were included in a forward stepwise regression predicting pre- to post-intervention change. The following baseline demographic and clinical factors were explored: age, sex, education, disease duration, disease course, and physical disability (EDSS). The following neuropsychological factors were examined: cognitive processing speed, personality trait Conscientiousness and its subdomains (Orderliness, Goal-Striving, Dependability), subjective cognitive impairment, executive function, depression, and fatigue. For descriptive purposes, a correlation matrix between these predictor variables was also generated and baseline cognitive processing speed for RRMS was compared to PMS participants using independent-samples t-tests.

Baseline brain MRI factors associated with response to treatment

The above analysis plan was then repeated for the following baseline brain MRI factors (T2LV, LVV, GMV, WMV, NCV, and dGMV).

Exploratory analyses of predictive factors on a sub-set of study participants with above threshold compliance and response to treatment

In a sub-sample of study participants (n=17) who improved on SDMT above a cut-off of 4 raw-score points (Strober et al., 2018) and were minimally compliant (>25%), the stepwise regression analyses were repeated, considering all clinical, neuropsychological, and MRI factors. This was repeated again, excluding subjects with greater than 3 years between time of MRI and baseline neuropsychological testing (n=14).

Baseline functional connectivity and structural network disruption

In a subgroup of participants that underwent resting state fMRI at baseline (n=25), attempts have been made to address the primary aim of exploring whether baseline functional connectivity moderates the relationship between structural network disruption and post-rehabilitation

improvement on SDMT. This was accomplished in two stages: (1) identification of a structural network of connected GM regions whose WM tract disruption at baseline predicts patient post-rehabilitation improvement on SDMT and (2) determining whether functional connectivity deviations within the network moderate this relationship. Furthermore, (3) an exploratory analysis, conventional network functional connectivity was also conducted.

(1) Identification of WM tract disruption network

The aim was to identify local patterns of baseline WM tract disruption which predict response to CR treatment. To do this the network-based statistics (NBS) tool (Zalesky, Fornito & Bullmore, 2010) was applied. The NBS tool, similar to conventional cluster statistics, is a method for controlling family-wise error while retaining power by accounting for network structure. First, in relation to change in SDMT performance, a test statistic is computed for each pair of connected GM regions. Then a network of structurally connected suprathreshold pairs is identified and a *p*-value is assigned to the entire identified component by comparing its size with the null distribution of maximal component size (5000 permutations). The analysis was performed including the covariates of age, sex, and GMV. GMV was included to control for overall disease burden in the identification of a localized network. Results were considered significant at an α-level of 0.05. Once the network was identified, a one-sample t-test was applied to determine if the proportion of disruption within the identified network was significantly greater than 0.

(2) Functional connectivity moderation of the structure-cognitive relationship

First, were calculated the mean proportion of WM tract disruption and the mean deviation of functional connectivity for GM pairs included in

the associated network (identified as described in the section above). Deviation of functional connectivity was reported as 'high' for above median values and 'low' for below median values. To determine whether the impact of structural network disruption is moderated by deviation in functional connectivity, multiple linear regression analysis was applied. For the regression, network WM tract disruption data was cubed-root transformed to maintain distribution normality. The regression model, predicting post-rehabilitation change in SDMT performance, included the covariates: age, sex, education (years), WM tract disruption of the associated network, deviation of functional in the network, and the interaction term between network WM tract disruption and functional network deviation. All β coefficients reported are standardized and R^2 values are adjusted for the number of covariates included in the model.

(3) Exploratory analysis, conventional network functional connectivity

To better understand the role of specific well-known functional brain networks, an additional exploratory analysis was completed in which was assessed the relationship between response to treatment and baseline functional connectivity with established resting state networks (Smith et al., 2009). This analysis was completed using FSL Dual Regression (Nickerson et al., 2017). First, a temporal-concatenated group independent component analysis was conducted to identify 20 spatial-temporal maps. These components were labeled according to previously established resting-state networks (Smith et al., 2009), and they were used to extract per-subject sets of time series using general linear modelling.

FSL randomize was then used with threshold-free cluster enhancement, 5000 permutations, to identify patterns of functional connectivity in association with the resting-state networks which are significantly associated with treatment response. Age and sex were included

as covariates in the model and corrected *p*-values using false discovery rate (FDR). Results were considered significant at α-level of 0.05.

Results

Sample characteristics, exposure, compliance, and pre- to post-intervention change

MS patients (70.6% female, 47.1% employed) participating in the study had a mean age of 56.1±11.8, disease duration of 21.6±9.8 years, and baseline SDMT score of 49.5±14.7. Baseline cognitive processing speed was significantly higher for participants with RRMS (mean=52.7, SD=13.6) relative to those with PMS (mean=42.6, SD=15.1; cohen's d=0.71, t=2.35, *p*=0.023). Mean speed on T25FW and 9HPT were 7.2±5.6 and 30.7±15.6 seconds respectively. Additional details are provided in Table 15.

	Patient Factor	Raw Scores	Z-scores
Demographics and Clinical Characteristics	Age	56.1±11.8	/
	Female/Male; % Female	36/15; 70.6	/
	Education (years)	15.6±2.3	/
	Right-handed/Left-handed; % Right-handed	45/6; 88.2	
	Disease Duration	21.6±9.8	/
	Disease Course		
	Relapse Remitting MS; %	35; 68.6	/
	Primary Progressive MS; %	4; 7.8	/
	Secondary Progressive MS; %	12; 23.5	/
	Physical disability, EDSS	4.0 (2.0-6.0)	/
Neuropsychological Metrics	Cognitive functioning (BICAMS)		
	Cognitive Processing Speed, SDMT	49.6±14.7	-0.46 ± 1.13
	Visual/Spatial Memory, BVMTR	23.3±8.0	0.22 ± 1.29
	Verbal Memory, CVLT-II	53.5±12.6	-0.17 ± 1.28
	Executive Functions (DKEFS)		
	Tower Test, Total Achievement	16.3±3.9	0.12 ± 0.87
	Card Sorting, Free Description	34.9±10.9	0.39 ± 1.02
	Letter Verbal Fluency, Total Correct	35.9±11.2	-0.03 ± 1.14
	Conscientiousness	47.1±12.5	-0.67 ± 0.93
	Orderliness	2.1±4.2	/
	Goal-Striving	3±2.5	/
	Dependability	4.2±3.4	/
	Subjective Cognitive Impairment, MSNQ-P	19.5±11.8	-0.66 ± 0.99
	Depression, BDI-FS	1.7±2.4	-0.56 ± 0.71
	Fatigue, FSS	4.1±1.7	-0.51 ± 1.28
Brain MRI	T2LV	15.0±18.5	/
	LVV	50.8±22.6	/
	GMV	737.4±50.6	/
	WMV	714.7±38.5	/
	NCV	600.4±38.5	/
	dGMV	54.0±5.7	/

Table 15. Demographic/clinical, neurological/neuropsychological, and brain MRI characteristics of study participants (N=51). Mean±SD are provided for continuous variables; median (IQR) is provided for EDSS.
SD = standard deviation; MS = Multiple Sclerosis; EDSS = Expanded Disability Status Scale; IQR = interquartile range; BICAMS = Brief International Cognitive Assessment for Multiple Sclerosis; SDMT = Symbol Digit Modalities Test; BVMTR = Brief Visuospatial Memory Test Revised; CVLT-II = California Verbal Learning Test II; DKEFS = Delis-Kaplan Executive

Function System; MSNQ-P = Patient-Report Multiple Sclerosis Neuropsychological Screening Questionnaire; BDI-FS = Beck Depression Inventory – Fast Screen; FSS = Fatigue Severity Scale; MRI = magnetic resonance imaging; T2LV = T2 lesion volume; LVV, lateral ventricular volume; GMV = gray matter volume; WMV = white matter volume; NCV = neocortical gray matter volume; dGMV = deep gray matter volume.

Study participants were exposed to a mean±SD of 561.4±398.3 exercises (median=566, IQR=225-870). Mean±SD weekly compliance for the study participants was 62.3±44.5. Compliance was significantly directly correlated with age (r=0.53, $p<0.001$) and inversely with the number of children raised at home (r=-0.45, $p=0.001$).

Significant improvement was observed on SDMT (cohen's d=0.55, t=3.91, $p<0.001$) from a mean±SD of 49.6±14.7 to 52.6±15.6 (Figure 21).

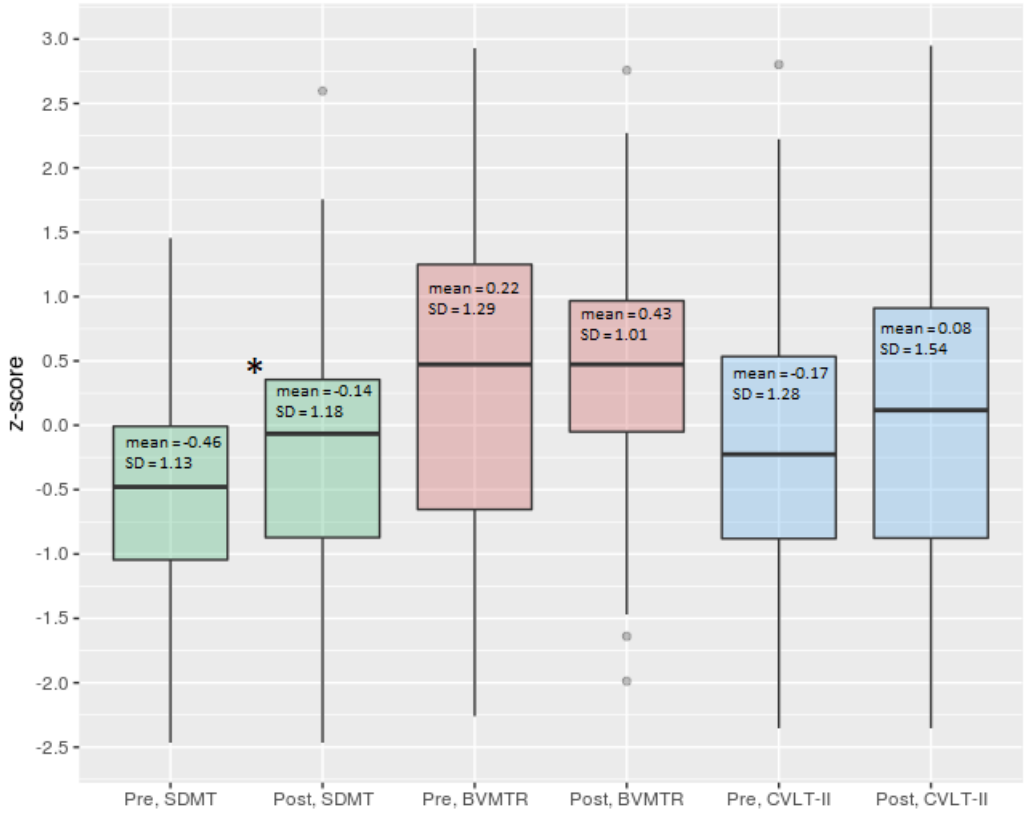

Figure 21. Performance on assessments of cognition pre- and post-rehabilitation. Study participants exhibited a statistically significant improvement on cognitive processing speed (SDMT).
* : $p < 0.001$
SDMT = Symbol Digit Modalities Test; BVMTR = Brief Visuospatial Memory Test Revised; CVLT-II = California Verbal Learning Test – II

Participants with RRMS improved by a mean±SD of 4.4±6.5 points on SDMT (p<0.001), exceeding the expected SDMT practice effect in MS (~0.5 points per repetition) by over 800% (Benedict et al., 2008a). In comparison, participants with PMS only improved by a mean±SD of 0.3±4.7 points (p=0.835) (Figure 22). Out of all the study participants (n=51), 22 (43.1%) exhibited a post-intervention improvement on SDMT above a threshold of 4 raw score points (Strober et al., 2018). Treatment response correlated positively with exposure (r=0.38, p=0.007), accounting for the effect of age and disease course. Nonetheless, response was highly heterogeneous (mean change±SD=3.1±6.2, range -6 to 21; Figure 22). A trending but non-significant improvement was also observed on BVMTR (cohen's d=0.22, t=1.58, p=0.12) and CVLT2 (cohen's d=0.26, t=1.86, p=0.069). Post-intervention improvement on SDMT was statistically significant for subjects who were cognitively impaired at baseline (n=14, mean change=3.4, cohen's d=0.74, t=2.75, p=0.016) and for those who were not (n=37, mean change = 3.0, cohen's d=0.48, t=2.96, p=0.005).

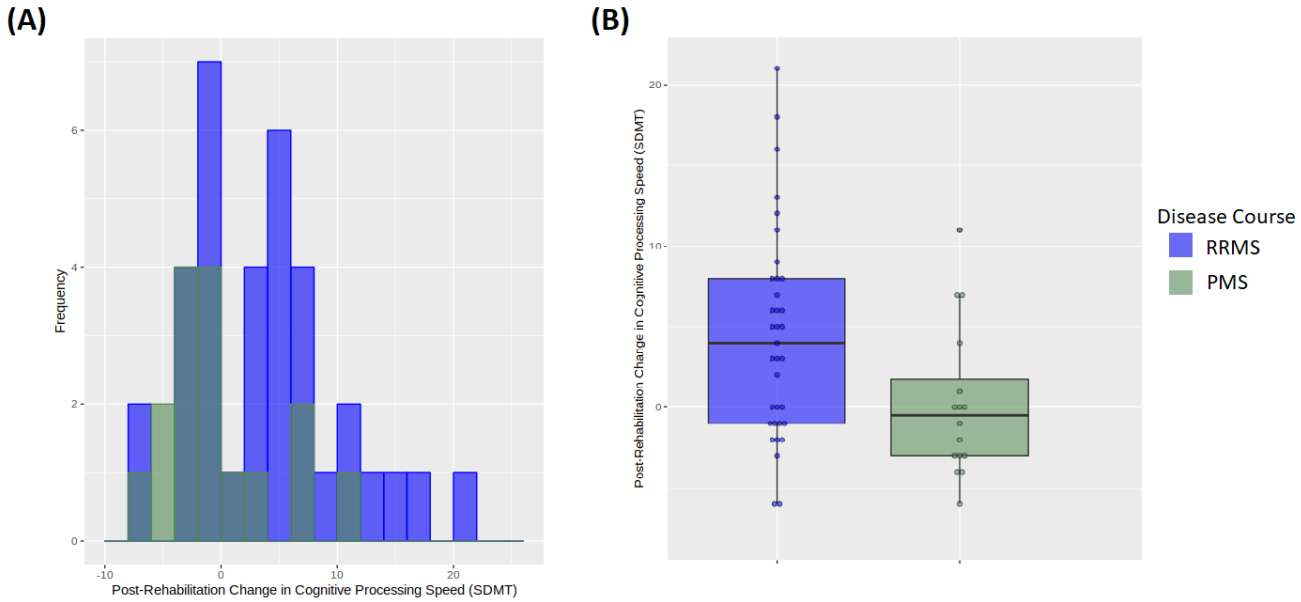

Figure 22. Disease course and treatment response heterogeneity. Greater post-rehabilitation improvement on cognitive processing speed (SDMT) was observed for people with RRMS relative to those with PMS. These differences and response heterogeneity

between the PMS and RRMS participants are illustrated here in a histogram (A) and scatter-box plot (B).
*: $p < 0.05$
SDMT = Symbol Digit Modalities Test; RRMS = Relapsing-Remitting Multiple Sclerosis; PMS = Progressive Multiple Sclerosis

Baseline demographic, clinical, and neuropsychological factors associated with response to treatment

Figure 23 shows a correlation matrix between each potential predictor considered, generated for descriptive purposes.

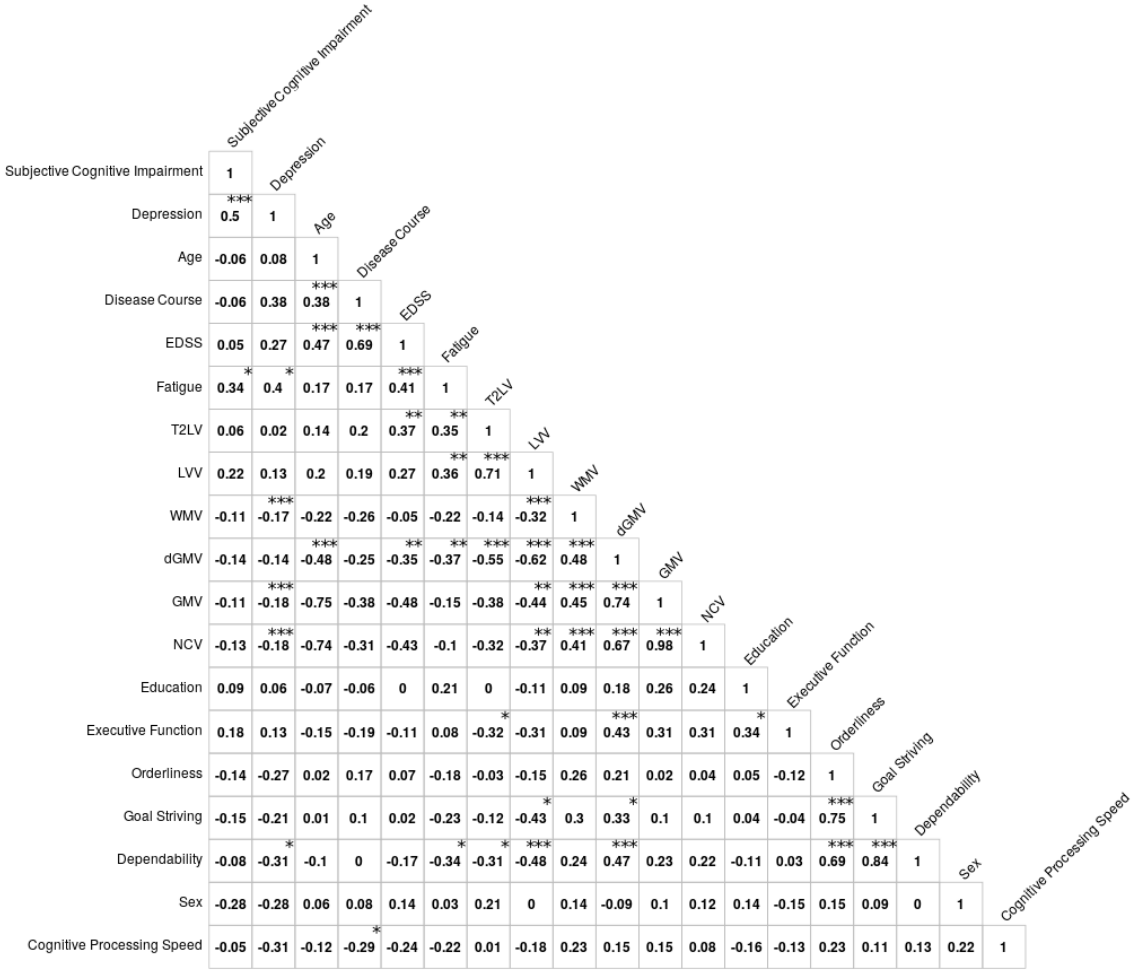

Figure 23. Association between baseline patient factors.
EDSS = Expanded Disability Status Scale; T2LV = T2 lesion volume; LVV = lateral ventricular volume; GMV = gray matter volume; WMV = white matter volume; NCV = neocortical gray matter volume; dGMV = deep gray matter volume

The associations between each baseline factor and response to treatment (improvement on SDMT) are provided in Table 16. A significant effect was observed for RRMS relative to PMS (t=2.29, p=0.026) and physical disability (EDSS, r_s=-0.31, p=0.027). The final model was a statistically significant predictor of treatment response (R^2=0.130, p=0.016) and included the following variables: disease course (RRMS vs PMS; β=-0.343, p=0.017) and trait Conscientiousness-Orderliness (β=0.292, p=0.040). In Figure 22 is provided a visualization of these group differences.

	Candidate Baseline Predictor	**Effect**	***p***
Demographics and Clinical Characteristics	Age	r=-0.151	0.291
	Sex (female vs male)	t=-1.467	0.149
	Education (Years)	r=-0.142	0.321
	Disease Duration (Years)	r=-0.101	0.480
	Disease Course (RRMS vs PMS)	t=2.294	**0.026**
	Physical Disability (EDSS)	r_s=-0.312	**0.027**
Neuropsychological Metrics	Baseline Cognitive Processing Speed (SDMT)	r=-0.058	0.684
	Executive Functions (DKEFS)	r=-0.213	0.134
	Conscientiousness-Orderliness	r=0.232	0.113
	Conscientiousness-Goal-Striving	r=0.110	0.445
	Conscientiousness-Dependability	r=0.126	0.395
	Subjective Cognitive Impairment (MSNQ-P)	r=0.120	0.402
Brain MRI	T2LV	r=-0.208	0.143
	LVV	r=-0.214	0.140
	GMV	r=0.311	**0.030**
	WMV	r=0.260	0.071
	NCV	r=0.245	0.089
	dGMV	r=0.283	**0.049**

Table 16. Association between baseline patient factors and response to restorative cognitive rehabilitation (post-rehabilitation change in cognitive processing speed, SDMT).
SDMT = Symbol Digit Modalities Test; RRMS = relapsing-remitting multiple sclerosis; PMS = progressive multiple sclerosis; EDSS = Expanded Disability Status Scale; DKEFS = Delis-Kaplan Executive Function System; MSNQ-P = Patient-Report Multiple Sclerosis Neuropsychological Screening Questionnaire; MRI = magnetic resonance imaging; T2LV = T2 lesion volume; LVV = lateral ventricular volume; GMV = gray matter volume; WMV = white matter volume; NCV = neocortical gray matter volume; dGMV = deep gray matter volume

Baseline conventional brain MRI factors associated with response to treatment

The associations between baseline brain MRI factors and response to treatment (improvement on SDMT) are provided in Table 16. A significant positive effect was observed for GMV (r=0.31, p=0.030) and dGMV (r=0.28, p=0.049). The final model (R^2=0.077, p=0.030) included GMV as a statistically significant predictor of treatment response (β=0.311, p=0.030; Figure 24A). Post-hoc analyses revealed that participants with PMS had significantly lower GMV relative to those with RRMS (t=2.74, p=0.009; Figure 24B) and the correlation between Conscientiousness-Orderliness and response to treatment was stronger for individuals with lower baseline GMV (below median, r=0.44, p=0.026) relative to those with higher baseline GMV (r=0.10, p=0.66).

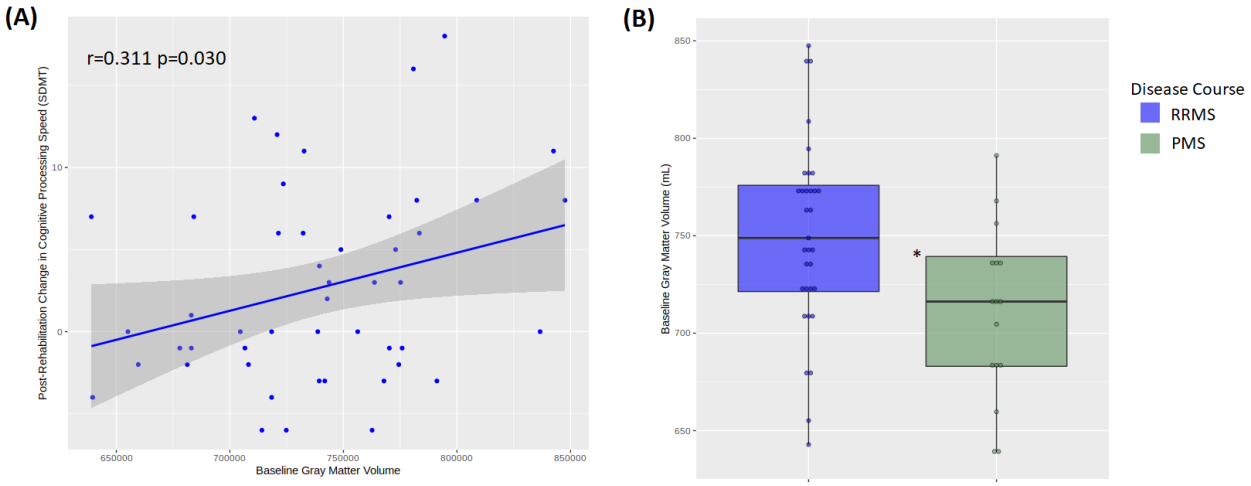

Figure 24. Gray matter volume in study participants. Greater post-rehabilitation improvement of cognitive processing speed (SDMT) was associated with higher baseline GMV (A). Subjects with PMS had reduced GMV at baseline relative to those with RRMS (B).
*: $p < 0.05$
SDMT = Symbol Digit Modalities Test; PMS = Progressive Multiple Sclerosis; RRMS = Relapsing-Remitting Multiple Sclerosis

Exploratory analyses of predictive factors on a sub-set of study participants (n=17) with above threshold compliance and response to treatment

In the sub-sample of study participants with above-threshold compliance and response to treatment (n=17; 76.4% RRMS), the final model was a statistically significant predictor of post-rehabilitation SDMT improvement (R^2=0.742, p<0.001). Within this final model, a significant positive effect was observed for dGMV (β=0.73, p<0.001; model 1 ΔR^2=0.376), trait Orderliness (β=0.60, p=0.001; model 2 ΔR^2=0.281), and executive functioning (β=0.31, p=0.041; model 3 ΔR^2=0.085). In the analysis repeated for subjects with the additional criteria of having MRI within 3 years of baseline neuropsychological assessment (n=14), the final model was a statistically significant predictor of post-rehabilitation improvement on SDMT (R^2=0.79, p<0.001). This model included dGMV (β=0.86, p<0.001) and trait Orderliness (β=0.61, p=0.001) as statistically significant predictors of post-rehabilitation improvement of SDMT performance.

Baseline advanced brain MRI factors associated with response to treatment

(1) Identification of WM tract disruption network

Network based analysis revealed a single network of connected GM regions whose structural disruption at baseline predicted patient response to rehabilitation, controlling for age, sex, and GMV. This network included four region-pairs in the left hemisphere and was largely centered on default mode network-related regions, the precuneus and cingulate (p=0.048). Details are provided in Table 17 and Figure 25. Median(IQR) pairwise disruption of WM tract streamlines in the network was 65.7%(22.9-95.6%) and mean±SD disruption was 59.1±36.9% (t=11.428, p<0.001). Tract disruption within this network was higher in participants with progressive MS relative to relapsing-remitting, approaching statistical significance (t=1.98, p=0.071).

Region-pair	% tract disruption	Deviation of FC	t value	*p*
L. Precuneus-L. Transverse Temporal Gyrus	60.9±39.5	0.832±0.627	3.43	**0.006**
L. Precuneus-L. Globus Pallidus	56.8±38.2	0.720±0.568	3.15	**0.006**
L. Cingulate Isthmus-L. Superior Temporal Gyrus	58.1±39.7	0.990±0.618	3.14	**0.009**
L. Cingulate Isthmus-L. Globus Pallidus	53.9±38.2	0.637±0.492	2.70	**0.014**

Table 17. Network white matter tract disruption. Accounting for age, sex, and gray matter volume, lower baseline white matter tract disruption within the described network of connected gray matter regions predicted greater post-rehabilitation improvement on Symbol Digit Modalities Test. The strength of association between tract disruption in each region-pair and patient response to the restorative rehabilitation is listed.
Mean±SD are provided for % tract disruption and deviation of FC.
L = left; R = right; FC = Functional Connectivity

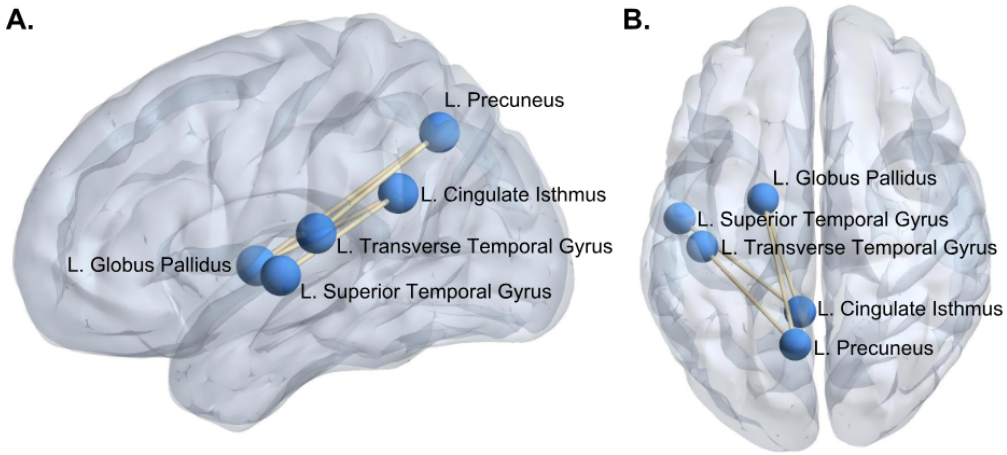

Figure 25. Accounting for age, sex, and gray matter volume, lower baseline white matter tract disruption within the pictured network of connected gray matter regions predicted greater post-rehabilitation Symbol Digit Modalities Test improvement. This network here illustrated (A. sagittal; B. horizontal) is centered on the left precuneus and cingulate.

(2) Functional connectivity moderation of structure-cognitive relationship

In the regression model predicting post-rehabilitation improvement of SDMT, (R^2=0.385, p=0.017), a statistically significant interaction effect between WM tract disruption and deviation in functional connectivity in the associated network was identified (β=-0.415, [95% CI: -0.764, -0.065], p=0.022). In Table 18 and Figure 26 are provided more details. Based on these results, subjects were divided into four groups according to median network structural disruption and functional deviation: (1) low structural disruption with high functional deviation (greatest treatment response), (2) low structural disruption with low functional deviation, (3) high structural disruption with low functional deviation, and (4) high structural disruption with high functional deviation (lowest treatment response). One-way ANOVA revealed these groups significantly differed in regard to baseline T2LV (F=17.92, p<0.001), such that group 1 had lowest baseline T2LV (group 1 mean=1.8mL; group 2 mean=3.1mL; group 3 mean=33.4mL; group 4 mean=35.7mL). The mean deviation in network functional connectivity did not significantly vary between groups 1 and 4 (t=-1.22, p=0.33), and the directions of the deviations were consistent for all network region-pairs across the groups.

	mean±SD	**β**	**β, 95% CI**	***p***
Age (years)	54.7±12.5	-0.299	-0.664, 0.067	0.103
Gender (F vs. M)	/	0.344	-0.022, 0.711	0.063
Education (year)	16.9±1.9	-0.325	-0.672, 0.021	0.064
Network WM Tract Disruption (%)	56.6±39.1	-0.210	-0.571, 0.151	0.237
Network Functional Deviation (|z|)	0.23±0.07	-0.081	-0.462, 0.301	0.661
Interaction: Tract Disruption-by-Functional Deviation	/	-0.415	-0.764, -0.065	**0.022**

Table 18. Functional Connectivity Moderation of Structure-Cognitive Relationship. In a model predicting patient response to restorative cognitive rehabilitation (post-rehabilitation improvement on Symbol Digit Modalities Test, SDMT), deviation in functional connectivity

moderated the relationship between baseline structural network disruption and post-rehabilitation improvement on SDMT.
CI = confidence interval.

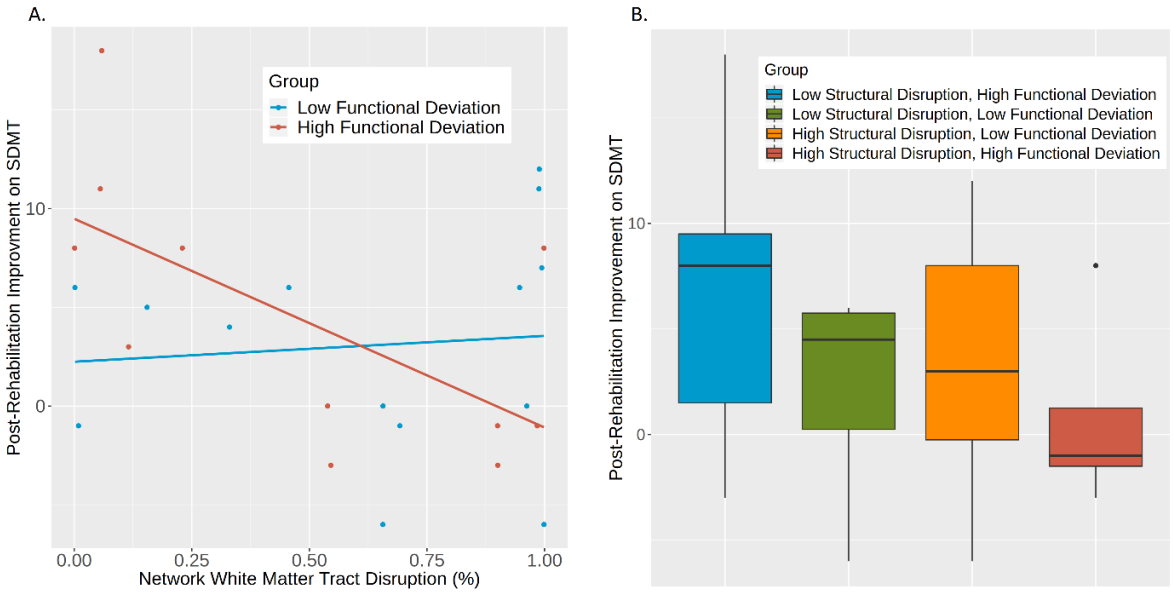

Figure 26 The relationship between baseline structural network disruption and post-rehabilitation improvement on SDMT was moderated by patient profiles of functional connectivity (A)(β=-0.415, p=0.022). Accounting for network functional deviation, four groups of patients were identified with varying degrees of post-rehabilitation improvement on the SDMT (B). Study participants with low structural disruption and high functional deviation at baseline responded best to the treatment. In contrast, those with higher degrees of structural disruption responded less well to treatment, especially given higher levels of functional deviation. In Figure 26B, the rectangles of the boxplots extend from the first to the third quartiles, the mid-lines represent the median, and the whiskers extend to 1.5 times the interquartile range.
*: $p < 0.05$

(3) Exploratory analysis, conventional network functional connectivity

Independent component analysis isolated 20 spatiotemporal maps, consistent with previously established resting-state networks (Smith et al., 2009). These components conformed to the boundaries of the visual (5 components), frontoparietal (2 components), default-mode (4 components), executive control (2 components), cerebellar (1 component), auditory (1 component), and somatosensory networks (5 components).

Baseline functional connectivity of several clusters of gray matter within the DMN significantly correlated with greater post-rehabilitation improvement on SDMT (mean t-stat=2.37, mean p=0.04, 137 voxels). The largest of these clusters included the superior frontal gyrus (MNI 4, 36, 52), frontal pole (MNI 30, 50, 22), and precentral gyrus (MNI 30, -22, 62). Figure 27 for details.

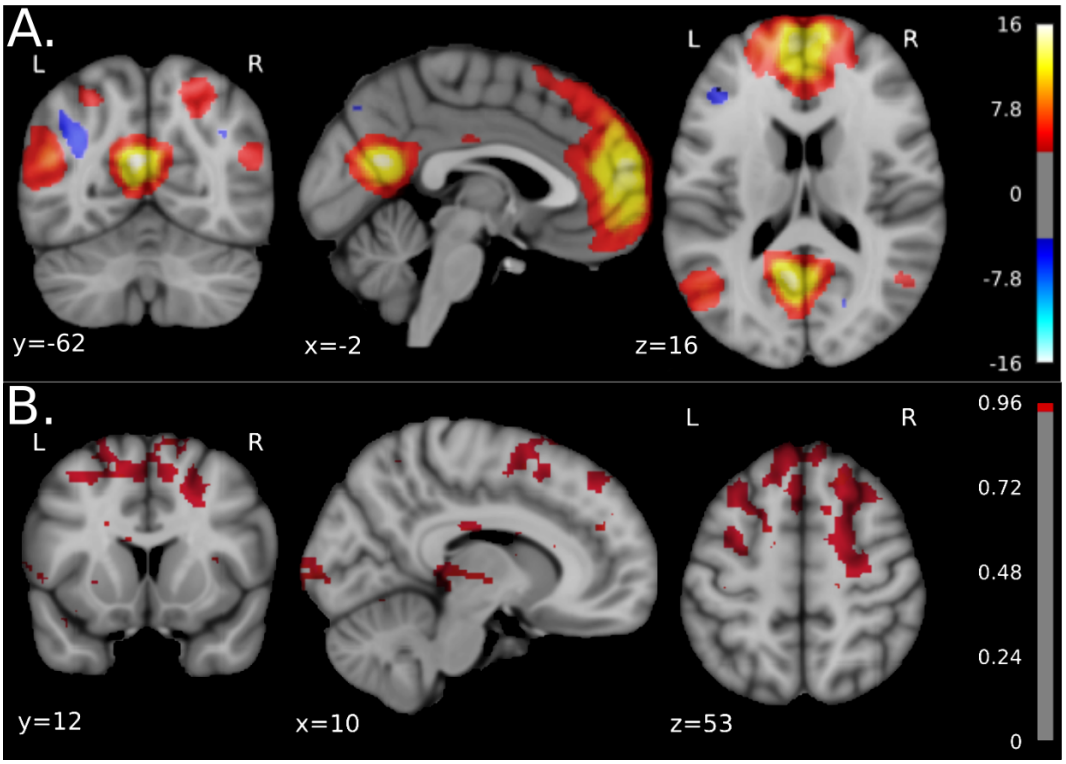

Figure 27. Functional connectivity of several clusters within the default mode network (A) was significantly associated with increased post-rehabilitation improvement on the Symbol Digit Modalities Test (B), accounting for age and sex. The largest clusters included the superior frontal gyrus (Montreal Neurological Institute, MNI 4, 36, 52), frontal pole (MNI 30, 50, 22), and precentral gyrus (MNI 30, -22, 62).
L = left; R = right

Discussion

MS patients improved cognitive processing speed following a restorative home-based cognitive training. In participants with RRMS, this was 800% over the expected practice effect on the SDMT (Benedict et al., 2008a) and is considered a clinically meaningful change (Strober et al.,

2018). Greater cognitive improvement was associated with baseline diagnosis of RRMS vs PMS, higher trait Orderliness (a subcomponent of Conscientiousness), and increased GMV (Figure 22 & Figure 24).

In the sub-sample of study participants with above threshold pre- to post-treatment improvement on SDMT and compliance (n=17; 76.4% RRMS), cognitive improvement was predicted by greater dGMV, higher trait Orderliness, and higher executive functioning. In this small sub-sample, a much higher degree of variance was explained by the final predictive model relative to the best model including all study participants (74.2% versus 13.0%), though the small sample size limits generalizability.

The positive association between baseline GMV and response to training is consistent with the importance of gray matter pathology in MS (Calabrese et al., 2015a) and cognition (Batista et al., 2012). Substantial gray matter atrophy likely contributes to an inability to functionally compensate for neuropathology and therefore limits the restorative capacity of the brain (Schoonheim, Meijer & Geurts, 2015). This result and the increased pre- to post-intervention change in RRMS (as compared to lower pre- to post-intervention change in PMS) is relevant because, in the past, restorative CR has been largely validated in studies which only consider people with RRMS (Mitolo et al., 2015). This view is reminiscent of the widely accepted notion that after traumatic brain injury, restorative approaches are most applicable during a period of spontaneous recovery of function and after a plateau, compensatory strategies are indicated (Eshel, Bowels & Ray, 2019). In MS, restorative CR may be better suited for RRMS patients whose neural capacity to benefit is spared in contrast to PMS characterized by more cerebral injury. Rather than the restorative technique investigated here, the compensatory approach may be more useful for people with PMS. In fact, a recent randomized controlled trial indicated class I evidence for the pre- to post-intervention change of a compensatory rehabilitation technique in PMS (Chiaravalloti, Moore & DeLuca, 2019). As

these findings are preliminary, future work is needed to further explore the pre- to post-intervention change of restorative versus compensatory approaches in people with varied disease course.

In this sample, RRMS were less cognitively impaired than PMS at baseline. This may have led us to believe that a higher level of cognitive functioning at baseline partially explains the difference in these groups responses to cognitive training. Nonetheless, baseline cognitive processing speed was not a significant predictor of response to the treatment itself and post-rehabilitation improvement on SDMT was not different for the cognitively-impaired subjects relative to those who were not. Based on these observations, it is unlikely that the group differences in response to training are due to differences in cognitive impairment alone. Future research is needed to clarify the relationship between baseline cognitive functioning and responsiveness to restorative cognitive training.

The effect of personality on response to training was also of particular interest. There are several possible explanations for the positive relationship between baseline Orderliness (a sub-component of Conscientiousness) and response to treatment. Conscientiousness is subsumed within the wider neurological reserve construct, as there are synergistic effects between Conscientiousness and traditional measures of cognitive reserve (Roy et al., 2016). High Conscientiousness might also be viewed as a correlate of a healthier brain: conscientiousness is associated with less neocortical atrophy (Benedict et al., 2008b) and spared network disruption in eloquent fronto-parietal networks which might otherwise contribute to cognitive adaptability (Fuchs et al., 2018). Additionally, Orderliness specifically relates to whether a person is methodical, organized, and efficient (Chapman, 2007). It is possible that these traits are contributing to increased treatment response because of increased moment-to-moment intentionality and purposefulness while completing exercises. The relationship between Conscientiousness and response to training is especially interesting in the context of MS, because

Conscientiousness is lower in MS patients and declines longitudinally in this population (Benedict et al., 2008b; Roy et al., 2018). Furthermore, lower Conscientiousness in MS patients is predictive of less favorable outcomes, such as increased unemployment and faster rates of cognitive decline (Strober et al., 2012; Fuchs et al., 2019b). Because Conscientiousness is negatively affected in MS patients and vice-versa, lower Conscientiousness itself predicts other unfavorable outcomes, it is often difficult to interpret the direction of causation. Future work is needed to clarify and to determine whether Conscientiousness-targeted interventions can help improve treatment outcomes.

Considering advanced MRI outcomes, in a subgroup of MS patients were found that lower WM tract disruption in a network of regions centered on the precuneus and posterior cingulate (key components of the default-mode network, DMN) predicted greater cognitive improvement ($p=0.048$) following the restorative CR. This relationship was moderated by profiles of functional connectivity within the network ($\beta=-0.415$, $p=0.022$), revealing four groups of subjects with varying degrees of response to rehabilitation: low structural network disruption with high functional deviation (greatest treatment response), low structural network disruption with low functional deviation, high structural network disruption with low functional deviation, and high structural network disruption with high functional deviation (lowest treatment response).

These results have several implications. For one, localized patterns of baseline structural network disruption between GM regions, including default-mode network regions, predict patient response to restorative CR. The variance explained by this network is above what is otherwise explained by GMV, a metric which predicts patient response to restorative CR better than other conventional MRI measures, as shown in this study. Thus, future studies aimed at determining which MS patients are most likely to benefit from restorative CR may benefit from exploring network-based metrics,

especially for the default-mode network, in addition to more conventional measures. The importance of the default-mode network in relation to cognitive functioning is a well-observed phenomenon in patients with MS (Rocca et al., 2010; Eijlers et al., 2017; Van Geest et al., 2018; Savini et al., 2019). This study builds on these findings by further demonstrating the importance of the default-mode network connectivity, not only to cognitive functioning, but also to individual's abilities to improve cognition with training. These findings may also help explain why people with progressive MS are less likely to respond to restorative CR, as they approached significantly greater disruption within the described network in the present study relative to individuals with relapsing-remitting MS.

These findings also indicate that patterns of functional connectivity moderate the relationship between baseline structural network damage and patient response to treatment. In the context of a previous work, which showed a relationship between preservation of functional connectivity and cognitive reserve in patients with MS (Fuchs et al., 2019a), the current findings may indicate that patterns of functional connectivity relate not only to cognitive reserve potential but also to the ability to improve cognitive functioning. It is interesting to see that our exploratory analysis further corroborated the importance of the default-mode network for determining which people are most likely to benefit from cognitive rehabilitation. These results were notably derived independently of the primary analysis plan and were acquired using a statistically driven approach. This exploratory analysis was undertaken to take a "function-first" analysis approach (rather than focused on networks of structurally disrupted networks) and to explore the relationship between response to rehabilitation and connectivity of known/established functional networks (Smith et al., 2009). In this analysis, it was also observed the importance of functional connectivity of the frontal pole and superior frontal gyrus with the default-mode network, which also further supports the importance of these regions for cognition in healthy

adults and in patients with MS (Duncan & Owen, 2000; Benedict et al., 2002).

Furthermore, the present findings build on this previous work because heretofore, preserved functional connectivity (i.e., most similar to HCs) was considered as an indication of maintained cognitive function despite structural brain insult. However, this study now also highlights that some individuals with lower structural network damage have patterns of functional connectivity that deviate from the HCs norms and also exhibit the greatest response to restorative cognitive training. In addition, individuals with functional connectivity most similar to HCs (low deviation), independently from their degree of structural network disruption, responded moderately well to the treatment. From this, one can deduce that when interpreting patterns of functional connectivity and whether deviations are adaptive or pathological, considering the context of such deviations is vital. Some MS patients with low structural network damage may benefit from adaptive functional reorganization, reflected as higher deviation compared to HCs. On the other hand, later in the disease after greater structural network damage, similar deviations may instead reflect a decompensation of functional connectivity.

Conflicting reports on the meaning of functional connectivity alterations have been reported many times in the literature (Hawellek et al., 2011; Rocca & Filippi, 2017; Schoonheim, 2017; Dobryakova, Rocca & Filippi, 2018; Meijer et al., 2018). Functional reorganization may be an adaptive response to early structural damage, whereas functional decompensation may occur during later stages of the disease with greater accumulation of structural damage (Schoonheim, Geurts & Barkhof, 2010). These results build on this hypothesis, suggesting that deviations in functional connectivity may be adaptive or maladaptive, depending on the degree of structural network disruption characterizing each person with MS.

These results provide preliminary evidence that baseline demographic, clinical, neuropsychological, and brain MRI factors are associated with response to restorative CR in MS patients. However, many other factors, i.e. emotional state (anxiety or stress), sleep quality, and familiarity with technology might contribute as well to moderate effects of cognitive intervention in MS. Given the added strength of the model performed on the sub-sample of supra-threshold study participants, it is possible that variability in treatment compliance and ceiling effects on treatment response also confound final predictions. Response to treatment may have also been limited to improvement on cognitive processing speed because the exercises chosen for the restorative cognitive intervention were selected to improve this domain.

The present study has several limitations. Firstly, patients compliance in this study was slightly lower than for a similarly structured intervention (Charvet et al., 2017), a difference possibly attributable to participants being asked to use their own home computers rather than laptops provided by study administrators. Additionally, compliance was not strictly enforced for participation in the study. It has been chosen to design the protocols in this way to increase the real-world clinical relevance of these findings. Use of home computers and restricted clinician involvement decreased the cost of treatment in terms of time, travel, and money and thus allows accessibility for a wider range of individuals. These results reflect a “real-world” application of Brain HQ, as it is likely to be utilized by MS patients who are steered toward the program and advised to use it on their own.

Though study participants experienced a median of 0 (IQR=0-0) relapses between the time of the MRI and the initial neuropsychological evaluation, future work could be improved by collecting MRI data closer to the time of baseline neuropsychological assessment. Future works should investigate a larger cohort of participants to provide more comparable samples of people with PMS and RRMS, include control treatment, and to

validate predictor variables in training sets and testing sets. It has been chosen not to include a control intervention so that it could adequately power the primary analyses – investigating predictors of response variability to a previously validated program (Charvet et al., 2017). Future studies should also further explore default-mode network disruption and patterns of functional connectivity to elucidate how deviations in functional connectivity affect cognitive functioning and potential for cognitive improvement in the context of varying structural brain insult. Moreover, it will be necessary in the future to compare the pre- to post-intervention change of restorative and compensatory training in RRMS versus PMS groups, at the same time in larger samples. Optimally, patients would be prescribed symptomatic treatment which is most appropriate for their needs, and future work could provide clinicians with the necessary information for making treatment decisions.

Conclusions

These results indicate that greater response to restorative cognitive training in MS is most apparent for individuals with RRMS (relative to those with PMS) and higher Orderliness at time of treatment initiation. The same was found for MS patients with greater gray matter volume previous to cognitive treatment initiation. Moreover, response to restorative CR is also predicted by structural network disruption between regions associated with the default-mode network, and this effect is moderated by patient profiles of functional connectivity. Deviations in functional connectivity relative to HCs may indicate cognitive reserve potential and may be adaptive or pathological, depending on the burden of structural brain insult characterizing each person with MS.

GENERAL CONCLUSIONS

For many years, the vast majority of studies on multiple sclerosis focused mainly on patients' physical manifestations and disability; however, in the last decades, a considerable interest in the topic of diagnosis and management of cognitive impairment has arisen. This doctoral thesis focused on cognitive alterations in MS patients with the aim of exploring emerging topics and providing relevant implications on diagnosis, prognosis, and rehabilitation.

The first study focused on MS patients early in the disease course, since the time of diagnosis, aiming at creating different patient profiles that will evolve through different disease progression pathways, combining neurologic, neuropsychological, neuroradiological, and bioumoral outcomes. This may contribute to better clarify the heterogeneity of MS in different patients, in order to know in advance how disease will progress in each patient, and will provide important clinical information in therapeutic interventions, including the choice of the most appropriate DMTs and the most suited rehabilitative approach. The combination of different outcomes may offer possible prognostic values of specific variables early in the disease course: previous works tried to combine different measures in order to obtain prognostic values, but the relevance of this project concerns in the large number of different outcomes selected, the large number and the different provenance of patients included in this multicentric project, and that all the patients included were assessed at the time of diagnosis.

Firstly, we began to focus on cognitive functioning at baseline. Results from this study showed that cognitive alterations are already present in MS patients since the time of diagnosis, even in the absence of formal cognitive impairment. The alterations in cognitive functioning were found both at a global level and at specific domain levels (i.e. memory,

attention/information processing speed, executive functions). Particularly relevant was the executive functions alteration, that might be neglected considering that the two most commonly used neuropsychological batteries in MS (i.e., the BRB-NT and the BICAMS) do not include any executive functions test. An extended neuropsychological assessment is of paramount importance from the beginning of the disease course, in order to early identify slight cognitive alterations that are missed in the traditional approach, in which cognition is considered in a dichotomic classification of "normal" vs. "impaired". However, there is no consensus on what has to be considered "normal cognition", since an enormous number of different classification criteria has been proposed; therefore, it seems preferable to consider cognitive functioning as a continuum, ranging from a minimum to a maximum, rather than dichotomize the cognitive variables using a large number of different criteria. Future perspectives of this study might include a greater number of MS patients, and likewise to include measures of MRI (i.e., brain atrophy) in order to investigate whether cognitive alterations reflect the same pattern of neuroradiological measures. Furthermore, longitudinal follow-up data of cognitive functioning might provide information on the progression of cognitive impairment, in association with disease progression and both inflammatory and neurodegenerative phenomena.

The second study, through the combination of neurologic, neuropsychological, neuroradiological, and bioumoral data, provided new insights about the association between these different outcomes. While the association between cognitive and MRI has been widely investigated in literature (strong correlation between cognitive impairment and both WM and GM damage), as well as the cognitive-neurologic association (correlation between cognitive impairment and physical disability, and the proved cognitive decline during a clinical relapse), the association between CSF and neuropsychological outcomes has never been deeply explored.

However, both LIGHT and parvalbumin were found to be associated with cognitive functioning at the time of diagnosis. Considering the role of LIGHT in the B-cell intrathecal activity, association between cognitive alterations and the presence of LIGHT in the CSF may reflect brain network dysfunction due to active inflammation processes. Parvalbumin, instead, may represent a valid biomarker for neurodegeneration: in addition to cognitive functioning, CSF parvalbumin levels are also correlated, both at baseline and at follow-up, with global and regional cortical thickness and with clinical disability. In this study, the level of NF-L was also investigated, as a well-known and wide used biomarker of neurodegenerative processes: however, correlation with NF-L were significant but weaker compared to parvalbumin results, suggesting that in our group of MS patients CSF parvalbumin level acted as a better biomarker of neurodegeneration. The relevance of the study was that it has been highlighted two specific biomarkers (one resembling an inflammatory aspect and the other resembling a neurodegenerative aspect) that, at baseline, are significantly associated with cognitive functioning, even more than other well-established biomarkers. Future perspectives might include a larger number of patients and, also, might include longitudinal data for cognitive functioning and greater time at follow-up for neurological and neuroradiological outcomes, to further investigate and establish the prognostic values of these two biomarkers.

Thirdly, the 4-years follow-up study highlighted the specific role of some CSF molecules (CXCL13, CXCL12, IFNγ, TNF, TWEAK, LIGHT, sCD163) in discriminating patients that will show a more severe disease course and a worse cortical pathology outcome over time. The predictive value since the time of diagnosis might help clinicians to know in advance how MS will progress and which is the most appropriate DMT for each patient, according to the hypothesis of shaping different MS profiles since the earliest stages of the disease and to identify the best pharmacological treatment for each patient. While CSF molecules showed a good predictive

value, for cognitive impairment at the time of diagnosis no prognostic effect was found at 4 years of follow-up. A possible explanation of this finding might be that patients at the time of diagnosis and after 4 years may have still a relative low level of accumulated neurodegeneration, due to the short disease duration, that will not exacerbate in severe cognitive worsening. However, the study highlighted the CSF-cytokines prognostic values in predicting new evidences of MS activity after the following 4 years. Future studies might include a higher number of patients and a longer follow-up period: this could be particularly relevant for cognitive impairment to occur, in order to investigate which sort of CSF profile showed at baseline those patients that will develop cognitive impairment at follow-up. Otherwise, to better understand the bioumoral prognostic value on cognitive alterations, future analyses could be performed using cognitive functioning as a variable, instead of dichotomizing cognitive impairment, considering that slight alterations in cognitive functioning arose sharply in advance compared to a more global cognitive impairment.

In MS patients, traditional cognitive assessment based on classical neuropsychological batteries (i.e. BICAMS, BRB-NT, MACFIMS, …) might be not sensitive enough to detect alterations in cognitive functioning. In this regard, the potential role of false memories was reported, a well-known psychological phenomenon, to better understand the characteristic of the semantic memory in MS patients. The study highlighted that MS patients, compared to a matched group of HCs, were less likely to produce false memories and were less confident to recognize stimuli never presented. These results might contribute in providing knowledge about semantic networks architecture in MS patients: a few numbers of nodes, besides alterations in connection between the nodes, might explain the reduced probability to produce false recalls and false recognition, nevertheless it is a normal effect that was found in HCs. It is important to underline that the

false memories effect found in the MS group could not be explained by verbal memory alterations, because the same results were found in MS patients without any evidence of cognitive impairment. Future studies might include follow-up data, in order to evaluate the impact of MS progression on semantic memory alterations; in addition, future works might investigate functioning of semantic memory in MS patients through MRI sequences of advanced structural neuroimaging, that could be able to provide knowledge on semantic network structure and organization. Further investigations by using a DRM paradigm might provoke false memories by means of visual materials instead of auditory stimuli, or might use lists of words phonologically related, instead of semantically related, to better understand different mechanisms underlying verbal memory in MS patients.

A group of MS patients without evidence of cognitive impairment were also evaluated with a specific protocol aimed to assess the social cognition domain, i.e. theory of mind, facial emotion recognition, and empathy. Despite all MS patients were classified as “cognitive normal”, they were characterized by a significantly lower performance in all the social cognition measures compared to a control group. By evaluating the MRI scans of these patients, it was found that theory of mind and facial emotion recognition impairment were associated with cortical lesion volume in the amygdalae (both left and right), which are the main brain hubs for processing and identify emotions and emotive status. In particular, identification of emotions of angry and afraid faces was difficult to be discriminated and elaborated by MS patients, probably due to the high association between amygdalae and these two emotional states. These results highlighted that social cognition alterations can occur in MS patients independently from the level of global cognitive functioning. Recently, social cognition has been inserted in the new version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) as one of the six main cognitive domains (together with complex attention, learning and memory,

executive functions, perceptual-motor function, and language), suggesting the importance of deeply investigating these abilities in the neuropsychological assessment. Future perspective of this study might investigate at follow-up both traditional cognitive domains and also social cognition processes, in order to understand whether, after a certain period of time, cognitive impairment and social cognition impairment still follow two different pathways of progression, as demonstrated at baseline, and whether the cortical lesion burden in the amygdalae at baseline, associated with social cognition impairment, will predict functional impairment in terms of social activities, employment and quality of life.

In the light of both these two studies on false memories and social cognition, it seems of fundamental importance to overcome the classical neuropsychological examination and further and deeper investigate wider cognitive aspects in MS patients by means of using specific paradigms.

In order to contrast cognitive decline associated with disease progression, cognitive rehabilitation programs should be administered promptly to MS patients. One of the major challenges for clinicians is to understand which patient is best suited for different existing types of cognitive rehabilitation. The single-arm study performed at the Buffalo Neuroimaging Analysis Center provided important evidence about which clinical, neuropsychological, structural MRI, and functional MRI variables seem to be predictive of a better improvement after a 3-months period of a restorative cognitive training. Behavioral results showed how MS patients with diagnosis of RRMS and higher conscientiousness personality trait reported greater cognitive impairment after the training period, while MRI outcomes better associated with cognitive improvement were found to be a higher gray matter volume and a lower WM tract disruption in a network of brain regions centered on the precuneus and posterior cingulate, both key areas of the default-mode network (previously demonstrated to have an

important role in cognitive functioning). Furthermore, fMRI data analysis showed how MS patients with the greatest post-treatment improvement were those with a high functional deviation from the HCs norms, meaning that deviations in functional connectivity may indicate relative cognitive reserve potential and may be adaptive or pathological, on the basis of the structural damage characterizing each MS patients. Taking together, these results provide encouraging evidence about the possibility to know in advance how much each patient would benefit from a cognitive rehabilitation program, allowing to maximize intervention effect and to choose the most suited rehabilitation treatment for each MS patient profile. Future perspectives of this study might include additional factors that could play a role in determining neuropsychological rehabilitation improvement, and, in addition, might recruit a larger number of MS patients. Moreover, since the rehabilitation software that was used showed a significantly greater improvement for RRMS compared to patients with a progressive phase, it would be interesting to run the same experimental design using a software that may be more specific for progressive patients (i.e. a neuropsychological rehabilitation treatment based on a compensatory approach).

In conclusion, despite MS patients are characterized by heterogenous burdens of neuroinflammation and neurodegeneration, these processes leads to different levels of brain damages that consequently result in alterations in cognitive functioning. From a translational perspective, the results of the studies presented suggest that a comprehensive neuropsychological assessment, therefore not limited to the traditional cognitive domains evaluations and in which a functional approach is preferred instead of a dichotomic one, should become a fundamental part of everyday clinical practice of MS patients, both at baseline and also at follow-up, in order to monitor cognitive functioning decline progression. Since cognitive alterations are present since the time of the diagnosis and lead to

devastating impact on psychological aspects of both patients and their caregivers, neuropsychological rehabilitation treatments should be offered to all MS patients as early as possible in the disease course.

REFERENCES

Abad, E., Sepulcre, J., Martinez-Lapiscina, E. H., Zubizarreta, I., Garcia-Ojalvo, J., & Villoslada, P. (2015). The analysis of semantic networks in multiple sclerosis identifies preferential damage of long-range connectivity. *Multiple Sclerosis and Related Disorders*, *4*(5), 387-94.

Aboulafia-Brakha, T., Christe, B., Martory, M. D., & Annoni, J. M. (2011). Theory of mind tasks and executive functions: a systematic review of group studies in neurology. *Journal of Neuropsychology*, *5*(1), 39-55.

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., et al. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in Neuroinformatics*, *8*(14), 1-10.

Aloisi, F., & Pujol-Borrell, R. (2006). Lymphoid neogenesis in chronic inflammatory diseases. *Nature Reviews Immunology*, *6*(3), 205-17.

Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *British Medical Journal*, *332*(7549), 1080.

Amato, M. P., Ponziani, G., Siracusa, G., & Sorbi, S. (2001). Cognitive dysfunction in early-onset multiple sclerosis: a reappraisal after 10 years. *Archives of Neurology*, *58*(10), 1602-6.

Amato, M. P., Portaccio, E., Goretti, B., Zipoli, V., Ricchiuti, L., De Caro, M. F., et al. (2006a). The Rao's Brief Repeatable Battery and Stroop Test: normative values with age, education and gender corrections in an Italian population. *Multiple Sclerosis Journal*, *12*(6), 787-93.

Amato, M. P., Zipoli, V., Goretti, B., Portaccio, E., DeCaro, M. F., Ricchiuti, L., et al. (2006b). Benign multiple sclerosis: cognitive, psychological and social aspects in a clinical cohort. *Journal of Neurology, 253*(8), 1054-9.

Amato, M. P., Zipoli, V., & Portaccio, E. (2006). Multiple sclerosis-related cognitive changes: a review of cross-sectional and longitudinal studies. *Journal of the Neurological Sciences*, *245*(1-2), 41-6.

Amato, M. P., Portaccio, E., Goretti, B., Zipoli, V., Hakiki, B., Giannini, M., et al. (2010). Cognitive impairment in early stages of multiple sclerosis. *Neurological Sciences*, *31*(2), 211-4.

Amato, M. P., Morra, V. B., Falautano, M., Ghezzi, A., Goretti, B., Patti, F., et al. (2018). Cognitive assessment in multiple sclerosis—an Italian consensus. *Neurological Sciences*, *39*(8), 1317-24.

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage*, *54*(3), 2033-44.

Azouvi, P., Jacquin-Courtois, S., & Luaute, J. (2017). Rehabilitation of unilateral neglect: Evidence-based medicine. *Annals of Physical and Rehabilitation Medicine*, *60*(3), 191-7.

Banati, M., Sandor, J., Mike, A., Illes, E., Bors, L., Feldmann, A., et al. (2010). Social cognition and Theory of Mind in patients with relapsing-remitting multiple sclerosis. *European Journal of Neurology*, *17*(3), 426-33.

Baraczka, K., Nekam, K., Pozsonyi, T., Szüts, I., & Ormos, G. (2004). Investigation of cytokine (tumor necrosis factor-alpha, interleukin-6, interleukin-10) concentrations in the cerebrospinal fluid of female patients with multiple sclerosis and systemic lupus erythematosus. *European Journal of Neurology*, *11*(1), 37-42.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42*(2), 241-51.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163-75.

Batista, S., Zivadinov, R., Hoogs, M., Bergsland, N., Heininen-Brown, M., Dwyer, M. G., et al. (2012). Basal ganglia, thalamus and neocortical atrophy predicting slowed cognitive processing in multiple sclerosis. *Journal of Neurology*, *259*(1), 139-46.

Batista, S., Alves, C., d'Almeida, O. C., Afonso, A., Félix-Morais, R., Pereira, J., et al. (2017a). Disconnection as a mechanism for social cognition impairment in multiple sclerosis. *Neurology*, *89*(1), 38-45.

Batista, S., d'Almeida, O. C., Afonso, A., Freitas, S., Macário, C., Sousa, L., et al. (2017b). Impairment of social cognition in multiple sclerosis: amygdala atrophy is the main predictor. *Multiple Sclerosis Journal*, *23*(10), 1358-66.

Batista, S., Freitas, S., Afonso, A., Macário, C., Sousa, L., Cunha, L., et al. (2017c). Theory of mind and executive functions are dissociated in multiple sclerosis. *Archives of Clinical Neuropsychology*, *33*(5), 541-51.

Beers, D. R., Ho, B. K., Siklós, L., Alexianu, M. E., Mosier, D. R., Mohamed, A. H., et al. (2001). Parvalbumin overexpression alters immune-mediated increases in intracellular calcium, and delays disease onset in a transgenic model of familial amyotrophic lateral sclerosis. *Journal of Neurochemistry*, *79*(3), 499-509.

Benedict, R. H. B. (1997). *Brief visuospatial memory test-revised: professional manual*. PAR.

Benedict, R. H. B., Bakshi, R., Simon, J. H., Priore, R., Miller, C., Munschauer, F. (2002). Frontal cortex atrophy predicts cognitive impairment in multiple sclerosis. *Journal of Neuropsychiatry and Clinical Neuroscience, 14*(1), 44-51.

Benedict, R. H. B., Fishman, I., McClellan, M. M., Bakshi, R., & Weinstock-Guttman, B. (2003). Validity of the beck depression inventory-fast screen in multiple sclerosis. *Multiple Sclerosis Journal*, *9*(4), 393-6.

Benedict, R. H. B., Cox, D., Thompson, L. L., Foley, F., Weinstock-Guttman, B., & Munschauer, F. (2004). Reliable screening for neuropsychological impairment in multiple sclerosis. *Multiple Sclerosis Journal*, *10*(6), 675-8.

Benedict, R. H. B., 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). *Journal of the International Neuropsychological Society*, *12*(4), 549-58.

Benedict, R. H. B., Duquin, J. A., Jurgensen, S., Rudick, R. A., Feitcher, J., Munschauer, F. E., et al. (2008a). Repeated assessment of neuropsychological deficits in multiple sclerosis using the Symbol Digit Modalities Test and the MS Neuropsychological Screening Questionnaire. *Multiple Sclerosis Journal*, *14*(7), 940-6.

Benedict, R. H. B., Hussein, S., Englert, J., Dwyer, M. G., Abdelrahman, N., Cox, J. L., et al. (2008b). Cortical atrophy and personality in multiple sclerosis. *Neuropsychology*, *22*(4), 432-41.

Benedict, R. H. B., & Zivadinov, R. (2011). Risk factors for and management of cognitive dysfunction in multiple sclerosis. *Nature Reviews Neurology*, *7*(6), 332-42.

Benedict, R. H. B., Amato, M. P., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., et al. (2012). Brief International Cognitive Assessment for MS (BICAMS): international standards for validation. *BMC Neurology*, *12*(1), 55-62.

Benedict, R. H. B., DeLuca, J., Enzinger, C., Geurts, J. J., Krupp, L. B., & Rao, S. M. (2017a). Neuropsychology of multiple sclerosis: looking back and moving forward. *Journal of the International Neuropsychological Society*, *23*(9-10), 832-42.

Benedict, R. H. B., DeLuca, J., Phillips, G., LaRocca, N., Hudson, L. D., Rudick, R., et al. (2017b). Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *Multiple Sclerosis Journal*, *23*(5), 721-33.

Bergsland, N., Horakova, D., Dwyer, M. G., Uher, T., Vaneckova, M., Tyblova, M., et al. (2018). Gray matter atrophy patterns in multiple sclerosis: A 10-year source-based morphometry study. *NeuroImage: Clinical*, *17*, 444-51.

Berneiser, J., Wendt, J., Grothe, M., Kessler, C., Hamm, A. O., & Dressel, A. (2014). Impaired recognition of emotional facial expressions in patients with multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *3*(4), 482-8.

Bernhardt, B. C., & Singer, T. (2012). The neural basis of empathy. *Annual Review of Neuroscience*, *35*, 1-23.

Bickart, K. C., Dickerson, B. C., & Barrett, L. F. (2014). The amygdala as a hub in brain networks that support social life. *Neuropsychologia*, *63*, 235-48.

Binder, L. M., Iverson, G. L., & Brooks, B. L. (2009). To err is human:"Abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology*, *24*(1), 31-46.

Bitanihirwe, B. K., Lim, M. P., Kelley, J. F., Kaneko, T., & Woo, T. U. W. (2009). Glutamatergic deficits and parvalbumin-containing inhibitory neurons in the prefrontal cortex in schizophrenia. *BMC Psychiatry*, *9*(71).

Bobholz, J. A., & Rao, S. M. (2003). Cognitive dysfunction in multiple sclerosis: a review of recent developments. *Current Opinion in Neurology*, *16*(3), 283-8.

Bora, E., Özakbaş, S., Velakoulis, D., & Walterfang, M. (2016). Social cognition in multiple sclerosis: a meta-analysis. *Neuropsychology Review*, *26*(2), 160-72.

Bottesi, G., Ghisi, M., Altoè, G., Conforti, E., Melli, G., & Sica, C. (2015). The Italian version of the Depression Anxiety Stress Scales-21: Factor structure and psychometric properties on community and clinical samples. *Comprehensive Psychiatry*, *60*, 170-81.

Brissart, H., Morele, E., Baumann, C., & Debouverie, M. (2012). Verbal episodic memory in 426 multiple sclerosis patients: impairment in encoding, retrieval or both? *Neurological Sciences*, *33*(5), 1117-23.

Bronnum-Hansen, H., Stenager, E., Hansen, T., & Koch-Henriksen, N. (2006). Survival and mortality rates among Danes with MS. *International MS Journal*, *13*(2), 66-71.

Browne, P., Chandraratna, D., Angood, C., Tremlett, H., Baker, C., Taylor, B. V., et al. (2014). Atlas of multiple sclerosis 2013: a growing global problem with widespread inequity. *Neurology*, *83*(11), 1022-24.

Caffarra, P., Vezzadini, G., Dieci, F., Zonato, F., & Venneri, A. (2002). Una versione abbreviata del test di Stroop: dati normativi nella popolazione italiana. *Nuova Rivista di Neurologia*, *12*(4), 111-5.

Calabrese, M., Atzori, M., Bernardi, V., Morra, A., Romualdi, C., Rinaldi, L., et al. (2007). Cortical atrophy is relevant in multiple sclerosis at clinical onset. *Journal of Neurology*, *254*(9), 1212-20.

Calabrese, M., Agosta, F., Rinaldi, F., Mattisi, I., Grossi, P., Favaretto, A., et al. (2009). Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis. *Archives of Neurology*, *66*(9), 1144-50.

Calabrese, M., Rinaldi, F., Mattisi, I., Grossi, P., Favaretto, A., Atzori, M., et al. (2010). Widespread cortical thinning characterizes patients with MS with mild cognitive impairment. *Neurology*, *74*(4), 321-8.

Calabrese, M., Poretto, V., Favaretto, A., Alessio, S., Bernardi, V., Romualdi, C., et al. (2012). Cortical lesion load associates with progression of disability in multiple sclerosis. *Brain*, *135*(10), 2952-61.

Calabrese, M., Magliozzi, R., Ciccarelli, O., Geurts, J. J., Reynolds, R., & Martin, R. (2015a). Exploring the origins of grey matter damage in multiple sclerosis. *Nature Reviews Neuroscience*, *16*(3), 147-58.

Calabrese, M., Reynolds, R., Magliozzi, R., Castellaro, M., Morra, A., Scalfari, A., et al. (2015b). Regional distribution and evolution of gray matter damage in different populations of multiple sclerosis patients. *PloS One*, *10*(8), e0135428.

Calabrese, M., & Pitteri, M. (2018). Cognition and fatigue in multiple sclerosis. In *Cognition and behavior in multiple sclerosis*. American Psychological Association, 127-48.

Calabresi, P. A. (2004). Diagnosis and management of multiple sclerosis. *American Family Physician*, *70*(10), 1935-44.

Cattelani, R., Dal Sasso, F., Corsini, D., & Posteraro, L. (2011). The Modified Five-Point Test: normative data for a sample of Italian healthy adults aged 16–60. *Neurological Sciences*, *32*(4), 595-601.

Cecchetto, C., Aiello, M., D'Amico, D., Cutuli, D., Cargnelutti, D., Eleopra, R., et al. (2014). Facial and bodily emotion recognition in multiple sclerosis: the role of alexithymia and other characteristics of the disease. *Journal of the International Neuropsychological Society*, *20*(10), 1004-14.

Chalah, M. A., & Ayache, S. S. (2017). Deficits in social cognition: an unveiled signature of multiple sclerosis. *Journal of the International Neuropsychological Society*, *23*(3), 266-86.

Chapman, B. P. (2007). Bandwidth and fidelity on the NEO-five factor inventory: replicability and reliability of item cluster subcomponents. *Journal of Personality Assessment*, *88*(2), 220-34.

Chard, D., & Miller, D. (2009). Grey matter pathology in clinically early multiple sclerosis: evidence from magnetic resonance imaging. *Journal of the Neurological Sciences*, *282*(1-2), 5-11.

Charil, A., Dagher, A., Lerch, J. P., Zijdenbos, A. P., Worsley, K. J., & Evans, A. C. (2007). Focal cortical atrophy in multiple sclerosis: relation to lesion load and disability. *Neuroimage*, *34*(2), 509-17.

Charvet, L. E., Yang, J., Shaw, M. T., Sherman, K., Haider, L., Xu, J., et al. (2017). Cognitive function in multiple sclerosis improves with telerehabilitation: Results from a randomized controlled trial. *PloS One*, *12*(5), e0177177.

Charvet, L. E., Shaw, M., Frontario, A., Langdon, D., & Krupp, L. B. (2018). Cognitive impairment in pediatric-onset multiple sclerosis is detected by the Brief International Cognitive Assessment for Multiple Sclerosis and computerized cognitive testing. *Multiple Sclerosis Journal*, *24*(4), 512-9.

Chiaravalloti, N. D., & DeLuca, J. (2008). Cognitive impairment in multiple sclerosis. *The Lancet Neurology*, *7*(12), 1139-51.

Chiaravalloti, N. D., Moore, N. B., & DeLuca, J. (2019). The efficacy of the modified Story Memory Technique in progressive MS. *Multiple Sclerosis Journal, 26*(3), 354-62.

Ciampi, E., Uribe-San-Martin, R., Vásquez, M., Ruiz-Tagle, A., Labbe, T., Cruz, J. P., et al. (2018). Relationship between Social Cognition and traditional cognitive impairment in Progressive Multiple Sclerosis and possible implicated neuroanatomical regions. *Multiple Sclerosis and Related Disorders*, *20*, 122-8.

Clements, R. J., McDonough, J., & Freeman, E. J. (2008). Distribution of parvalbumin and calretinin immunoreactive interneurons in motor cortex from multiple sclerosis post-mortem tissue. *Experimental Brain Research*, *187*(3), 459-65.

Costa, A., Bagoj, E., Monaco, M., Zabberoni, S., De Rosa, S., Papantonio, A. M., et al. (2014). Standardization and normative data obtained in the Italian population for a new verbal fluency instrument, the phonemic/semantic alternate fluency test. *Neurological Sciences*, *35*(3), 365-72.

Cotter, J., Firth, J., Enzinger, C., Kontopantelis, E., Yung, A. R., Elliott, R., et al. (2016). Social cognition in multiple sclerosis: a systematic review and meta-analysis. *Neurology*, *87*(16), 1727-36.

Daams, M., Steenwijk, M. D., Schoonheim, M. M., Wattjes, M. P., Balk, L. J., Tewarie, P. K., et al. (2016). Multi-parametric structural magnetic resonance imaging in relation to cognitive dysfunction in long-standing multiple sclerosis. *Multiple Sclerosis Journal, 22*(5), 608-19.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage*, *9*(2), 179-94.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*(1), 17-22.

Dekkers, J., Bayley, P., Dick, J. R. T., Schwaller, B., Berchtold, M. W., & Greensmith, L. (2004). Over-expression of parvalbumin in transgenic mice rescues motoneurons from injury-induced cell death. *Neuroscience*, *123*(2), 459-66.

Delis, D. C. (2000). California verbal learning test. *Adult version manual.* Psychological Corporation.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System®(D-KEFS®): Examiner's Manual: Flexibility of Thinking, Concept Formation, Problem Solving, Planning, Creativity, Impulse Control, Inhibition*. Pearson.

DeLuca, J., Gaudino, E. A., Diamond, B. J., Christodoulou, C., & Engel, R. A. (1998). Acquisition and storage deficits in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *20*(3), 376-90.

Deloire, M., Ruet, A., Hamel, D., Bonnet, M., & Brochet, B. (2010). Early cognitive impairment in multiple sclerosis predicts disability outcome several years later. *Multiple Sclerosis Journal*, *16*(5), 581-7.

De Meijer, L., Merlo, D., Skibina, O., Grobbee, E. J., Gale, J., Haartsen, J., et al. (2018). Monitoring cognitive change in multiple sclerosis using a computerized cognitive battery. *Multiple Sclerosis Journal – Experimental, Translational and Clinical*, *4*(4), 1-9.

De Stefano, N., Guidi, L., Stromillo, M. L., Bartolozzi, M. L., & Federico, A. (2003). Imaging neuronal and axonal degeneration in multiple sclerosis. *Neurological Sciences*, *24*(5), s283-6.

Dobryakova, E., Rocca, M. A., & Filippi, M. (2018). Cerebral reorganization and cognition in multiple sclerosis.

Duan, H., Luo, Y., Hao, H., Feng, L., Zhang, Y., Lu, D., et al. (2013). Soluble CD146 in cerebrospinal fluid of active multiple sclerosis. *Neuroscience*, *235*, 16-26.

Dulau, C., Deloire, M., Diaz, H., Saubusse, A., Charre-Morin, J., Prouteau, A., et al. (2017). Social cognition according to cognitive impairment in different clinical phenotypes of multiple sclerosis. *Journal of Neurology*, *264*(4), 740-8.

Duncan, J., Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neuroscience, 23*(10), 475-83.

Dutta, R., McDonough, J., Yin, X., Peterson, J., Chang, A., Torres, T., et al. (2006). Mitochondrial dysfunction as a cause of axonal degeneration in multiple sclerosis patients. *Annals of Neurology*, *59*(3), 478-89.

Eijlers, A. J., Meijer, K. A., Wassenaar, T. M., Steenwijk, M. D., Uitdehaag, B. M., Barkhof, F., et al. (2017). Increased default-mode network centrality in cognitively impaired multiple sclerosis patients. *Neurology*, *88*(10), 952-60.

Eijlers, A. J., van Geest, Q., Dekker, I., Steenwijk, M. D., Meijer, K. A., Hulst, H. E., et al. (2018). Predicting cognitive decline in multiple sclerosis: a 5-year follow-up study. *Brain*, *141*(9), 2605-18.

Esen, N., Rainey-Barger, E. K., Huber, A. K., Blakely, P. K., & Irani, D. N. (2014). Type-I interferons suppress microglial production of the lymphoid chemokine, CXCL13. *Glia*, *62*(9), 1452-62.

Eshaghi, A., Marinescu, R. V., Young, A. L., Firth, N. C., Prados, F., Jorge Cardoso, M., et al. (2018). Progression of regional gray matter atrophy in multiple sclerosis. *Brain*, *141*(6), 1665-77.

Eshel, I., Bowles, A. O., & Ray, M. R. (2019). Rehabilitation of cognitive dysfunction following traumatic brain injury. *Physical Medicine and Rehabilitation Clinics*, *30*(1), 189-206.

Farina, G., Magliozzi, R., Pitteri, M., Reynolds, R., Rossi, S., Gajofatto, A., et al. (2017). Increased cortical lesion load and intrathecal inflammation is associated with oligoclonal bands in multiple sclerosis patients: a combined CSF and MRI study. *Journal of Neuroinflammation*, *14*(1), 40-51.

Feinstein, A. (2011). Multiple sclerosis and depression. *Multiple Sclerosis Journal*, *17*(11), 1276-81.

Ferraro, D., Galli, V., Vitetta, F., Simone, A. M., Bedin, R., Del Giovane, C., et al. (2015). Cerebrospinal fluid CXCL13 in clinically isolated syndrome patients: association with oligoclonal IgM bands and prediction of multiple sclerosis diagnosis. *Journal of Neuroimmunology*, *283*, 64-9.

Fialová, L., Bartos, A., Švarcová, J., Zimova, D., Kotoucova, J., & Malbohan, I. (2013). Serum and cerebrospinal fluid light neurofilaments and antibodies against them in clinically isolated syndrome and multiple sclerosis. *Journal of Neuroimmunology*, *262*(1-2), 113-20.

Filippi, M., van den Heuvel, M. P., Fornito, A., He, Y., Pol, H. E. H., Agosta, F., et al. (2013). Assessment of system dysfunction in the brain through MRI-based connectomics. *The Lancet Neurology*, *12*(12), 1189-99.

Filippi, M., Preziosa, P., & Rocca, M. A. (2018). MRI in multiple sclerosis: what is changing? *Current Opinion in Neurology*, *31*(4), 386-95.

Fischer, M., Kunkel, A., Bublak, P., Faiss, J. H., Hoffmann, F., Sailer, M., et al. (2014). How reliable is the classification of cognitive impairment across different criteria in early and late stages of multiple sclerosis? *Journal of the Neurological Sciences*, *343*(1-2), 91-9.

Fischl, B., & Dale, A. M. (2000). Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences*, *97*(20), 11050-5.

Fischl, B. (2012). FreeSurfer. *Neuroimage*, *62*(2), 774-81.

Fisniku, L. K., Chard, D. T., Jackson, J. S., Anderson, V. M., Altmann, D. R., Miszkiel, K. A., et al. (2008). Gray matter atrophy is related to long-term disability in multiple sclerosis. *Annals of Neurology*, *64*(3), 247-54.

Forn, C., Rocca, M. A., Valsasina, P., Boscá, I., Casanova, B., Sanjuan, A., et al. (2012). Functional magnetic resonance imaging correlates of cognitive performance in patients with a clinically isolated syndrome suggestive of multiple sclerosis at presentation: an activation and connectivity study. *Multiple Sclerosis Journal*, *18*(2), 153-63.

Fox, R. J., Miller, D. H., Phillips, J. T., Hutchinson, M., Havrdova, E., Kita, M., et al. (2012). Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. *New England Journal of Medicine*, *367*(12), 1087-97.

Frischer, J. M., Bramow, S., Dal-Bianco, A., Lucchinetti, C. F., Rauschka, H., Schmidbauer, M., et al. (2009). The relation between inflammation and neurodegeneration in multiple sclerosis brains. *Brain*, *132*(5), 1175-89.

Fuchs, T. A., Dwyer, M. G., Kuceyeski, A., Choudhery, S., Carolus, K., Li, X., et al. (2018). White matter tract network disruption explains reduced conscientiousness in multiple sclerosis. *Human Brain Mapping*, *39*(9), 3682-90.

Fuchs, T. A., Benedict, R. H. B., Bartnik, A., Choudhery, S., Li, X., Mallory, M., et al. (2019a). Preserved network functional connectivity underlies cognitive reserve in multiple sclerosis. *Human Brain Mapping, 40*(18), 5231-41.

Fuchs, T. A., Wojcik, C., Wilding, G. E., Pol, J., Dwyer, M. G., Weinstock-Guttman, B., et al. (2019b). Trait Conscientiousness predicts rate of longitudinal SDMT decline in multiple sclerosis. *Multiple Sclerosis Journal, 26*(2), 245-52.

Gaetani, L., Blennow, K., Calabresi, P., Di Filippo, M., Parnetti, L., & Zetterberg, H. (2019). Neurofilament light chain as a biomarker in neurological disorders. *Journal of Neurology, Neurosurgery & Psychiatry*, *90*(8), 870-81.

Gardner, C., Magliozzi, R., Durrenberger, P. F., Howell, O. W., Rundle, J., & Reynolds, R. (2013). Cortical grey matter demyelination can be induced by elevated

pro-inflammatory cytokines in the subarachnoid space of MOG-immunized rats. *Brain*, *136*(12), 3596-608.

Gajofatto, A., & Benedetti, M. D. (2015). Treatment strategies for multiple sclerosis: when to start, when to change, when to stop? *World Journal of Clinical Cases*, *3*(7), 545-55.

Gelfand, J. M. (2014). Multiple sclerosis: diagnosis, differential diagnosis, and clinical presentation. In *Handbook of clinical neurology*, *122*, 269-90.

Genova, H. M., Cagna, C. J., Chiaravalloti, N. D., DeLuca, J., & Lengenfelder, J. (2016). Dynamic assessment of social cognition in individuals with multiple sclerosis: a pilot study. *Journal of the International Neuropsychological Society*, *22*(1), 83-8.

Geurts, J. J., Stys, P. K., Minagar, A., Amor, S., & Zivadinov, R. (2009). Gray matter pathology in (chronic) MS: modern views on an early observation. *Journal of the Neurological Sciences*, *282*(1-2), 12-20.

Geurts, J. J. G., Roosendaal, S. D., Calabrese, M., Ciccarelli, O., Agosta, F., Chard, D. T., et al. (2011). Consensus recommendations for MS cortical lesion scoring using double inversion recovery MRI. *Neurology*, *76*(5), 418-24.

Geurts, J. J., Calabrese, M., Fisher, E., & Rudick, R. A. (2012). Measurement and clinical effect of grey matter pathology in multiple sclerosis. *The Lancet Neurology*, *11*(12), 1082-92.

Giovannoni, G., Tomic, D., Bright, J. R., & Havrdová, E. (2017). "No evident disease activity": the use of combined assessments in the management of patients with multiple sclerosis. *Multiple Sclerosis Journal*, *23*(9), 1179-87.

Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: a validation study. *Cognition and Emotion*, *22*(6), 1094-118.

Golan, D., Wilken, J., Fratto, T., Kane, R., Srinivasan, J., Zarif, M., et al. (2019). Validity and Potential Added Value of Multi-Domain Computerized Cognitive Assessment for Patients with Multiple Sclerosis. *Multiple Sclerosis and Related Disorders, 30*, 154-62.

Goldman, M. D., Motl, R. W., Scagnelli, J., Pula, J. H., Sosnoff, J. J., & Cadavid, D. (2013). Clinically meaningful performance benchmarks in MS: timed 25-foot walk and the real world. *Neurology*, *81*(21), 1856-63.

Goñi, J., Arrondo, G., Sepulcre, J., Martincorena, I., de Mendizábal, N. V., Corominas-Murtra, B., et al. (2011). The semantic organization of the animal category: evidence from semantic verbal fluency and network theory. *Cognitive Processing*, *12*(2), 183-96.

Goverover, Y., Chiaravalloti, N. D., O'Brien, A. R., & DeLuca, J. (2018). Evidenced-based cognitive rehabilitation for persons with multiple sclerosis: an updated review of the literature from 2007 to 2016. *Archives of Physical Medicine and Rehabilitation*, *99*(2), 390-407.

Graumann, U., Reynolds, R., Steck, A. J., & Schaeren-Wiemers, N. (2003). Molecular changes in normal appearing white matter in multiple sclerosis are characteristic of neuroprotective mechanisms against hypoxic insult. *Brain Pathology*, *13*(4), 554-73.

Green, R., Adler, A., Banwell, B. L., Fabri, T. L., Yeh, E. A., Collins, D. L., et al. (2018). Involvement of the Amygdala in Memory and Psychosocial Functioning in Pediatric-Onset Multiple Sclerosis. *Developmental Neuropsychology*, *43*(6), 524-34.

Griffanti, L., Douaud, G., Bijsterbosch, J., Evangelisti, S., Alfaro-Almagro, F., Glasser, M. F., et al. (2017). Hand classification of fMRI ICA noise components. *Neuroimage*, *154*, 188-205.

Grzegorski, T., & Losy, J. (2017). Cognitive impairment in multiple sclerosis–a review of current knowledge and recent research. *Reviews in the Neurosciences*, *28*(8), 845-60.

Haider, L., Zrzavy, T., Hametner, S., Höftberger, R., Bagnato, F., Grabner, G., et al. (2016). The topograpy of demyelination and neurodegeneration in the multiple sclerosis brain. *Brain, 139*(3), 807-15.

Handel, A. E., Giovannoni, G., Ebers, G. C., & Ramagopalan, S. V. (2010). Environmental factors and their timing in adult-onset multiple sclerosis. *Nature Reviews Neurology*, *6*(3), 156-66.

Harris, V. K., Diamanduros, A., Good, P., Zakin, E., Chalivendra, V., Sadiq, S. A. (2010). Bri2-23 is a potential cerebrospinal fluid biomarker in multiple sclerosis. *Neurobiology of Disease, 40*(1), 331-9.

Harrison, D. M., Roy, S., Oh, J., Izbudak, I., Pham, D., Courtney, S., et al. (2015). Association of cortical lesion burden on 7-T magnetic resonance imaging with cognition and disability in multiple sclerosis. *JAMA Neurology*, *72*(9), 1004-12.

Havrdova, E., Galetta, S., Stefoski, D., & Comi, G. (2010). Freedom from disease activity in multiple sclerosis. *Neurology*, *74*(17), S3-7.

Hawellek, D. J., Hipp, J. F., Lewis, C. M., Corbetta, M., & Engel, A. K. (2011). Increased functional connectivity indicates the severity of cognitive impairment in multiple sclerosis. *Proceedings of the National Academy of Sciences*, *108*(47), 19066-71.

Henry, J. D., Phillips, L. H., Beatty, W. W., McDonald, S., Longley, W. A., Joscelyne, A. M. Y., et al. (2009). Evidence for deficits in facial affect recognition and theory of mind in multiple sclerosis. *Journal of the International Neuropsychological Society*, *15*(2), 277-85.

Hommes, O. R., Sørensen, P. S., Fazekas, F., Enriquez, M. M., Koelmel, H. W., Fernandez, O., et al. (2004). Intravenous immunoglobulin in secondary progressive multiple sclerosis: randomised placebo-controlled trial. *The Lancet*, *364*(9440), 1149-56.

Howell, O. W., Rundle, J. L., Garg, A., Komada, M., Brophy, P. J., & Reynolds, R. (2010). Activated microglia mediate axoglial disruption that contributes to axonal injury in multiple sclerosis. *Journal of Neuropathology & Experimental Neurology*, *69*(10), 1017-33.

Howell, O. W., Reeves, C. A., Nicholas, R., Carassiti, D., Radotra, B., Gentleman, S. M., et al. (2011). Meningeal inflammation is widespread and linked to cortical pathology in multiple sclerosis. *Brain*, *134*(9), 2755-71.

Hu, G., Liu, Y., Li, H., Zhao, D., Yang, L., Shen, J., et al. (2010). Adenovirus-mediated LIGHT gene modification in murine B-cell lymphoma elicits a potent antitumor effect. *Cellular & Molecular Immunology*, *7*(4), 296-305.

Hu, H., Gan, J., & Jonas, P. (2014). Fast-spiking, parvalbumin+ GABAergic interneurons: From cellular design to microcircuit function. *Science*, *345*(6196), 1255263.

Hulst, H. E., Gehring, K., Uitdehaag, B. M., Visser, L. H., Polman, C. H., Barkhof, F., et al. (2014). Indicators for cognitive performance and subjective cognitive complaints in multiple sclerosis: a role for advanced MRI? *Multiple Sclerosis Journal*, *20*(8), 1131-4.

Jakimovski, D., Kuhle, J., Ramanathan, M., Barro, C., Tomic, D., Hagemeier, J., et al. (2019). Serum neurofilament light chain levels associations with gray matter pathology: a 5-year longitudinal study. *Annals of Clinical and Translational Neurology, 6*(9), 1757-70.

Jiang, X., Lachance, M., & Rossignol, E. (2016). Involvement of cortical fast-spiking parvalbumin-positive basket cells in epilepsy. *Progress in Brain Research, 226*, 81-126.

Kalatha, T., Arnaoutoglou, M., Koukoulidis, T., Hatzifilippou, E., Bouras, E., Baloyannis, S., et al. (2019). Does cognitive dysfunction correlate with neurofilament light polypeptide levels in the CSF of patients with multiple sclerosis? *Journal of International Medical Research, 47*(5), 2187-98.

Kalb, R., Beier, M., Benedict, R. H. B., Charvet, L., Costello, K., Feinstein, A., et al. (2018). Recommendations for cognitive screening and management in multiple sclerosis care. *Multiple Sclerosis Journal*, *24*(13), 1665-80.

Kantarci, O., & Wingerchuk, D. (2006). Epidemiology and natural history of multiple sclerosis: new insights. *Current Opinion in Neurology*, *19*(3), 248-54.

Kappos, L. (1998). Placebo-controlled multicentre randomised trial of interferon β-1b in treatment of secondary progressive multiple sclerosis. *The Lancet*, *352*(9139), 1491-97.

Klaver, R., Popescu, V., Voorn, P., Galis-de Graaf, Y., van der Valk, P., de Vries, H. E., et al. (2015). Neuronal and axonal loss in normal-appearing gray matter and subpial lesions in multiple sclerosis. *Journal of Neuropathology & Experimental Neurology*, *74*(5), 453-8.

Komori, M., Blake, A., Greenwood, M., Lin, Y. C., Kosa, P., Ghazali, D., et al. (2015). Cerebrospinal fluid markers reveal intrathecal inflammation in progressive multiple sclerosis. *Annals of Neurology*, *78*(1), 3-20.

Kos, D., Kerckhofs, E., Nagels, G., D'hooghe, M. B., & Ilsbroukx, S. (2008). Origin of fatigue in multiple sclerosis: review of the literature. *Neurorehabilitation and Neural Repair*, *22*(1), 91-100.

Kraemer, M., Herold, M., Uekermann, J., Kis, B., Wiltfang, J., Daum, I., et al. (2013). Theory of mind and empathy in patients at an early stage of relapsing remitting multiple sclerosis. *Clinical Neurology and Neurosurgery*, *115*(7), 1016-22.

Krumbholz, M., Theil, D., Cepok, S., Hemmer, B., Kivisäkk, P., Ransohoff, R. M., et al. (2005). Chemokines in multiple sclerosis: CXCL12 and CXCL13 up-regulation is differentially linked to CNS immune cell recruitment. *Brain*, *129*(1), 200-11.

Krupp, L. B., LaRocca, N. G., Muir-Nash, J., & Steinberg, A. D. (1989). The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology*, *46*(10), 1121-3.

Kuceyeski, A., Maruta, J., Relkin, N., & Raj, A. (2013). The Network Modification (NeMo) Tool: elucidating the effect of white matter integrity changes on cortical and subcortical structural connectivity. *Brain Connectivity*, *3*(5), 451-63.

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*, *33*(11), 1444-52.

Kutzelnigg, A., Lucchinetti, C. F., Stadelmann, C., Brück, W., Rauschka, H., Bergmann, M., et al. (2005). Cortical demyelination and diffuse white matter injury in multiple sclerosis. *Brain*, *128*(11), 2705-12.

Kutzelnigg, A., & Lassmann, H. (2014). Pathology of multiple sclerosis and related inflammatory demyelinating diseases. *Handbook of Clinical Neurology, 122*, 15-58.

Langdon, D. W., Amato, M. P., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., et al. (2012). Recommendations for a brief international cognitive assessment for multiple sclerosis (BICAMS). *Multiple Sclerosis Journal, 18*(6), 891-8.

Lanoue, A. C., Blatt, G. J., & Soghomonian, J. J. (2013). Decreased parvalbumin mRNA expression in dorsolateral prefrontal cortex in Parkinson′ s disease. *Brain Research, 1531*, 37-47.

Lassmann, H. (2018). Pathogenic mechanisms associated with different clinical courses of multiple sclerosis. *Frontiers in Immunology*, *9*(3116), 1-14.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: reliability and validity of the Empathy Quotient. *Psychological Medicine*, *34*(5), 911-20.

Lisak, R. P., Benjamins, J. A., Nedelkoska, L., Barger, J. L., Ragheb, S., Fan, B., et al. (2012). Secretory products of multiple sclerosis B cells are cytotoxic to oligodendroglia in vitro. *Journal of Neuroimmunology*, *246*(1-2), 85-95.

Lublin, F. D., & Reingold, S. C. (1996). Defining the clinical course of multiple sclerosis: results of an international survey. *Neurology*, *46*(4), 907-11.

Lublin, F. D. (2014). New multiple sclerosis phenotypic classification. *European Neurology*, *72*(1), S1-5.

Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sørensen, P. S., Thompson, A. J., et al. (2014). Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology*, *83*(3), 278-86.

Lucchinetti, C. F., Popescu, B. F., Bunyan, R. F., Moll, N. M., Roemer, S. F., Lassmann, H., et al. (2011). Inflammatory cortical demyelination in early multiple sclerosis. *New England Journal of Medicine*, *365*(23), 2188-97.

Lundqvist, D., Flykt, A., & Öhman, A. (1998). The Karolinska directed emotional faces. *KDEF: Department of Clinical Neuroscience, Psychology Section Karolinska Institute Stockholm*.

Macías Islas, M. Á., & Ciampi, E. (2019). Assessment and impact of cognitive impairment in multiple sclerosis: an overview. *Biomedicines*, *7*(22), 1-19.

Magliozzi, R., Howell, O., Vora, A., Serafini, B., Nicholas, R., Puopolo, M., et al. (2007). Meningeal B-cell follicles in secondary progressive multiple sclerosis associate with early onset of disease and severe cortical pathology. *Brain*, *130*(4), 1089-104.

Magliozzi, R., Howell, O. W., Reeves, C., Roncaroli, F., Nicholas, R., Serafini, B., et al. (2010). A gradient of neuronal loss and meningeal inflammation in multiple sclerosis. *Annals of Neurology*, *68*(4), 477-93.

Magliozzi, R., Howell, O. W., Nicholas, R., Cruciani, C., Castellaro, M., Romualdi, C., et al. (2018). Inflammatory intrathecal profiles and cortical damage in multiple sclerosis. *Annals of Neurology*, *83*(4), 739-55.

Magliozzi, R., Howell, O. W., Durrenberger, P., Aricò, E., James, R., Cruciani, C., et al. (2019). Meningeal inflammation changes the balance of TNF signalling in cortical grey matter in multiple Sclerosis. *Journal of Neuroinflammation, 16*(259), 1-16.

Magliozzi, R., Cross, A. H. (2020). Can CSF biomarkers predict future MS disease activity and severity*? Multiple Sclerosis Journal*, 1-9.

Mahad, D. H., Trapp, B. D., & Lassmann, H. (2015). Pathological mechanisms in progressive multiple sclerosis. *The Lancet Neurology*, *14*(2), 183-93.

Mateer, C. A. (2005). Fundamentals of cognitive rehabilitation. *Effectiveness of Rehabilitation for Cognitive Deficits*, *21*, 29.

Mathiowetz, V., Weber, K., Kashman, N., & Volland, G. (1985). Adult norms for the nine hole peg test of finger dexterity. *The Occupational Therapy Journal of Research*, *5*(1), 24-38.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, *52*(1), 81-90.

Meijer, K. A., van Geest, Q., Eijlers, A. J. C., Geurts, J. J. G., Schoonheim, M. M., & Hulst, H. E. (2018). Is impaired information processing speed a matter of structural or functional damage in MS? *NeuroImage: Clinical*, *20*, 844-50.

Michel, L., Touil, H., Pikor, N. B., Gommerman, J. L., Prat, A., & Bar-Or, A. (2015). B cells in the multiple sclerosis central nervous system: trafficking and contribution to CNS-compartmentalized inflammation. *Frontiers in Immunology*, *6*(636), 1-12.

Mikol, D. D., Barkhof, F., Chang, P., Coyle, P. K., Jeffery, D. R., Schwid, S. R., et al. (2008). Comparison of subcutaneous interferon beta-1a with glatiramer acetate in patients with relapsing multiple sclerosis (the Rebif vs Glatiramer Acetate in Relapsing MS Disease [REGARD] study): a multicentre, randomised, parallel, open-label trial. *The Lancet Neurology*, *7*(10), 903-14.

Mitolo, M., Venneri, A., Wilkinson, I. D., & Sharrack, B. (2015). Cognitive rehabilitation in multiple sclerosis: a systematic review. *Journal of the Neurological Sciences*, *354*(1-2), 1-9.

Moccia, M., Lanzillo, R., Palladino, R., Chang, K. C. M., Costabile, T., Russo, C., et al. (2016). Cognitive impairment at diagnosis predicts 10-year multiple sclerosis progression. *Multiple Sclerosis Journal*, *22*(5), 659-67.

Modvig, S., Degn, M., Roed, H., Sørensen, T. L., Larsson, H. B. W., Langkilde, A. R., et al. (2015). Cerebrospinal fluid levels of chitinase 3-like 1 and neurofilament light chain predict multiple sclerosis development and disability after optic neuritis. *Multiple Sclerosis Journal*, *21*(14), 1761-70.

Møllgaard, M., Degn, M., Sellebjerg, F., Frederiksen, J. L., Modvig., S. (2016). Cerebrospinal fluid chitinase-3-like 2 and chitotriosidase are potential prognostic

biomarkers in early multiple sclerosis. *European Journal of Neurology, 23*(5), 898-905.

Monaco, M., Costa, A., Caltagirone, C., & Carlesimo, G. A. (2013). Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. *Neurological Sciences*, *34*(5), 749-54.

Mondini, S., Mapelli, D., Vestri, A., Arcara, G., & Bisiacchi, P. S. (2011). I test neuropsicologici dell'ENB-2. *Esame Neuropsicologico Breve 2*.

Mori, F., Rossi, S., Sancesario, G., Codecà, C., Mataluni, G., Monteleone, F., et al. (2011). Cognitive and cortical plasticity deficits correlate with altered amyloid-β CSF levels in multiple sclerosis. *Neuropsychopharmacology, 36*(3), 559-68.

Neuhaus, M., Bagutti, S., Yaldizli, Ö., Zwahlen, D., Schaub, S., Frey, B., et al. (2018). Characterization of social cognition impairment in multiple sclerosis. *European Journal of Neurology*, *25*(1), 90-6.

Nickerson, L. D., Smith, S. M., Öngür, D., & Beckmann, C. F. (2017). Using dual regression to investigate network shape and amplitude in functional connectivity analyses. *Frontiers in Neuroscience*, *13*(11), 115.

Nomi, J. S., Scherfeld, D., Friederichs, S., Schäfer, R., Franz, M., Wittsack, H. J., et al. (2008). On the neural networks of empathy: A principal component analysis of an fMRI study. *Behavioral and Brain Functions*, *4*(1), 41.

Noseworthy, J. H., Lucchinetti, C., Rodriguez, M., & Weinshenker, B. G. (2000). Medical progress. Multiple sclerosis. *New England Journal of Medicine*, *343*(13), 938-52.

Olsson, T., Barcellos, L. F., & Alfredsson, L. (2017). Interactions between genetic, lifestyle and environmental risk factors for multiple sclerosis. *Nature Reviews Neurology*, *13*(1), 25-36.

Opsahl, J. A., Vaudel, M., Guldbrandsen, A., Aasebø, E., Van Pesch, V., Franciotta, D., et al. (2016). Label-free analysis of human cerebrospinal fluid addressing various normalization strategies revealing protein groups affected by multiple sclerosis. *Proteomics, 16*(7), 1154-65.

Orton, S. M., Ramagopalan, S. V., Brocklebank, D., Herrera, B. M., Dyment, D. A., Yee, I. M., et al. C. (2010). Effect of immigration on multiple sclerosis sex ratio in Canada: the Canadian Collaborative Study. *Journal of Neurology, Neurosurgery & Psychiatry*, *81*(1), 31-6.

Ouellet, J., Scherzer, P. B., Rouleau, I., Metras, P., Bertrand-Gauvin, C., Djerroud, N., et al. (2010). Assessment of social cognition in patients with multiple sclerosis. *Journal of the International Neuropsychological Society*, *16*(2), 287-96.

Parmenter, B. A., Testa, S. M., Schretlen, D. J., Weinstock-Guttman, B., & Benedict, R. H. B. (2010). The utility of regression-based norms in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *Journal of the International Neuropsychological Society*, *16*(1), 6-16.

Patwardhan, M. B., Matchar, D. B., Samsa, G. P., McCrory, D. C., Williams, R. G., & Li, T. T. (2005). Cost of multiple sclerosis by level of disability: a review of literature. *Multiple Sclerosis Journal*, *11*(2), 232-9.

Peterson, J. W., Bö, L., Mörk, S., Chang, A., & Trapp, B. D. (2001). Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions. *Annals of Neurology*, *50*(3), 389-400.

Petzold, A., Steenwijk, M. D., Eikelenboom, J. M., Wattjes, M. P., & Uitdehaag, B. M. (2016). Elevated CSF neurofilament proteins predict brain atrophy: a 15-year follow-up study. *Multiple Sclerosis Journal*, *22*(9), 1154-62.

Pham, D. L., & Prince, J. L. (1999). Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Transactions on Medical Imaging*, *18*(9), 737-52.

Pham, D. L., Xu, C., & Prince, J. L. (2000). Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, *2*(1), 315-37.

Phillips, L. H., Henry, J. D., Scott, C., Summers, F., Whyte, M., & Cook, M. (2011). Specific impairments of emotion perception in multiple sclerosis. *Neuropsychology*, *25*(1), 131-6.

Phillips, L. H., Henry, J. D., Nouzova, E., Cooper, C., Radlak, B., & Summers, F. (2014). Difficulties with emotion regulation in multiple sclerosis: Links to executive function, mood, and quality of life. *Journal of Clinical and Experimental Neuropsychology*, *36*(8), 831-42.

Pietroboni, A. M., Caprioli, M., Carandini, T., Scarioni, M., Ghezzi, L., Arighi, A., et al. (2019). CSF β-amyloid predicts prognosis in patients with multiple sclerosis. *Multiple Sclerosis Journal*, *25*(9), 1223-31.

Pitteri, M., Romualdi, C., Magliozzi, R., Monaco, S., & Calabrese, M. (2017). Cognitive impairment predicts disability progression and cortical thinning in MS: An 8-year study. *Multiple Sclerosis Journal*, *23*(6), 848-54.

Pitteri, M., Chen, P., Passarini, L., Albanese, S., Meneghello, F., & Barrett, A. M. (2018). Conventional and functional assessment of spatial neglect: Clinical practice suggestions. *Neuropsychology*, *32*(7), 835-42.

Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., et al. (2011). Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology*, *69*(2), 292-302.

Pöttgen, J., Dziobek, I., Reh, S., Heesen, C., & Gold, S. M. (2013). Impaired social cognition in multiple sclerosis. *Journal of Neurology, Neurosurgery and Psychiatry*, *84*(5), 523-28.

Pozzilli, C., Romano, S., & Cannoni, S. (2002). Epidemiology and current treatment of multiple sclerosis in Europe today. *Journal of Rehabilitation Research and Development*, *39*(2), 175-86.

Preziosa, P., Rocca, M. A., Pagani, E., Stromillo, M. L., Enzinger, C., Gallo, A., et al. (2016). Structural MRI correlates of cognitive impairment in patients with multiple sclerosis: a multicenter study. *Human Brain Mapping*, *37*(4), 1627-44.

Quintana, E., Coll, C., Salavedra-Pont, J., Muñoz-San Martín, M., Robles-Cedeño, R., Tomàs-Roig, J., et al. (2018). Cognitive impairment in early stages of multiple sclerosis is associated with high cerebrospinal fluid levels of chitinase 3-like 1 and neurofilament light chain. *European Journal of Neurology*, *25*(9), 1189-91.

Raimo, S., Trojano, L., Pappacena, S., Alaia, R., Spitaleri, D., Grossi, D., et al. (2017). Neuropsychological correlates of theory of mind deficits in patients with multiple sclerosis. *Neuropsychology*, *31*(7), 811-21.

Rainey-Barger, E. K., Rumble, J. M., Lalor, S. J., Esen, N., Segal, B. M., & Irani, D. N. (2011). The lymphoid chemokine, CXCL13, is dispensable for the initial recruitment of B cells to the acutely inflamed central nervous system. *Brain, Behavior, and Immunity*, *25*(5), 922-31.

Rao, S. M. (1990). A manual for the brief repeatable battery of neuropsychological tests in multiple sclerosis. *Milwaukee: Medical College of Wisconsin*.

Rao, S. M., Grafman, J., DiGiulio, D., Mittenberg, W., Bernardin, L., Leo, G. J., et al. (1993). Memory dysfunction in multiple sclerosis: Its relation to working memory, semantic encoding, and implicit learning. *Neuropsychology*, *7*(3), 364-74.

Rao, S. M. (2004). Cognitive function in patients with multiple sclerosis: impairment and treatment. *International Journal of MS Care*, *6*(1), 9-22.

Regan, D. M., Silver, R., & Murray, T. J. (1977). Visual acuity and contrast sensitivity in multiple sclerosis--hidden visual loss: an auxiliary diagnostic test. *Brain*, *100*(3), 563-79.

Reuter, M., & Fischl, B. (2011). Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage*, *57*(1), 19-21.

Riccitelli, G., Rocca, M. A., Pagani, E., Rodegher, M. E., Rossi, P., Falini, A., et al. (2011). Cognitive impairment in multiple sclerosis is associated to different patterns of gray matter atrophy according to clinical phenotype. *Human Brain Mapping*, *32*(10), 1535-43.

Rimkus, C. M., Schoonheim, M. M., Steenwijk, M. D., Vrenken, H., Eijlers, A. J., Killestein, J., et al. (2019). Gray matter networks and cognitive impairment in multiple sclerosis. *Multiple Sclerosis Journal*, *25*(3), 382-91.

Roca, C. A., Su, T. P., Elpern, S., McFarland, H., Rubinow, D. R. (1999). Cerebrospinal fluid somatostatin, mood, and cognition in multiple sclerosis. *Biological Psychiatry, 46*(4), 551-6.

Rocca, M. A., Valsasina, P., Absinta, M., Riccitelli, G., Rodegher, M. E., Misci, P., et al. (2010). Default-mode network dysfunction and cognitive impairment in progressive MS. *Neurology*, *74*(16), 1252-9.

Rocca, M. A., Amato, M. P., De Stefano, N., Enzinger, C., Geurts, J. J., Penner, I. K., et al. (2015). Clinical and imaging assessment of cognitive dysfunction in multiple sclerosis. *The Lancet Neurology*, *14*(3), 302-17.

Rocca, M. A., & Filippi, M. (2017). Functional reorganization is a maladaptive response to injury–YES. *Multiple Sclerosis Journal*, *23*(2), 191-3.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of experimental psychology: Learning, Memory, and Cognition*, *21*(4), 803-14.

Roediger, H. L., & McDermott, K. B. (2000). Distortions of memory. *The Oxford Handbook of Memory*, 149-62.

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories.

Rojas, J. I., Patrucco, L., Miguez, J., & Cristiano, E. (2016). Brain atrophy in multiple sclerosis: therapeutic, cognitive and clinical impact. *Arquivos de Neuro-psiquiatria*, *74*(3), 235-43.

Rovaris, M., Inglese, M., Van Schijndel, R. A., Sormani, M. P., Rodegher, M., Comi, G et al. (2000). Sensitivity and reproducibility of volume change measurements of different brain portions on magnetic resonance imaging in patients with multiple sclerosis. *Journal of Neurology*, *247*(12), 960-5.

Roy, S., Schwartz, C. E., Duberstein, P., Dwyer, M. G., Zivadinov, R., Bergsland, N., et al. (2016). Synergistic effects of reserve and adaptive personality in multiple sclerosis. *Journal of the International Neuropsychological Society*, *22*(9), 920-7.

Roy, S., Drake, A., Fuchs, T., Dwyer, M. G., Zivadinov, R., Chapman, B. P., et al. (2018). Longitudinal personality change associated with cognitive decline in multiple sclerosis. *Multiple Sclerosis Journal*, *24*(14), 1909-12.

Ruggieri, S., Pontecorvo, S., Tortorella, C., & Gasperini, C. (2018). Induction treatment strategy in multiple sclerosis: a review of past experiences and future perspectives. *Multiple Sclerosis and Demyelinating Disorders*, *3*(1), 5.

Sailer, M., Fischl, B., Salat, D., Tempelmann, C., Schönfeld, M. A., Busa, E., et al. (2003). Focal thinning of the cerebral cortex in multiple sclerosis. *Brain, 126*(8), 1734-44.

Sandry J, Akbar N, Zuppichini M, DeLuca J. Cognitive rehabilitation in multiple sclerosis. *Research Progress in Alzheimer's Disease and Dementia*. *2016*(6), 195-233.

Santangelo, G., Siciliano, M., Pedone, R., Vitale, C., Falco, F., Bisogno, R., et al. (2015). Normative data for the Montreal Cognitive Assessment in an Italian population sample. *Neurological Sciences*, *36*(4), 585-91.

Savini, G., Pardini, M., Castellazzi, G., Lascialfari, A., Chard, D., D'Angelo, E., et al. (2019). Default mode network structural integrity and cerebellar connectivity predict information processing speed deficit in multiple sclerosis. *Frontiers in Cellular Neuroscience*, *13*, 21.

Sawcer, S., Hellenthal, G., Pirinen, M., Spencer, C. C., Patsopoulos, N. A., Moutsianas, L., et al. (2011). Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature, 476*(7359), 214-9.

Scalfari, A., Romualdi, C., Nicholas, R. S., Mattoscio, M., Magliozzi, R., Morra, A., et al. (2018). The cortical damage, early relapses, and onset of the progressive phase in multiple sclerosis. *Neurology*, *90*(24), e2107-18.

Schirmer, L., Velmeshev, D., Holmqvist, S., Kaufmann, M., Werneburg, S., Jung, D., et al. (2019). Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature*, *573*(7772), 75-82.

Schoonheim, M. M., Geurts, J. J., & Barkhof, F. (2010). The limits of functional reorganization in multiple sclerosis. *Neurology*, *74*(16), 1246-7.

Schoonheim, M. M., Meijer, K. A., & Geurts, J. J. (2015). Network collapse and cognitive impairment in multiple sclerosis. *Frontiers in Neurology*, *6*, 82.

Schoonheim, M. M. (2017). Functional reorganization is a maladaptive response to injury–Commentary. *Multiple Sclerosis Journal*, *23*(2), 194-6.

Sellebjerg, F., Royen, L., Soelberg Sørensen, P., Oturai, A. B., & Jensen, P. E. H. (2018). Prognostic value of cerebrospinal fluid neurofilament light chain and chitinase-3-like-1 in newly diagnosed patients with multiple sclerosis. *Multiple Sclerosis Journal*, 25(11), 1444-51.

Sharief, M. K., & Hentges, R. (1991). Association between tumor necrosis factor-α and disease progression in patients with multiple sclerosis. *New England Journal of Medicine*, *325*(7), 467-72.

Serafini, B., Rosicarelli, B., Magliozzi, R., Stigliano, E., & Aloisi, F. (2004). Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis. *Brain Pathology*, *14*(2), 164-74.

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

Smith, S. M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P. M., Federico, A., et al. (2002). Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage*, *17*(1), 479-89.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23*(1), S208-19.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., et al. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*, *106*(31), 13040-5.

Sohlberg, M. M., & Mateer, C. A. (2001). Improving attention and managing attentional problems: Adapting rehabilitation techniques to adults with ADD. *Annals of the New York Academy of Sciences*, *931*(1), 359-75.

Sohlberg, M. M., & Mateer, C. A. (2017). *Cognitive rehabilitation: An integrative neuropsychological approach*. Guilford Publications.

Sospedra, M., & Martin, R. (2005). Immunology of multiple sclerosis. *Annual Review of Immunology*, *23*, 683-747.

Stilund, M., Reuschlein, A. K., Christensen, T., Møller, H. J., Rasmussen, P. V., & Petersen, T. (2014). Soluble CD163 as a marker of macrophage activity in newly diagnosed patients with multiple sclerosis. *PloS One*, *9*(6), e98588.

Stilund, M., Gjelstrup, M. C., Petersen, T., Møller, H. J., Rasmussen, P. V., & Christensen, T. (2015). Biomarkers of inflammation and axonal degeneration/damage in patients with newly diagnosed multiple sclerosis: contributions of the soluble CD163 CSF/serum ratio to a biomarker panel. *PloS One*, *10*(4), e0119681.

Strober, L. B., Christodoulou, C., Benedict, R. H. B., Westervelt, H. J., Melville, P., Scherl, W. F., et al. (2012). Unemployment in multiple sclerosis: the contribution of personality and disease. *Multiple Sclerosis Journal*, *18*(5), 647-53.

Strober, L., DeLuca, J., Benedict, R. H. B., Jacobs, A., Cohen, J. A., Chiaravalloti, N., et al. (2018). Symbol Digit Modalities Test: A valid clinical trial endpoint for measuring cognition in multiple sclerosis. *Multiple Sclerosis Journal, 25*(13), 1781-90.

Studer-Luethi, B., Jaeggi, S. M., Buschkuehl, M., & Perrig, W. J. (2012). Influence of neuroticism and conscientiousness on working memory training outcome. *Personality and Individual Differences*, *53*(1), 44-9.

Sumowski, J. F., Benedict, R. H. B., Enzinger, C., Filippi, M., Geurts, J. J., Hamalainen, P., et al. (2018). Cognition in multiple sclerosis: State of the field and priorities for the future. *Neurology*, *90*(6), 278-88.

Teunissen, C. E., Petzold, A., Bennett, J. L., Berven, F. S., Brundin, L., Comabella, M., et al. (2009). A consensus protocol for the standardization of cerebrospinal fluid collection and biobanking. *Neurology*, *73*(22), 1914-22.

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., et al. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology*, *17*(2), 162-73.

Thornton, A. E., Raz, N., & Tucker, K. A. (2002). Memory in multiple sclerosis: contextual encoding deficits. *Journal of the International Neuropsychological Society*, *8*(3), 395-409.

Tillema, J. M., Hulst, H. E., Rocca, M. A., Vrenken, H., Steenwijk, M. D., Damjanovic, D., et al. (2016). Regional cortical thinning in multiple sclerosis and its relation with cognitive impairment: a multicenter study. *Multiple Sclerosis Journal, 22*(7), 901-9.

Van Geest, Q., Douw, L., van't Klooster, S., Leurs, C. E., Genova, H. M., Wylie, G. R., et al. (2018). Information processing speed in multiple sclerosis: Relevance of default mode network dynamics. *NeuroImage: Clinical*, *19*, 507-15.

Vellante, M., Baron-Cohen, S., Melis, M., Marrone, M., Petretto, D. R., Masala, C., et al. (2013). The “Reading the Mind in the Eyes” test: systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry*, *18*(4), 326-54.

Weber, E., Goverover, Y., & DeLuca, J. (2019). Beyond cognitive dysfunction: Relevance of ecological validity of neuropsychological tests in multiple sclerosis. *Multiple Sclerosis Journal*, *25*(10), 1412-9.

Wegner, C., Esiri, M. M., Chance, S. A., Palace, J., & Matthews, P. M. (2006). Neocortical neuronal, synaptic, and glial loss in multiple sclerosis. *Neurology*, *67*(6), 960-7.

Weinshenker, B. G., Issa, M., & Baskerville, J. (1996). Meta-analysis of the placebo-treated groups in clinical trials of progressive MS. *Neurology*, *46*(6), 1613-9.

Weinstock-Guttman, B. (2013). An update on new and emerging therapies for relapsing-remitting multiple sclerosis. *The American Journal of Managed Care*, *19*(17), S343-54.

Weinstock-Guttman, B., Eckert, S., & Benedict, R. H. B. (2018). A decline in cognitive function should lead to a change in disease-modifying therapy - Yes. *Multiple Sclerosis Journal, 24*(13), 1681-2 .

Wilson, B. A., & Betteridge, S. (2019). *Essentials of Neuropsychological Rehabilitation*. Guilford Press.

Zalesky, A., Fornito, A., & Bullmore, E. T. (2010). Network-based statistic: identifying differences in brain networks. *Neuroimage*, *53*(4), 1197-207.

Zipoli, V., Goretti, B., Hakiki, B., Siracusa, G., Sorbi, S., Portaccio, E., et al. (2010). Cognitive impairment predicts conversion to multiple sclerosis in clinically isolated syndromes. *Multiple Sclerosis Journal*, *16*(1), 62-7.

Zivadinov, R., Heininen-Brown, M., Schirda, C. V., Poloni, G. U., Bergsland, N., Magnano, C. R., et al. (2012). Abnormal subcortical deep-gray matter susceptibility-weighted imaging filtered phase measurements in patients with multiple sclerosis: a case-control study. *Neuroimage*, *59*(1), 331-9.

Zivadinov, R., Tekwe, C., Bergsland, N., Dolezal, O., Havrdova, E., Krasensky, J., et al. (2013). Bimonthly evolution of cortical atrophy in early relapsing-remitting

multiple sclerosis over 2 years: a longitudinal study. *Multiple Sclerosis International*, *2013*, 231345.

Zivadinov, R., Ramasamy, D. P., Benedict, R. H. B., Polak, P., Hagemeier, J., Magnano, C., et al. (2016a). Cerebral microbleeds in multiple sclerosis evaluated on susceptibility-weighted images and quantitative susceptibility maps: a case-control study. *Radiology*, *281*(3), 884-95.

Zivadinov, R., Uher, T., Hagemeier, J., Vaneckova, M., Ramasamy, D. P., Tyblova, M., et al. (2016b). A serial 10-year follow-up study of brain atrophy and disability progression in RRMS patients. *Multiple Sclerosis Journal*, *22*(13), 1709-18.