Predicting Long-term Disability in Multiple Sclerosis:

A Narrative Review of Current Evidence and Future Directions

Bianca Weinstock-Guttman, MD; Maria Pia Sormani, PhD; Pavle Repovic, MD, PhD

From the Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, Buffalo, NY, USA (BW-G); Department of Health Sciences, University of Genova, Genoa, Italy (MPS); IRCCS Ospedale Policlinico San Martino, Genoa, Italy (MPS); and Swedish Medical Center at Seattle, Seattle, WA, USA (PR). Correspondence: Bianca Weinstock-Guttman, MD, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, State University of New York at Buffalo, 1010 Main St, Buffalo, NY 14202, USA; e-mail: bw8@buffalo.edu.

Running head: Predicting Long-term Disability in MS

DOI: 10.7224/1537-2073.2020-114

© 2021 Consortium of Multiple Sclerosis Centers.
Practice Points

- Although the Expanded Disability Status Scale (EDSS) is a widely accepted prognostic tool for MS, it has been criticized for its poor reliability and limited sensitivity to changes in certain aspects of disease progression.

- Using prognostic factors beyond the EDSS may improve patient care and inform future trials in MS. Specifically, supplementing the EDSS with patient-reported physical measures, assessing fatigue, and using simultaneous cognitive and motor function testing may help in gauging MS disease progression.
Abstract

The ability to reliably monitor disease progression in patients with multiple sclerosis (MS) is integral to patient care. The Expanded Disability Status Scale (EDSS) is a commonly used tool to assess the disability status of patients with MS; however, it has limited sensitivity in detecting subtle changes in disability levels and, as a result, does not consistently provide clinicians with accurate insight into disease progression. At the 2019 European Committee for Treatment and Research in Multiple Sclerosis meeting in Stockholm, Sweden, a panel of neurologists met to discuss the limitations of the EDSS as a short-term predictor of MS progression. Prior to this panel discussion, a targeted literature review was conducted to evaluate published evidence on prognostic measures such as fatigue, physical assessments, and measures that are more taxing for patients, all of which may be useful to clinicians at different stages of the course of MS. This article summarizes currently available evidence in support of these measures. In addition, this article highlights the current state of expert clinical consensus regarding the current approaches used to predict and monitor disease progression and offers insight for future studies to assist clinicians in accurately monitoring disease progression in patients with MS. *Int J MS Care.*
Multiple sclerosis (MS) is a chronic neurologic disease wherein overactive immune cells cause inflammation, demyelination, and axonal damage in the central nervous system. The resulting degenerative disease course requires accurate, sensitive, and comprehensive measures for the purposes of monitoring progression and making decisions for patient care.

The Expanded Disability Status Scale (EDSS), an ordinal, clinician-administered assessment scale with scores ranging from 0 to 10, is widely used to assess the status of patient disability in MS. Short-term sustained changes in disability status are often reported in terms of confirmed disability progression (CDP), the presence of which corresponds to an increase in EDSS score maintained on repeat evaluation at 3-6 months. However, several studies have criticized the EDSS and CDP for limitations related their prognostic value, including their lack of accuracy, limited sensitivity to change at certain disability levels of the disease, focus on physical ability, and nonstandardized interpretability.\(^1\)\(^-\)\(^3\) As a binary indicator, CDP in particular may offer limited value for prediction of disability progression. Although the EDSS does have some predictive value, certain aspects such as intrarater and interrater variability leave room for improvement.\(^4\)\(^,\)\(^5\) Given these limitations, there is need for an expansion in the tools used to gauge disease progression in clinical practice as well as clinical trials evaluating the efficacy of new MS treatments.

In this narrative review, we first summarize the available evidence of the predictive power of a select number of commonly used measures to assess MS progression. The goal of this review is not a systematic assessment but rather a targeted review of recent literature that focuses
on specific end points related to MS progression. We then discuss available evidence on key limitations of the EDSS as a prognostic measure and highlight some measures that may assist in improving the sensitivity of this scale in the clinical management of MS. We highlight general viewpoints from a panel of four experts in MS research, which included the authors (B.W.-G., M.P.S., P.R.) and Dr Enrique Alvarez of the University of Colorado, that convened at the 2019 European Committee for Treatment and Research in Multiple Sclerosis meeting in Stockholm, Sweden, held September 13-19, 2019. At this meeting, the panel discussed the utility of clinical methods currently used to predict and monitor MS progression and suggested general directions for future studies that may help clinicians effectively monitor MS disease worsening.

**Targeted Literature Review of Prognostic Measures**

A targeted literature review using ProQuest and Google Scholar was conducted prior to the expert panel discussion to identify evidence on prognostic measures other than the EDSS that may be predictive of long-term MS disease progression. The targeted literature review largely focused on articles published between 2014 and 2019, although articles published prior to 2014 were included to capture relevant evidence. Articles not published in English were excluded from consideration. All identified studies were screened, reviewed, and synthesized based on the following measures: no evidence of disease activity (NEDA); EDSS and CDP; brain volume loss; gray matter; white matter; lesions; serum neurofilament light; cognitive impairment; pregnancy and demographics, specifically age, sex, and education; relapse and annualized
relapse rate; disease duration; macular volume; and Montgomery-Åsberg Depression Rating Scale. A summary of key search terms associated with each measure is provided in Table 1.

The impact of each assessment used to assess the progression of MS and the patient populations evaluated are summarized in Table S1, which is published in the online version of this article at ijmsc.org. For each measure assessed, the evidence varied regarding its utility in assessing disease progression. Among the studies that met the inclusion criteria, the evidence regarding the predictive power of NEDA and minimal evidence of disease activity was generally inconsistent.\(^5\)\(^-\)\(^9\) Brain volume loss generally showed positive predictive power on long-term MS progression, particularly regarding increases in EDSS score\(^10\) and assessments that combine brain atrophy and retinal thinning.\(^11\) Although assessments of gray matter reductions were also predictive of MS progression, the evidence was partially based on associations and correlations.\(^12\),\(^13\) Similarly, there was some evidence supporting the predictive power of white matter on long-term disease progression, however it was based on associations.\(^12\) Limited evidence was available in favor of the predictive power of other measures of brain volume such as cerebrospinal fluid, parenchyma,\(^12\) and ventricular fractions,\(^14\) as well as macular volume.\(^15\) Cognitive impairment,\(^16\) demographic characteristics (age, sex, education),\(^12,14,16\) and measures assessing lesion volume,\(^10,14,17\) serum neurofilament light,\(^18\) and disease duration\(^12,16\) did not consistently predict disease progression, particularly in the long term. Although the evidence was mixed regarding the predictive power of annualized relapse rate,\(^16\) there was some evidence to support the predictive power of relapse on long-term disease progression.\(^5\) There was some evidence to support the predictive power of pregnancy on long-term disease progression; however, the research was largely based on congress presentations.\(^19\)
Depression Rating Scale was found to have no predictive power in long-term assessments of MS progression.\textsuperscript{14}

\textbf{The EDSS}

\textbf{Limitations of EDSS as a Prognostic Measure}

There is ample body of evidence supporting the limitations of the EDSS as a prognostic measure. Foremost, the EDSS is meant to function as a measure of irreversible disability in MS; however, the literature shows that it fails to serve this basic purpose.\textsuperscript{1-3}

Recently, the placebo arms of 31 randomized controlled trials in relapsing-remitting MS and secondary progressive MS patient populations were analyzed, and the results showed significant rates of EDSS improvement, sometimes as high as rates of EDSS worsening.\textsuperscript{1} To further illustrate the inaccuracy and instability of this instrument, the EDSS has been shown to overestimate the accumulation of permanent disability by up to 30\%.\textsuperscript{3} The EDSS also has plateau scores at 6.0 and 6.5, which over time has discouraged researchers from including patients with these scores in clinical trials, given the need to quantify change in progression in these trials.

\textbf{Inaccuracies Associated with EDSS Use}
Further evidence shows that increases in EDSS scores do not accurately identify patients with irreversible, long-term disease progression. A study conducted in persons with relapsing-remitting MS assessed multiple definitions of sustained progression using the EDSS. Between 15.8% and 42.2% of these individuals had sustained progression over 3.7 years, but nearly 50% of them did not maintain progression for the duration of follow-up. Relapses or changes in provider could not explain the poor performance of the EDSS, suggesting that the use of the EDSS and CDP as outcomes for clinical trials or observational studies could lead to incorrect conclusions due to the potential instability of EDSS scores.

The EDSS has also been criticized for not being a comprehensive measure of all dimensions of MS. While lower scale values are influenced by impairments detected by a neurologic examination, values above 4 are mainly based on walking ability, and values above 6 are based on patient handicaps. The EDSS does not adequately capture the dimensions of cognition, upper extremity function, or fatigue, which are believed to be relevant predictors of long-term disease progression in MS. Moreover, studies recommend separate consideration of the lower and upper value ranges of the EDSS because EDSS scores of 6 and above are less sensitive to change in disease severity. The evidence on the prognostic limitations of the EDSS across several domains highlights the scope for improved prediction of long-term disease progression in MS.

Panel’s Viewpoints on Alternatives to EDSS as Prognostic Measures
Because of the potential prognostic limitations of the EDSS, it is the panel’s opinion that alternative measures are needed to better classify disease severity in patients with MS and to assess disease progression. Although there is a large and growing body of evidence on the prognostic value of factors beyond the EDSS, there is no widely-held understanding of whether these factors improve predictions of long-term disability beyond the EDSS. To detect clinically meaningful changes in patients with MS, prognostic measures will need increased reliability and sensitivity. To these ends, current research suggests that approaches incorporating fatigue and lower limb function, as well as combination measures, either as a supplement or alternative to the EDSS, may be promising avenues to explore. In the following sections, we provide literature to support the panel’s viewpoints on the alternatives to the EDSS as prognostic measures.

Incorporation of Fatigue and Lower Limb Function

The related matters of fatigue and lower extremity dysfunction, including patient-reported leg weakness as well as functional limitations identified via examinations, may be of particular interest among alternative measures of disease progression with promising predictive value. A preliminary study found that in older people with MS, fatigue and limited lower extremity function predicted conversion from relapsing-remitting MS to secondary progressive MS within 5 years. The study evaluated 155 persons aged 50 years or older with relapsing-remitting MS and a disease duration of at least 15 years. Fatigue was reported to be four times as likely in people with disease progression (92%) compared to those who did not progress (68%). Additionally, limitations in lower extremity function were reported to be three times as likely in
people with disease progression at the beginning of the study (53%) compared with those who did not progress (22%). Researchers noted that although the precise cause of fatigue in patients with MS has not been fully established, it is most likely a result of the underlying complex inflammatory and neurodegenerative processes that characterize the disease. These included sequelae from MS—demyelination, axonal injury, and inflammatory response—in addition to other factors such as depression and sleep disturbances. Consequently, fatigue and lower extremity dysfunction may be more sensitive indicators of the extent of central nervous system injury, but they may not be fully assessed or gauged with high sensitivity during a short neurologic examination.

Use of Combination Measures to Supplement EDSS

An alternative approach to improved prediction of disease progression could involve supplementing the EDSS with physical measures such as the Nine-Hole Peg Test (NHPT), Timed 25-Foot Walk test, or low-contrast letter acuity test. In the example of the NHPT, people with worsening MS frequently report impairment in manual dexterity. The NHPT is commonly used in MS research and clinical practice and is considered to be a gold standard measure of manual dexterity. Studies have indicated consistently high interrater and test-retest reliability of the NHPT as well as its ability to distinguish patients with MS who have different levels of upper limb impairment. A 20% change in NHPT test score is used to define clinically meaningful worsening; however, this definition needs further validation in all stages of the disease.
Another potential avenue for future research is combining cognitive measures such as the Symbol Digit Modalities Test (SDMT) or the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS) with taxing motor tasks and walking tests.\(^{31,32}\) A 10-year retrospective longitudinal study reported cognitive impairment, measured by SDMT, as a predictor of disability progression and secondary progressive MS conversion in newly diagnosed relapsing-remitting MS.\(^{16}\) The study included 155 persons with relapsing-remitting MS, of whom 67 (43.2\%) reached an EDSS score of 4.0, and 34 (21.9\%) converted to secondary progressive MS during the follow-up period.\(^{16}\) Individuals with cognitive impairment at MS diagnosis were three times as likely to reach an EDSS score of 4.0 and two times as likely to convert to secondary progressive MS compared with cognitively preserved individuals.\(^{16}\)

Lastly, the predictive validity of magnetic resonance imaging and clinical scoring assessments may help clinicians identify treatment failure and potentially assist with treatment optimization.\(^{33}\) For example, the Rio scoring system, which combines evidence of magnetic resonance imaging lesions, presence of relapse, and increases in EDSS scores within 12 months of treatment initiation, has been used to predict ongoing disease activity and, ultimately, which patients are at risk for a suboptimal response to therapy over time.\(^{33}\) The sum of each parameter (0-3) distinguishes patients’ risk from low to high.\(^{33}\) The modified Rio score, a simplified version of the Rio score that excludes the EDSS and modifies relapse and magnetic resonance imaging lesion criteria, has been reported to have a high positive predictive value of disease progression within 3 years of treatment initiation.\(^{33}\) Results from a longitudinal study validating the utility of the scoring system found that in 222 patients with relapsing-remitting MS treated with interferon,
patients with a risk score of 0 had a 24% probability of MS progression whereas those with a score of 2 or higher had their risk of progression increase to 65%.\textsuperscript{33}

**Direction for Future Studies**

Although our current understanding of MS disease progression does not allow for a single most promising measure to be identified, research should continue to explore promising themes and measures to improve clinical and real-world assessments of MS treatments in the near term.\textsuperscript{34} An improved understanding of individuals at higher risk of disease progression may eventually result in more tailored treatment options for patients and establish an advanced approach to economic evaluations in MS.

**Financial Disclosures:** Dr Weinstock-Guttman has received grant/research support from, and participated in speakers' bureaus and/or served as a consultant for, Biogen, EMD Serono, Novartis, Genentech, Celgene, AbbVie, Sanofi Genzyme, and Mallinckrodt and serves on the editorial board for *BMJ Neurology, Children, CNS Drugs, MS International,* and *Frontiers Epidemiology.* Dr Sormani has received consulting fees from Biogen, Merck, Novartis, Roche, Sanofi, Celgene, MedDay, GeNeuro, and Mylan. Dr Repovic declares no conflicts of interest.

**Funding/Support:** Financial support in the preparation of this review was provided by Celgene.
This in-press manuscript has been peer reviewed and accepted for publication by the International Journal of MS Care and appears here in nearly final form. It has been edited and received author approval. Essential corrections may still be made later in the proof stage, before publication in a print issue. Once published in an issue, the paper will be removed from the Online First section and appear in that issue's table of contents. Meanwhile, the manuscript is citable using the DOI, which appears on the first page.

References

1. Ebers GC, Heigenhauser L, Daumer M, Lederer C, Noseworthy JH. Disability as an outcome in MS clinical trials. Neurology. 2008;71:624-631.

2. Healy BC, Engler D, Glanz B, Musallam A, Chitnis T. Assessment of definitions of sustained disease progression in relapsing-remitting multiple sclerosis. Mult Scler Int. 2013;2013:189624.

3. Kalincik T, Cutter G, Spelman T, et al. Defining reliable disability outcomes in multiple sclerosis. Brain. 2015;138(pt 11):3287-3298.

4. Lavery AM, Verhey LH, Waldman AT. Outcome measures in relapsing-remitting multiple sclerosis: capturing disability and disease progression in clinical trials. Mult Scler Int. 2014;2014:262350.

5. Rio J, Rovira A, Tintore M, et al. Disability progression markers over 6-12 years in interferon-β-treated multiple sclerosis patients. Mult Scler. 2018;24:322-330.

6. Prosperini L, Mancinelli CR, Sgarlata E, et al. Minimal or no evidence of disease activity: which target to prevent long-term disability in multiple sclerosis? Poster (769) presented at: Seventh Joint ECTRIMS-ACTRIMS Meeting; October 25-18, 2017; Paris, France. ECTRIMS Online Library. 2017.

7. Uher T, Havrdova E, Sobisek L, et al. Is no evidence of disease activity an achievable goal in MS patients on intramuscular interferon beta-1a treatment over long-term follow-up? Mult Scler. 2017;23:242-252.
This in-press manuscript has been peer reviewed and accepted for publication by the International Journal of MS Care and appears here in nearly final form. It has been edited and received author approval. Essential corrections may still be made later in the proof stage, before publication in a print issue. Once published in an issue, the paper will be removed from the Online First section and appear in that issue’s table of contents. Meanwhile, the manuscript is citable using the DOI, which appears on the first page.

8. Kappos L, Stefano N, Sormani M, et al. NEDA as a predictor of conversion to secondary progressive multiple sclerosis in patients with relapsing-remitting multiple sclerosis. Poster (P1217) presented at: 32nd Congress of ECTRIMS; September 14-17, 2016; London, UK. ECTRIMS Online Library. 2016

9. Goodin D, Reder A, Koelbach R, et al. Contribution of clinical and MRI to evaluation of NEDA (no evidence of disease activity) from the pivotal trial of interferon beta-1b in multiple sclerosis. Poster (1592) presented at: 32nd Congress of ECTRIMS; September 14-17, 2016; London, UK. ECTRIMS Online Library. 2016

10. Horakova D, Dwyer MG, Havrdova E, et al. Gray matter atrophy and disability progression in patients with early relapsing-remitting multiple sclerosis: a 5-year longitudinal study. J Neurol Sci. 2009;282:112-119.

11. Pulido-Valdeolivas I, Andorra M, Gomez-Andres D, et al. Retinal and brain damage during multiple sclerosis course: inflammatory activity is a key factor in the first 5 years. Sci Rep. 2020;10:13333.

12. Lavorgna L, Bonavita S, Ippolito D, et al. Clinical and magnetic resonance imaging predictors of disease progression in multiple sclerosis: a nine-year follow-up study. Mult Scler. 2014;20:220-226.

13. Uher T, Vaneckova M, Krasensky J, et al. Pathological cut-offs of global and regional brain volume loss in multiple sclerosis. Mult Scler. 2019;25:541-553.

14. Deloire MS, Ruet A, Hamel D, Bonnet M, Doussset V, Brochet B. MRI predictors of cognitive outcome in early multiple sclerosis. Neurology. 2011;76:1161-1167.
15. Rothman A, Murphy OC, Fitzgerald KC, et al. Retinal measurements predict 10-year disability in multiple sclerosis. *Ann Clin Transl Neurol*. 2019;6:222-232.

16. Moccia M, Lanzillo R, Palladino R, et al. Cognitive impairment at diagnosis predicts 10-year multiple sclerosis progression. *Mult Scler*. 2016;22:659-667.

17. Popescu V, Agosta F, Hulst HE, et al. Brain atrophy and lesion load predict long term disability in multiple sclerosis. *J Neurol Neurosurg Psychiatry*. 2013;84:1082-1091.

18. Chitnis T, Gonzalez C, Healy BC, et al. Neurofilament light chain serum levels correlate with 10-year MRI outcomes in multiple sclerosis. *Ann Clin Transl Neurol*. 2018;5:1478-1491.

19. Jokubaitis V, Spelman T, Kalincik T, et al. 10-year progression predictors after first treatment initiation in a relapsing-remitting MS cohort (poster 281). *Mult Scler*. 2015;21(suppl):76-653. 31st Congress of ECTRIMS.

20. Cinar BP, Yorgun YG. What we learned from the history of multiple sclerosis measurement: Expanded Disability Status Scale. *Noro Psikiyatr Ars*. 2018;55(suppl 1):S69-S75.

21. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology*. 1983;33:1444-1452.

22. Vaughn C, Kavak K, Bushra A, et al. Self-reported fatigue and lower limb problems predictive of conversion to secondary progressive multiple sclerosis (SPMS) in an aging sample of MS patients. Presented at 69th AAN Annual Meeting; April 22-28, 2017; Boston.
23. Feys P, Lamers I, Francis G, et al. The Nine-Hole Peg Test as a manual dexterity performance measure for multiple sclerosis. *Mult Scler*. 2017;23:711-720.

24. Meyer-Moock S, Feng Y-S, Maeurer M, Dippel F-W, Kohlmann T. Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. *BMC Neurol.* 2014;14:58.

25. Wang G, Cutter GR, Cofield SS, et al. Baseline EDSS proportions in MS clinical trials affect the overall outcome and power: a cautionary note. *Mult Scler*. 2017;23:982-987.

26. Vaughn CB, Kavak KS, Dwyer MG, et al. Fatigue at enrollment predicts EDSS worsening in the New York State Multiple Sclerosis Consortium. *Mult Scler*. 2020;26:99-108.

27. Cavallari M, Palotai M, Glanz BI, et al. Fatigue predicts disease worsening in relapsing-remitting multiple sclerosis patients. *Mult Scler*. 2016;22:1841-1849.

28. Hulst HE, Thompson AJ, Geurts JJ. The measure tells the tale: clinical outcome measures in multiple sclerosis. Editorial. *Mult Scler*. 2017;23:626-627.

29. Motl RW, Cohen JA, Benedict R, et al. Validity of the timed 25-foot walk as an ambulatory performance outcome measure for multiple sclerosis. *Mult Scler*. 2017;23:704-710.

30. Balcer LJ, Raynowska J, Nolan R, et al. Validity of low-contrast letter acuity as a visual performance outcome measure for multiple sclerosis. *Mult Scler*. 2017;23:734-747.
31. Strober L, DeLuca J, Benedict RHB, et al. Symbol Digit Modalities Test: a valid clinical trial endpoint for measuring cognition in multiple sclerosis. *Mult Scler.* 2019;25:1781-1790.

32. Saccà F, Costabile T, Carotenuto A, et al. The EDSS integration with the Brief International Cognitive Assessment for Multiple Sclerosis and orientation tests. *Mult Scler.* 2017;23:1289-1296.

33. Sormani MP, Rio J, Tintore M, et al. Scoring treatment response in patients with relapsing multiple sclerosis. *Mult Scler.* 2013;19:605-612.

34. Tutuncu M, Altintas A, Dogan BV, et al. The use of Modified Rio score for determining treatment failure in patients with multiple sclerosis: retrospective descriptive case series study. *Acta Neurol Belg.* 2020. doi: 10.1007/s13760-020-01476-2. Online ahead of print.
Table 1. Summary of key search terms used in the targeted literature review

| Measure                      | Key search terms                                                                 |
|------------------------------|----------------------------------------------------------------------------------|
| ARR                          | Annual relapse rate, long term, disability progression, multiple sclerosis, MS, relapsing remitting, RRMS, RMS |
| Brain volume loss            | Brain volume loss, BVL, cognition, cognitive decline, cognitive impairment, physical impairment, physical decline, disease progression, disease worsening, long term effects, long term outcomes, long term disability, multiple sclerosis, MS |
| Cognitive impairment         | Cognitive impairment, symbol digit modalities test, SDMT, multiple sclerosis, MS, relapsing remitting, RRMS, RMS |
| Cognitive reserve            | Cognitive reserve, multiple sclerosis, MS, relapsing remitting, RRMS, RMS         |
| EDSS + CDP                   | CDP, disability progression, EDSS, expanded disability status scale, multiple sclerosis, MS, relapsing remitting, RRMS, RMS |
| Gray matter                  | Grey matter, gray matter, cognition, cognitive decline, cognitive impairment, physical impairment, physical decline, disease progression, disease worsening, long term effects, long term outcomes, long term disability, multiple sclerosis, MS |
| Lesions                      | T1, T2, GAD, lesion assessment, lesion load, annualized relapse rate, multiple sclerosis, MS |
| MS disability measures       | Multiple sclerosis, disability measures                                          |
| MSFC                         | Multiple Sclerosis Functional Composite, MSFC, long term, multiple sclerosis, MS, relapsing remitting, RRMS, RMS |
| NEDA                         | No evidence of disease activity, no evident disease activity, NEDA, long term, multiple sclerosis, MS, relapsing remitting, RRMS, RMS |
| Serum neurofilament light    | Serum neurofilament light, SNfL, multiple sclerosis, MS                           |

Abbreviations: ARR, annual relapse rate; BVL, brain volume loss; CDP, confirmed disability progression; EDSS, Expanded Disability Status Scale; GAD, gadolinium; MS, multiple sclerosis; Multiple Sclerosis Functional Composite; NEDA, no evidence of disease activity; RMS, relapsing MS; RRMS, relapsing-remitting MS; SDMT, Symbol Digit Modalities Test; SNfL, serum neurofilament light.


**Table S1. Summary of studies identified following targeted literature review**

| Study            | Number of patients evaluated | Measure | Assessment | Main Finding(s) |
|------------------|-----------------------------|---------|------------|-----------------|
| Horakova 2009¹⁰  | 181 patients with early RRMS | -Percentage of brain volume loss | -Predictive of increase in EDSS observed over 5-year period |
|                  |                             | -Lateral ventricle volume loss |                      |
| Uher 2019¹³      | 964 patients with RRMS     | -Annualized percent whole brain volume loss | -Correlation with annualized absolute increase in EDSS |
|                  |                             | -Annualized percent thalamic volume loss |                      |
|                  |                             | -Annualized percent corpus callosum volume loss |                      |
|                  |                             | -Annualized gray matter volume loss | -Correlation with higher annualized absolute EDSS |
| Lavorgna 2014¹²  | 241 patients with RRMS    | -Reduction in gray matter fraction¹ | Predictive of: |
|                  |                             |                      | -Higher probability of conversion from RRMS to SPMS |
|                  |                             |                      | -Higher probability of EDSS progression² |
|                  |                             |                      | -Achievement of EDSS score of 4 |
|                  |                             |                      | -Higher risk of reaching EDSS score of 4 |
|                  | Brain volume measures      | -Lower gray matter fraction | Association with: |
|                  |                             |                      | -Conversion from RRMS to SPMS |
|                  |                             |                      | -Achievement of EDSS score of 4 |
|                  |                             | -Lower white matter fraction | Association with: |
|                  |                             |                      | -Conversion from RRMS to SPMS |
|                  |                             |                      | -Achievement of EDSS score of 4 |
|                  |                             | -Higher cerebrospinal fluid fraction | Association with: |
|                  |                             |                      | -Conversion from RRMS to SPMS |
|                  |                             |                      | -Achievement of EDSS score of 4 |
| Deloire 2011¹⁴  | 44 patients with MS        | -Brain parenchymal fraction | No predictive evidence of: |
|                  |                             | -Ventricular fraction  | -Change in memory z score |
|                  |                             | -Normal-appearing brain tissue-magnetization transfer ratio | -Change in information processing speed |
| Author(s)            | Number of Patients | MRI/EEG/Neuropsychological Features                                                                 | Outcome Measures                                                                 |
|----------------------|--------------------|---------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| Pulido-Valdeolivas   | 161 patients with MS | Annualized rates of brain volume loss (MRI, T1 and T2 lesions) and retinal thinning (optical coherence tomography) | Predictive of early MS progression (within 5 years)                               |
| Popescu 2013         | 97 patients with RRMS | Higher T2 lesion volume                                                                           | Assessments predictive of:                                                        |
| Horakova 2009        | 181 patients with early RRMS | Higher T2 lesion volume                                                                           | Correlated with an increase in EDSS (over 5 years)                               |
| Deloire 2011         | 44 patients with MS | Lesion load, Lesion magnetization transfer ratio                                                  | No predictive evidence of:                                                        |
| Rio 2018             | 283 patients with RRMS | 3 or more new T2 lesions, 2 or more gadolinium-enhancing lesions                                | Predictive of sustained EDSS worsening of at least 2 points (year 8)               |
| Goodin 2016          | 376 patients with MS | T2 burden of disease, Change in T2 burden of disease                                              | Predictive of negative disability outcomes (year 16)                              |
| Prosperini 2017      | 987 patients with RRMS | Evidence of disease activity, Minimal evidence of disease activity                                | Predictive of higher risk of reaching EDSS score of 6 or higher (year 10)         |
| Uher 2017            | 162 patients following first demyelinating event suggestive of MS | Loss of NEDA-3, NEDA-4                                                                          | Assessment predictive of:                                                         |
| Kappos 2016          | 2355 patients with RRMS | NEDA-4                                                                                           | No predictive evidence of sustained disability progression (year 10)              |
|                      |                    |                                                                                                   | Predictive of longer time to secondary progressive MS conversion (over 8 years)   |
| Study       | Participants | Findings |
|-------------|--------------|----------|
| Goodin 2016 | 376 patients with MS | - NEDA-3\textsuperscript{10} – No predictive of evidence of longer time to secondary progressive MS conversion (over 8 years)  
- Clinical NEDA\textsuperscript{11} – Predictive of negative disability outcomes (year 16)  
- No predictive evidence of time to death (over 21 years)  
- NEDA-3\textsuperscript{a}  
- NEDA-3b\textsuperscript{13}  
- NEDA-4\textsuperscript{14}  
- No predictive evidence of negative disability outcomes (year 16)  
- No predictive evidence of survival (year 21) |
| Rio 2018    | 283 patients with RRMS | - Minimal evidence of disease activity -1\textsuperscript{15} – Predictive of sustained EDSS worsening of at least 2 points (year 8)  
- Minimal evidence of disease activity -2\textsuperscript{16}  
- Evidence of disease activity – No predictive evidence of sustained EDSS worsening of at least 2 points (year 8) |
| Study          | Participants | Key Findings and Correlations |
|---------------|--------------|-------------------------------|
| Chitnis 2018  | 122 patients with MS | Higher average yearly NfL measurements, Higher average yearly NfL levels, Higher NfL levels, Yearly and average yearly NfL measurements, Average yearly NfL, NfL at each year. Assessments were predictive of: Reduction in brain parenchymal fraction, Higher T2-lesion load, Higher MFIS but only years 1-3 were predictive over the 10-year period evaluated. |
| Moccia 2016   | 155 patients with RRMS | Lower SDMT, Higher number of failed impaired tests of the BRB-N, Cognitive impairment (measured within 6 months from MS diagnosis), Lower SDMT, Higher number of failed impaired tests of the BRB-N, Lower SRT-D, Failure of at least 3 tests of the BRB-N. Assessment predicted: Achievement of EDSS score of 4 (year 10), Conversion from RRMS to SPMS (year 10). No predictive evidence of: Occurrence of clinical relapse (year 10). |
| Lavorgna 2014 | 241 patients with RRMS | Higher age. Association with: Achievement of EDSS score of 4. Not associated with: Conversion from RRMS to SPMS. |
| Study          | Number of Patients | Demographic/Outcome Measures | Association/Not Associated with: |
|---------------|--------------------|-----------------------------|----------------------------------|
| **Moccia 2016** | 155 patients with RRMS | Higher age, Gender | Achievement of EDSS score of 4, Conversion from RRMS to SPMS |
| **Deloire 2011** | 44 patients with MS | Age, Gender, Education | No predictive evidence of change in: Memory z score, Change in information processing speed |
| **Jokubaitis 2015** | 2271 patients with MS | Pregnancy | Predictive of a decrease in median EDSS score (over 8 years) |
| **Jokubaitis 2015** | 2271 patients with MS | Higher annualized relapse rate | Predictive of an increase in mean EDSS score (over 8 and 10 years) |
| **Moccia 2016** | 155 patients with RRMS | Higher annualized relapse rate | Association with achievement of EDSS score of 4 (year 10), Not associated with a conversion from RRMS to SPMS (year 10) |
| **Rio 2018** | 283 patients with RRMS | Presence of one relapse without changes in EDSS score (year 1) | Predictive of: Conversion from RRMS to secondary MS (year 12), Five-step EDSS worsening (year 12) |
| **Lavorgna 2014** | 241 patients with RRMS | Longer disease duration | Association with achievement of EDSS score of 4 |
| **Moccia 2016** | 155 patients with RRMS | Disease duration | No association with: Achievement of EDSS score of 4.0 (year 10), Conversion from RRMS to SPMS (year 10) |
| **Rothman 2019** | 140 patients with RRMS | Lower total macular volume | Predictive of EDSS worsening, Correlated with higher EDSS (year 10) |
Deloire 2011

**44 patients with MS**

| Montgomery-Åsberg Depression Rating Scale | No predictive evidence of:
|------------------------------------------|-------------------------------------|
|------------------------------------------|-------------------------------------|
| Change in memory z score                |                                    |
| Change in information processing speed   |                                    |

Abbreviations: BRB-N, Rao Brief Repeatable Battery of Neuropsychological Tests in Multiple Sclerosis (version A); MFIS, Modified Fatigue Impact Scale; MRI, magnetic resonance imaging; MS, multiple sclerosis; NEDA, no evidence of disease activity; NfL, neurofilament light; RRMS, relapsing remitting MS; SDMT, Symbol Digit Modalities Test; SRT-D, Selective Reminding Test-Delayed; T25-FW, Timed 25-Foot Walk

Notes:

1. Fraction of intracranial volume
2. One-point increase for EDSS score ≤ 5.5 and a 0.5-point increase (confirmed after at least 3 months) for EDSS score > 5.5
3. Compared to patients with NEDA (absence of relapses, of confirmed EDSS worsening and of MRI activity)
4. Evidence of disease activity: occurrence of ≥ 1 relapses with either confirmed EDSS worsening or ≥ 1 gadolinium-enhancing lesions or ≥3 new T2 lesions
5. Minimal evidence of disease activity: either 1 relapse with ≤2 new T2 lesions or < 3 new T2 lesions or < 2 gadolinium-enhancing lesions, in the absence confirmed EDSS worsening
6. NEDA-3: absence of new clinical relapse, disability progression, and active MRI lesions during follow-up, and no evidence of disease activity
7. Sustained disability progression was defined as an increase in EDSS by 1.0 point (if baseline EDSS score > 0) or 1.5 points (if baseline EDSS score = 0), confirmed after 12 months
8. NEDA-4: absence of new clinical relapse, disability progression, active MRI lesions as well as increased whole brain volume loss during follow-up
9. NEDA-4: no MRI lesion activity, no relapses, no confirmed progression of disability, and annual brain volume loss < 0.4%
10. NEDA-3: no MRI lesion activity, no relapses, and no confirmed progression of disability
11. Clinical NEDA: no relapses, and no EDSS worsening
12. NEDA-3a: no relapses and no new T2-active lesions during treatment, and no confirmed 1-point EDSS progression
13. NEDA-3B: no relapses and no EDSS/T2 burden of disease worsening
14. NEDA-a4: no relapses; no EDSS/T2 burden of disease worsening, and no enlargement in 3rd ventricle size
15. Minimal evidence of disease activity-1: presence of one relapse with 0 or 1-2 new T2 lesions
16. Minimal evidence of disease activity-2: presence of <3 new T2 lesions or <2 gadolinium-enhancing lesions