The Barcelona baseline risk score to predict long-term prognosis after a first demyelinating event: a prospective observational study

Carmen Tur,[a,*] Pere Carbonell-Mirabent,[a] Susana Otero-Romero,[a] Álvaro Cobo-Calvo,[a] María Jesús Arévalo,[a] Helena Ariño,[a] Georgina Arrambide,[a] Cristina Auger,[b,d] René Carvajal,[a] Joaquín Castilló,[a] Manuel Comabella,[a] Ingrid Galán,[a] Luciana Midaglia,[a] Carlos Nos,[a] Agustín Pappolla,[a] Deborah Pareto,[b,d] Jordi Río,[a] Breogán Rodríguez-Acevedo,[a] Estibaliz Saez de Gordoa,[b,d] Ángela Vidal-Jordana,[a] Ana Zabalza,[a] Jaume Sastre-Garriga,[a] Àlex Rovira,[b,d] Xavier Montalban,[a,c,e] and Mar Tintoré[a,c,e]

[a]Multiple Sclerosis Centre of Catalonia (Cemcat), Department of Neurology, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain
[b]Section of Neuroradiology, Department of Radiology, Vall d'Hebron University Hospital, Spain
[c]Universitat de Vic-Universitat Central de Catalunya (UVic-UCC), Spain
[d]Universitat Autònoma de Barcelona, Barcelona, Spain

The Lancet Regional Health - Europe 2025;53: 101302
Published Online xxx
https://doi.org/10.1016/j.lanepe.2025.101302

Summary

Background In multiple sclerosis (MS), predicting at symptom onset who will develop early and severe disability is an unmet need with significant therapeutic implications. Here we propose the Barcelona-Baseline Risk Score (BRS) model to predict long-term disease outcomes in a flexible and generalisable manner.

Methods Using prospectively acquired data from the Barcelona first-attack cohort, we created the Barcelona-BRS model as a set of six Weibull survival models of time to an Expanded Disability Status Scale score of 3.0, built with flexible combinations of predictors, including sex, age at first attack, and number and topography of T2 lesions, among others, adaptable to data availability. Data-driven risk groups were identified and compared in terms of long-term clinical and MRI outcomes, including relapse-associated worsening (RAW), progression independent of relapse activity (PIRA), conversion to secondary progressive MS (SPMS), lesional and brain volumetric data, and patient-reported/administered clinical scores, through Kaplan–Meier and mixed-effects models. Finally, we externally validated our model in a completely unseen cohort.

Findings We included 1074 patients (737 [69%] female, mean age: 31.7 years) with a first demyelinating attack. Over a median follow-up of 11.9 years, 375 (35%), 298 (28%), and 94 (8.8%) developed RAW, PIRA, and SPMS, respectively. Weibull models included age at first attack, number of brain T2 lesions, and disability at first visit as main predictors. Four data-driven groups of increasing risk of unfavourable outcomes were created: Light-Green-BRS (N = 258), Dark-Green-BRS (N = 319), Orange-BRS (N = 321), and Red-BRS (N = 176), which, over time, behaved significantly differently across disability, quality of life, and MRI measures, being the Red-BRS the group with worst outcomes ($p < 0.01$). The results in the external validation cohort (N = 139, 100 female [72%], 34 years) mirrored those of the original one.

Interpretation The robustness, flexibility, and generalisability of the Barcelona-BRS model support its consideration as a ready-to-use tool for clinical practice.

Funding None.

Keywords: Multiple sclerosis; Predictive model; Disease progression; MRI; Progression independent of relapse activity; Relapse-associated worsening

*Corresponding author. Multiple Sclerosis Centre of Catalonia (Cemcat), Department of Neurology/Neuroimmunology, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona, Barcelona, Spain.
E-mail address: ctur@cem-cat.org (C. Tur).
[e]Joint last authors.

Research in context

Evidence before this study

Over the last 25 years, several predictive models of mid- and long-term prognosis in multiple sclerosis (MS) have been proposed, with the ultimate aim of helping clinicians in their decision-making processes while deepening into the mechanisms driving disability accumulation. To evaluate the current evidence of predictive models on MS, we carried out a structured review of the published scientific literature in MEDLINE, Scopus, Google Scholar, and PubMed for relevant reports published in any language from Jan 1, 1995, to September 1, 2024, using search terms that included "multiple sclerosis", "demyelinating attack", "inflammatory-demyelinating disease", "clinically isolated syndrome", or "demyelination", AND "predictive model", "prediction", "prognostic model", "risk stratification", or "risk assessment tool". We found that many of the models proposed so far have been defined in the context of treatment response, and most of them have focused on just a few predictors of the same kind at a time instead of integrating all possible predictors into the same model. This has hampered the development of comprehensive risk scores to be used in the clinic, which would be, on the other hand, very useful for an optimal approach to individual patients' disease, in line with the concept of personalised medicine.

Added value of this study

In this paper, we present the Barcelona-BRS model, which effectively classifies patients into distinct risk groups for several mid- and long-term clinical and MRI outcomes, confirming the association between older age at first attack, male sex, spinal cord (or multifocal) first-attack topographies, greater disability scores at first visit, greater T2 lesion burden and contrast-enhancing lesions with an unfavourable prognosis. Here we provide evidence of the robustness and generalisability of the model, through externally validating its coefficients in a completely unseen cohort of people with early MS. Model coefficients, which can be directly applied to individual patients to estimate their personalised risk, and data-driven thresholds to be applied to such risks, to classify people with MS into risk groups (BRS groups), are also provided in this paper.

Implications of all the available evidence

The proposed Barcelona-BRS model is a risk assessment tool to be used at the first demyelinating attack, which is based on traditional risk factors such as age, sex, and lesion burden. It provides a ready-to-use solution for clinicians aiming to predict long-term outcomes and personalise treatment plans for MS patients.

Introduction

Multiple sclerosis (MS) is a condition characterised by inflammation and demyelination of the central nervous system (CNS), with significant clinical and pathological variability.[1] It commonly affects young adults, leading to long-term disability,[2,3] and impacts quality of life and work ability, causing societal and economic consequences.[2] In 85% of cases, MS begins with a clinically isolated syndrome (CIS), often followed by relapses in a relapsing-remitting phase.[2] These relapses can result in permanent disability if recovery is incomplete (relapse-associated worsening, RAW). However, the primary mechanism of disability in MS is progression independent of relapse activity (PIRA).[4–6]

Recent years have seen a rise in disease-modifying treatments for MS.[3] Some drugs can effectively reduce CNS inflammation and alter the disease's natural history. High-efficacy treatments, particularly in early stages, have shown the most success in preventing disability,[7] though they carry risks of side effects and high costs. For milder cases, moderate-efficacy treatments may suffice, though predicting the best treatment strategy early in the disease remains challenging.

Over the last few years, big efforts have been made to predict MS prognosis, mainly focusing on treatment response.[8–13] However, comprehensive models integrating multiple factors for personalised care are lacking. In other fields, such as cardiovascular research and ischaemic heart disease, tools like the Framingham score have advanced clinical management. A similar risk score for MS prognosis, adaptable across clinical settings, is yet to be introduced.

This study aims to build the Barcelona-Baseline Risk Score (Barcelona-BRS) to predict long-term prognosis at the first demyelinating attack, using clinical, biological, and radiological data. It consists of a flexible risk assessment tool at the first attack, valid for almost any clinical setting, even with limited available data. The ultimate goal of the Barcelona-BRS, tested in an external, independent cohort, is to assist clinicians with the decision-making process concerning personalised therapeutic approaches for MS through risk stratification at the first demyelinating attack.

Methods

Study design and participants

This retrospective analysis includes longitudinal data from patients with a first demyelinating attack suggestive of MS who have been prospectively followed up in the ongoing Barcelona first-attack cohort. The study also involves applying the Barcelona-BRS model to an independent cohort from the Multiple Sclerosis Partners Advancing Technology and Health Solutions (MS PATHS) dataset. It was created between November 2016 and January 2022 as a learning health initiative that

systematically gathered clinical and imaging data from routine practice at ten MS centers in the United States, Germany, and Spain.[14]

Barcelona first-attack cohort

Patients from the Multiple Sclerosis Center of Catalonia (Cemcat) experiencing their first demyelinating attack between 1994 and 2023, under the age of 50, were included.[11] Patients had to be assessed within three months of the first attack, with brain T2 lesion data available within 3–5 months.[11]

External validation cohort

We used a cohort extracted from the large MS PATHS longitudinal database (2016 onwards),[14] consisting of standardised routinely acquired clinical and MRI data of people with a confirmed diagnosis of MS.[14] We included patients under 50 at the first attack, with time between that the and the first MRI assessment of less than two years. Of note, since the Vall d'Hebron University Hospital was part of the MS PATHS project, we excluded all patients belonging to our hospital to ensure validation independence.

Study variables

Demographic and clinical data

We obtained information on demographic data, including sex and age at first demyelinating attack; clinical data, including date and topography of the first demyelinating event, presence and dates of relapses (recorded at each visit), disability status according to the Expanded Disability Status Scale (EDSS)[15] at baseline (i.e., within three months after the first demyelinating attack) and then (at least) annually; and data on disease-modifying treatment (DMT) exposure, including starting dates and types of DMT exposure (i.e., high efficacy drugs: monoclonal antibodies, mitoxantrone, and cyclophosphamide; and low/moderate efficacy ones: the rest). Additionally, in a subset of patients of the Barcelona first-attack cohort who participated in the MS PATHS study between 2016 and 2022 (i.e., the Barcelona-MS PATHS subcohort from now on), we collected data on: general disability, through the self-reported Patient Determined Disease Steps (PDDS)[16] (at least once); neuroperformance, through the self-administered Multiple Sclerosis Performance Tests (MSPT),[14] which included the walking speed test (WST), manual dexterity test (MDT), processing speed test (PST), and the contrast sensitivity test (CST) (at least once); and quality of life, through the self-reported NeuroQoL questionnaire (at least once).

Brain and spinal cord MRI data

All patients underwent brain MRI scans (T1-and T2-weighted sequences, at 1.5 or 3T MRI scanners) at the first demyelinating attack (within five months of symptom onset), 12 months later, and then according to clinical practice: every 12 months for most treated patients, or at least every 5 years if untreated. Most patients also underwent spinal cord MRI, although this was only systematically performed at the first attack in all patients after 2007, regardless of the topography of this first attack. The data obtained from the brain and spinal cord MRI included: T2 lesion number and topography at the first attack, presence of contrast-enhancing lesions (CEL) at the first attack, and number of new brain T2 lesions over the follow-up.

Additionally, for the Barcelona-MS PATHS subcohort, brain volumetric measures, including T2 lesion volume and brain parenchymal fraction (BPF), were obtained from 3D T2-FLAIR and 3D T1-weighted images (isotropic voxel resolution 1 mm^3) acquired in a Siemens 3T MR scanner. All MR parameters can be found somewhere else.[14] As exploratory outcomes, we analysed grey matter fraction (GMF), cortical GMF (CGMF), deep GMF (DGMF), thalamic fraction, and white matter fraction (WMF).

Laboratory data

Some patients were assessed on the presence of IgG oligoclonal bands (OBs) in the cerebrospinal fluid (CSF) and serum at the first attack, tested on agarose gel isoelectric focusing combined with immunoblotting, as previously reported.[11] Presence of CSF OBs in the absence of serum OBs was defined as 'positive OBs'.

Statistical methods

Creation of the predictive tool

We visualised our predictive tool, the Barcelona-BRS model, as an *umbrella* model made of several survival Weibull models able to adapt to different clinical settings with variable amounts of available data. It was conceptualised as a risk stratification tool, so that patients could be classified into risk groups (or BRS groups, from now on) (Fig. 1).

We first defined a set of potential predictors at the first attack/early MS stages, including demographic (age, sex), clinical (disability score at first visit, topography of first attack), brain MRI (number of T2 lesions, presence of infratentorial T2 and contrast-enhancing T1 lesions), spinal cord MRI (presence of T2 lesions), and laboratory variables (positive OBs). Then, to identify the most informative variables to be included in the Barcelona-BRS model, a survival random forest model of time to EDSS 3.0 was built using the full cohort (N = 1074) including all possible predictor variables. Random forest models are able to order the variables associated with a given outcome, in this case time to EDSS 3.0, based on the strength of the association. Those variables with greatest importance (individually) were selected (Fig. 1; Supplementary Figure S1). Of note, disease-modifying treatment (DMT) exposure was not included in the Weibull models, which only

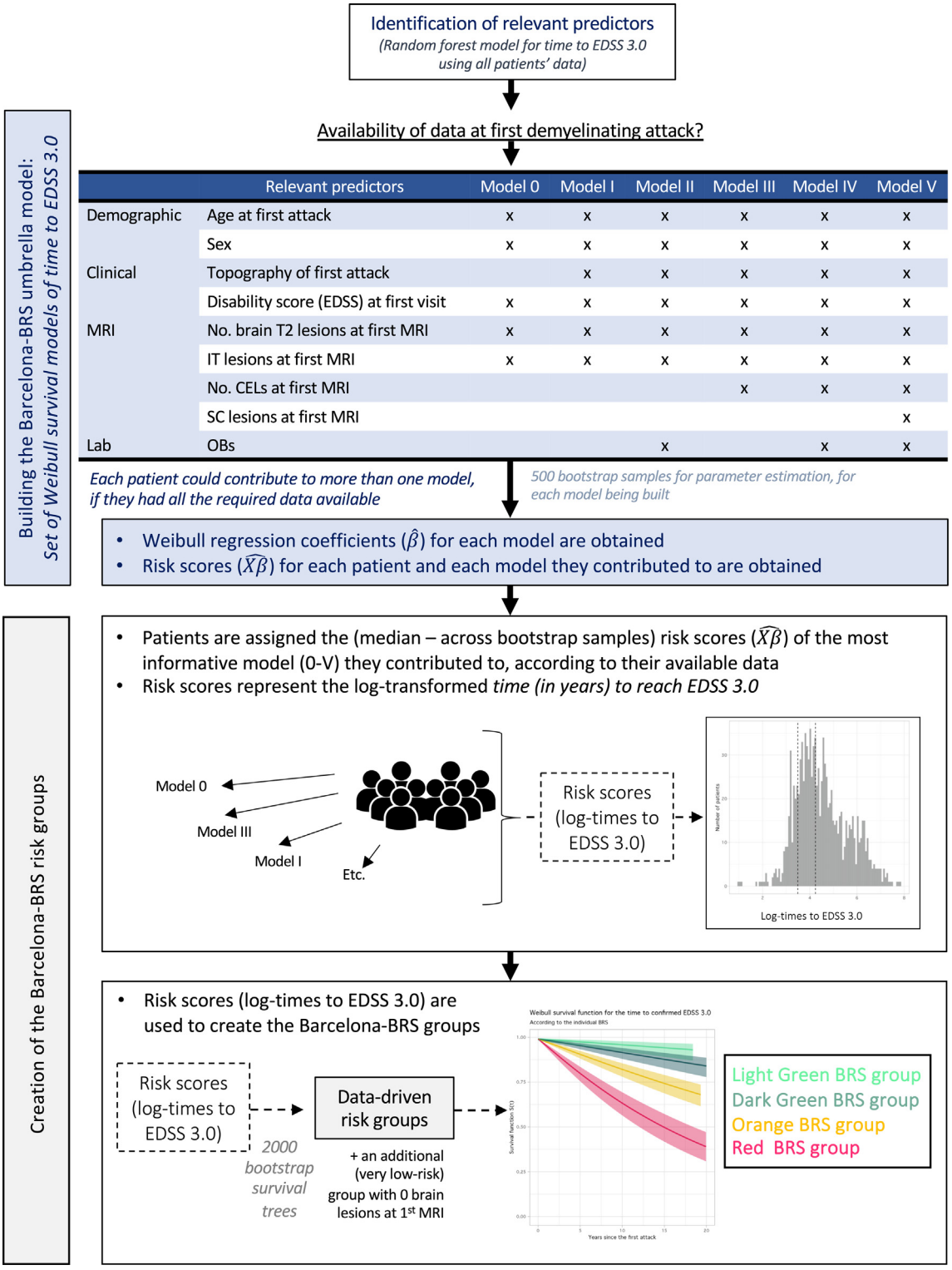

***Fig. 1:* Study design.** Schematic view of the path to creating the Barcelona-BRS umbrella model and obtaining the derived Barcelona-BRS risk groups. *Abbreviations:* BRS: baseline risk score; CELs: Contrast-enhancing lesions; EDSS: Expanded Disability Status Scale; IT: Infratentorial lesions; MRI: Magnetic Resonance Imaging; OBs: oligoclonal bands; SC: spinal cord.

accounted for variables that could be obtained at the time of the first attack, but once the BRS groups were created, to predict long-term outcomes, as explained below (see *Assessment of the association between DMT exposure and prediction of long-term outcomes*). This allowed us to assess the effect of DMT as a time-varying covariate in a more efficient way (since the number of parameters to be estimated was reduced) and in each BRS group separately.

The first Weibull model of time to EDSS 3.0 ('model V') was built with all possible relevant variables as predictors (Fig. 1; Supplementary Table S1). It was created with 500 bootstrap samples without replacement of size equal to 80% of the size of the sub-cohort with non-missing data. Then, for each patient, we estimated a varying number (maximum = 500) of '*log-transformed* times (in years) to EDSS 3.0', and the median value of these log-times was assigned to each patient (see Supplementary Methods 1.1 for more details). Afterwards, five additional models (models 0, I, II, III, IV) were created following the same methodology (500 bootstrap samples for parameter estimation) but including a decreasing number of predictor variables (Fig. 1). The order in which variables were removed from model V until reaching model 0 was partly based on data availability, using our cohort as a representative example of common practice. Model 0 contained the minimum set of 'important' variables from our perspective required to predict the outcome, considering that it was not possible to assess all possible combinations of predictors in separate models. Consequently, these additional models (from IV to 0) were built using increasingly larger sub-cohorts. This approach allowed us to make use of all patients in the cohort, without the need of imputing missing data, and provided the flexibility needed for the applicability of the tool (Fig. 1). Since the sub-cohorts used to build the predictive Weibull models 0-V were not mutually exclusive, patients could be assigned several (from 1 to 6) estimated median log-times to EDSS 3.0, depending on the amount of available data. The median log-time of the *most complex model* (i.e., the model with highest number of predictors) was considered the most accurate median log-time to EDSS 3.0.

After assigning the most accurate median log-time to EDSS 3.0 to each patient, we investigated the best data-driven thresholds of such log-times to be able to classify patients into risk groups (i.e., BRS groups). We carried out this through running conditional inference trees on 2000 bootstrap samples, without imposing the numbers of groups to be created (see Supplementary Methods 1.2 for more details). For each one of these 2000 bootstrap survival trees, we recorded the number of terminal nodes (i.e., groups) of risk. The most frequent number of groups was considered as the number of data-driven risk categories. The median cut-point values that defined these terminal nodes were considered as the best data-driven thresholds to classify patients into risk groups. These were then recorded to be applied to the external validation cohort later. Afterwards, the group with the lowest risk of reaching EDSS 3.0 was further split into two groups based on the presence (low risk) or absence (very low risk) of brain T2 lesions. Each final BRS group was assigned a colour, which was then used throughout the paper.

Of note, since patients were included in our first-attack cohort based on clinical grounds, some baseline MRI studies could be normal, as explained below.

Definition of long-term clinical and MRI outcomes

For all patients, we obtained information on the following outcomes:

- McDonald 2017 criteria for MS;[17]
- second clinical attack;
- confirmed and sustained disability worsening (CSDW), which required the definition of confirmed disability worsening (CDW) as an increase in EDSS of 1.5, 1.0, or 0.5 EDSS points if baseline EDSS was 0, 1.0–5.0 (both included), or ≥5.5, respectively, confirmed after at least 6 months; then, CSDW was defined as a CDW where the last EDSS was equal to/higher than the EDSS score showing the increase used to define CDW;
- confirmed & sustained EDSS 3.0: event of EDSS 3.0 confirmed after at least 6 months and sustained until the end of the follow-up;
- first event of RAW, either after the first attack: EDSS score ≥1.5 during the first year after the first attack, confirmed also during the first year, or in a subsequent relapse: CDW that did not occur within a relapse-free period (RFP), defined as that period starting 3 months after a relapse (or 6 months after the first attack) until the next attack, where the confirmation score could fall within or outside a RFP.
- first event of PIRA, defined as a CDW occurring within a RFP and confirmed within a RFP;[5]
- secondary progressive MS (SPMS), defined as disability worsening (DW) and the reaching of an EDSS 4.0, always using only EDSS scores that are 1 month after a relapse, confirmed at 6 months;[18]
- EDSS trajectories over time;
- annual rate of new T2 lesions >2 (this threshold has proven clinically relevant in studies on treatment response, as its predictive power is evident when observed during DMT exposure[19]).
- cumulative number of new T2 lesion trajectories.

Additional outcomes from the Barcelona-MS PATHS subcohort included PDDS, MSPT scores (WST, MDT, PST, CST), quality of life scores (NeuroQoL t-scores), brain T2 lesion volume and brain parenchymal fractions at the last MS PATHS assessment.

Prediction of long-term multidimensional outcomes

For binary events, we built Kaplan Meier survival models and estimated the percentages of event occurrence (1-survival probability) at each follow-up year for each BRS group. For longitudinal continuous data, we built linear mixed-effects models with random intercept and random slope (although, for visualisation purposes, we used splines mixed models). In these models, the variable being explored as outcome of interest was considered as the dependent variable, and the variable *time from first attack* (in years) was considered as the main explanatory variable, together with the BRS group variable and the interaction term *time* X BRS group. Whenever the interaction term was significant ($p < 0.05$), we assumed BRS groups were significantly associated with the longitudinal trajectories. Since BRS groups already incorporate all risk factors, these models were not further adjusted by any clinical or demographic characteristics. Finally, we built linear regression models for continuous clinical and MRI variables measured at a given time point (i.e., at the last MS PATHS evaluation), including PDDS, MSPT, NeuroQoL, and brain volumetric data, adjusting for disease duration at that time point.

Assessment of the association between DMT exposure and prediction of long-term outcomes

Cox proportional hazards models were built, where the binary outcome (e.g., developing a second attack) was the dependent variable and the BRS group was the main predictor of time to reaching the binary outcome, together with DMT exposure as a time-dependent variable. Whenever the DMT variable was significant ($p < 0.05$), we assumed there was an association between DMT exposure and the risk of developing the outcome (immediately) afterwards. Furthermore, survival Kaplan–Meier models for each binary outcome (and the derived heatmaps of event occurrence at every year after) were built for patients with an early DMT exposure (within 12 months of symptom onset) and those with a late DMT exposure (DMT initiation after 3 years of symptom onset).

External validation analysis

Study variables included demographic data (sex and age at symptom onset), clinical data (PDDS, from which we estimated the first EDSS, MSPT, and NeuroQoL), brain MRI data from 3D T2-FLAIR and 3D T1-weighted images (T2 lesion volume/number of T2 lesions estimated from T2 lesion volume and other volumetric measures). Some of these variables were used as predictor variables of the BRS model (although some of them required some *translation* to be used in the BRS [see Supplementary Table S2]). The coefficients of the most appropriate BRS model based on data availability in the validation cohort (Supplementary Table S3), which in our case was model 0, were directly applied, without any re-estimation or re-calibration, to predict the log-times to EDSS 3.0 in that external cohort. These predicted log-times to EDSS 3.0 were then used to create the BRS groups, after applying the data-driven thresholds obtained in the original cohort. Some other variables were instead used as long-term outcomes (e.g., PDDS/MSPT/NeuroQoL data, T2 lesion volumes, and brain tissue fractions at last follow-up). These outcomes were predicted by the BRS group through linear regression models adjusted for disease duration. We used bootstrap-based methods to estimate the uncertainty around our predictions (please see Supplementary Methods 1.3 for more details). As a sensitivity analysis, we investigated the consistency across BRS model predictions after performing data imputation in the validation cohort.

All analyses were carried out in RStudio 2021.09.0 Build 351. Bootstrap-based analyses were carried out manually. The rest of the analyses were carried out using freely available R packages.

Ethics approval

The ethics committee of Vall d'Hebron Hospital approved the study (XMG-INT-2014-01), and informed consent was obtained from all participants.

Role of funding source

None of the funding sources disclosed has had a major role in the writing of the manuscript or in the decision to submit it for publication. None of the funding sources disclosed has had any role in data collection, analysis or interpretation, study design, patient recruitment, or any aspect pertinent to the study. None of the authors has been paid to write this article by any pharmaceutical company.

Results

Descriptive statistics

Out of a total of 1561 patients who were seen at Cemcat with a first episode suggestive of a demyelinating attack until February 2024, we finally included 1074 patients (Fig. 2). Over a median follow-up period of 11.9 years (IQR 6.1, 19.4), 538 (50%) had a second attack, 375 (35%) developed some RAW (i.e., at the first attack or at subsequent relapses), although only 128 (12%) developed RAW at subsequent relapses, and 298 (28%) developed PIRA. Table 1 shows the main demographic, clinical, and MRI data at study baseline and over the follow-up. Of the 1074 patients, 450 (42%) participated in the MS PATHS study, conforming the *Barcelona-MS PATHS subcohort*.

Creation of the predictive tool

Six survival Weibull models were built as described in the Methods section. The most complex model (model V) was built with the smallest number of patients (N = 407).

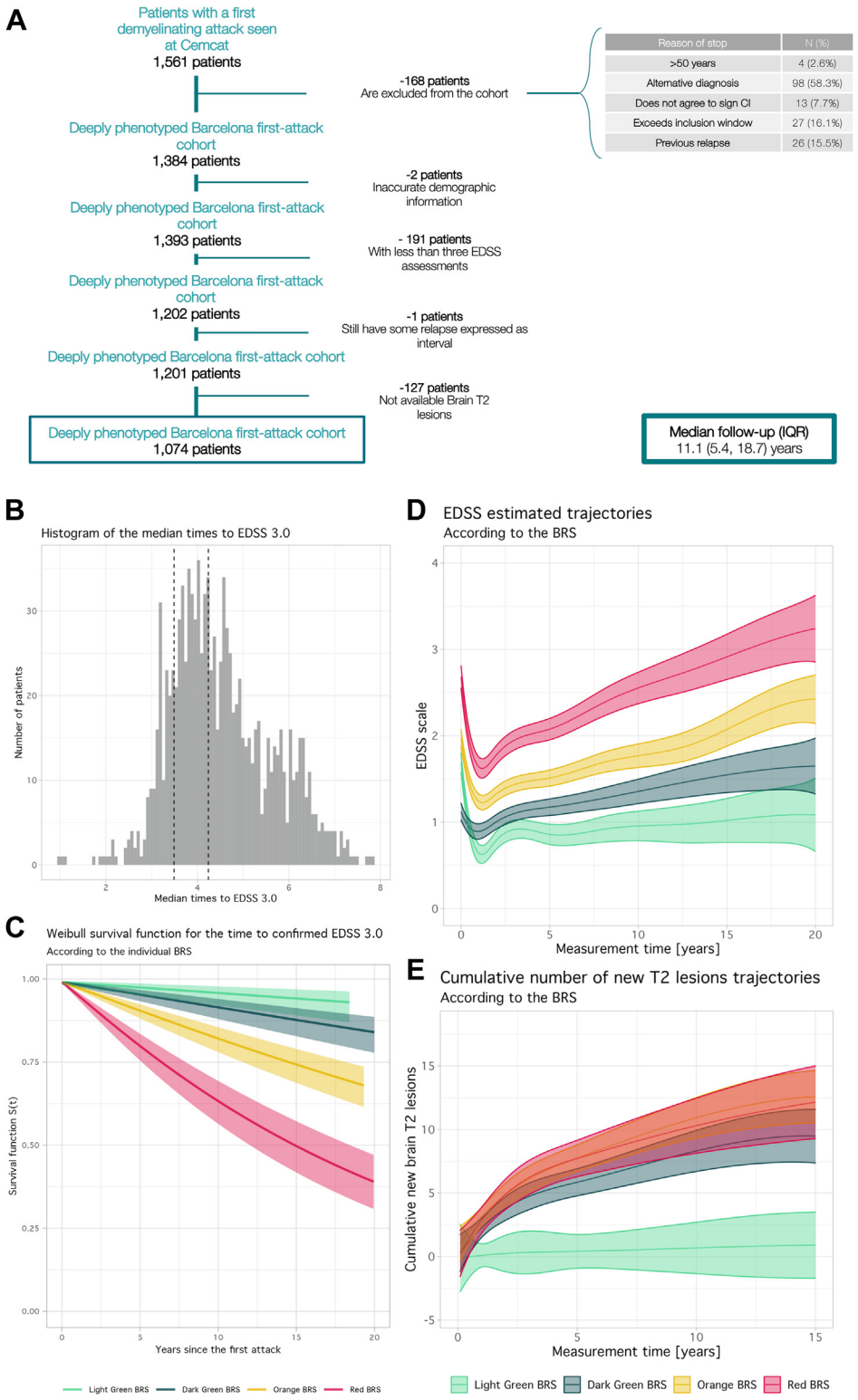

Fig. 2: **A.** Patient Flowchart. Out of a total of 1561 patients with a first demyelinating attack seen at Cemcat, 1074 finally met our inclusion criteria. *Abbreviations:* IQR: interquartile range. **B.** This panel shows the most informative times (in years) to EDSS 3.0, which correspond to the median times (across bootstrap samples) of the most informative model for each patient (A). Fig. 3A also shows the data-driven thresholds to

However, it showed the greatest Harrell's C (0.79), indicating the highest accuracy. Instead, model 0 was built with the largest number of patients possible (N = 1074—all) but had the lowest Harrell's C value (0.72). Supplementary Table S3 shows the regression coefficients for all Weibull models. Fig. 2 shows the distribution of the most informative median log-times to EDSS 3.0. In the Barcelona first-attack cohort, no patients were assigned the log-times to EDSS 3.0 from model 0 since all patients had the minimum required data to build at least model I.

When investigating the best data-driven thresholds to create risk groups, we found that patients could be classified as:

- Red BRS: median log-time to EDSS 3.0: ≤3.49 years, N = 176 [16.4%],
- Orange BRS (median log-time to EDSS 3.0: >3.49 and ≤4.24 years, N = 321 [29.9%]),
- Green BRS (median log-time to EDSS 3.0: >4.24 years, N = 577 [53.7%]), being this group further split into Light Green BRS (N = 258 [24.0%]) and Dark Green BRS (N = 319 [29.7%]), as explained (Fig. 2).

Table 1 shows the baseline and follow-up patient characteristics for each BRS risk group.

Prediction of clinical outcomes

All long-term clinical outcomes differed across the BRS groups. For instance, whereas 77 (44%) of the patients in the Red BRS group reached confirmed and sustained EDSS 3.0 during the follow-up, this outcome was only achieved by 9 (3.5%) of the Light Green BRS one. Similarly, whereas 42% of patients within the Red BRS group had ≥1 PIRA event during the follow-up, only 19% of the Light Green BRS experienced that outcome.

Kaplan Meier estimations of the cumulative probabilities of reaching each one of the clinical outcomes at each year of follow-up until year 20 were significantly different across BRS groups (all log-rank p-values $p < 0.01$) (Fig. 3). The four BRS groups also behaved differently in terms of EDSS trajectories over time, being the Light Green BRS group the one with the smallest annual EDSS increase: –0.026 (95% CI –0.043, –0.010), p = 0.002, and the Red BRS group the one with the highest: 0.055 (0.037, 0.073), $p < 0.001$ (Fig. 2; Supplementary Table S4).

Finally, using only the Barcelona-MS PATHS subcohort, the four BRS groups also behaved differently regarding the self-reported and self-administered tests: Light Green BRS patients showed the mildest PDDS and neuroperformance scores, followed by the Dark Green BRS, Orange BRS, and Red BRS groups (Supplementary Figure S2). For the NeuroQoL scores, the four BRS groups also behaved differently, especially for the function scores: the perception of upper and lower extremity function was significantly worse in the Red BRS group than in the Light Green BRS one, being the Dark Green BRS and Orange BRS consistently between both extreme BRS groups (Supplementary Figure S3).

Prediction of MRI outcomes

Light Green BRS patients showed the lowest rate of reaching >2 new T2 lesions/year, followed by the Dark Green BRS group, and very closely by both the Orange BRS and the Red BRS groups (log-rank p-value <0.001) (Fig. 3). Regarding the cumulative number of new T2 lesions trajectories over time, the Light Green BRS group also behaved very differently from the other three, especially from the Orange BRS and Red BRS groups, which behaved almost identically (Fig. 2; Supplementary Table S5).

Using the Barcelona-MS PATHS subcohort, we also found that the four BRS groups behaved differently in terms of T2 lesion volumes assessed at the last MS PATHS assessment, especially total T2 lesion volume and periventricular T2 lesion volume (Supplementary Figure S2): patients from the Red BRS risk group had, on average, 11.16 (95% CI 5.11, 17.22) mL greater lesion volume than the Light Green BRS group at last MS PATHS assessment ($p < 0.001$). Dark Green BRS and Orange BRS had intermediate behaviours, worse than Light Green BRS but better than Red BRS. Similarly, Red BRS patients had smaller global and regional brain parenchymal fractions than the Light Green BRS. These differences were mainly evident for the BPF (–0.020 [95% CI –0.036, –0.003], p = 0.019) (Supplementary Figure S2) and the DGMF (–0.002 [–0.003, –0.0006], p = 0.005) (Supplementary Figure S4).

create risk groups, thanks to 2000 survival trees. As explained in the results section, we found three data-driven groups: Red BRS (estimated median time to EDSS 3.0 below or equal to 3.49 years, N = 176 [16.4%]), Orange BRS (estimated median time above 3.49 and below or equal 4.23 years, N = 321 [29.9%]), and green-BRS (median time above 4.23 years, N = 577 [53.7%]). **C.** Weibull model of time to EDSS 3.0 for the four Barcelona-BRS groups. **D.** Estimated EDSS trajectories over time for the four Barcelona-BRS risk groups. For visualisation purposes, this figure shows the trajectories using mixed splines models. However, regression estimates of the slopes over time have been obtained through linear mixed-effects models (see Supplementary Table S3). **E.** Estimated trajectories of cumulative number of new T2 lesions over time for each one of the four Barcelona-BRS risk groups from the first demyelinating attack (time = 0 years). As can be observed, and as also shown in Fig. 4 (heatmap for different outcomes), the light green-BRS group behaved very differently from the other three, and especially from the Orange BRS and Red BRS groups, which behaved almost identically. The regression estimates of the slopes over time have been obtained through linear mixed-effects models (see Supplementary Table S4). *Abbreviations:* BRS: baseline risk score; EDSS: Expanded Disability Status Scale.

Characteristic	Overall, N = 1074	Light green BRS, N = 258	Dark green BRS, N = 319	Orange BRS, N = 321	Red BRS, N = 176	p-value
Baseline characteristics						
Sex, N (%)						<0.001
Female	737 (69%)	174 (67%)	247 (77%)	230 (72%)	86 (49%)	
Male	337 (31%)	84 (33%)	72 (23%)	91 (28%)	90 (51%)	
Age at first demyelinating event in years, median (IQR)	31.7 (25.9, 38.6)	32.5 (25.4, 38.9)	30.0 (25.8, 35.4)	31.7 (25.6, 38.8)	35.0 (29.7, 41.6)	**<0.001**
First EDSS, median (IQR)	2.0 (1.0, 2.5)	2.0 (1.0, 2.5)	1.0 (0, 2.0)	2.0 (1.5, 3.0)	3.0 (2.0, 3.5)	**<0.001**
Calendar year of the first demyelinating event	2007.0 (2001.2, 2014.0)	2008.0 (2001.0, 2013.0)	2008.0 (2003.0, 2015.0)	2006.0 (2001.0, 2013.0)	2007.0 (2000.0, 2014.0)	**0.010**
Topography of the first demyelinating event, N (%)						**<0.001**
Optic nerve	371 (35%)	142 (55%)	133 (42%)	68 (21%)	28 (16%)	
Spinal Cord	311 (29%)	67 (26%)	48 (15%)	119 (37%)	77 (44%)	
Brainstem	266 (25%)	33 (13%)	107 (34%)	100 (31%)	26 (15%)	
Other	126 (12%)	16 (6.2%)	31 (9.7%)	34 (11%)	45 (26%)	
Categorized baseline number of brain T2 lesions, N (%)						**<0.001**
0 lesions	258 (24%)	258 (100%)	0 (0%)	0 (0%)	0 (0%)	
1–3 lesions	155 (14%)	0 (0%)	109 (34%)	42 (13%)	4 (2.3%)	
4–8 lesions	141 (13%)	0 (0%)	61 (19%)	61 (19%)	19 (11%)	
≥9 lesions	520 (48%)	0 (0%)	149 (47%)	218 (68%)	153 (87%)	
Number of brain T2 lesions at baseline, mean (SD)	24.3 (34.6)	0.0 (0.0)	17.7 (20.8)	37.2 (44.3)	45.6 (36.6)	**<0.001**
Unknown	501	139	127	150	85	
OBs, N (%)						**<0.001**
Negative	337 (36%)	164 (77%)	111 (40%)	47 (17%)	15 (9.6%)	
Positive	591 (64%)	48 (23%)	170 (60%)	232 (83%)	141 (90%)	
Unknown	146	46	38	42	20	
Brain contrast enhancing lesions, mean (SD)	1.4 (3.5)	0.0 (0.0)	0.7 (2.3)	1.8 (3.7)	3.1 (5.2)	**<0.001**
Unknown	294	160	38	70	26	
Categorized number of brain contrast enhancing lesions, N (%)						**<0.001**
>2 CEL	109 (14%)	0 (0%)	18 (6.4%)	43 (17%)	48 (32%)	
1–2 CEL	175 (22%)	0 (0%)	58 (21%)	75 (30%)	42 (28%)	
No CEL	496 (64%)	98 (100%)	205 (73%)	133 (53%)	60 (40%)	
Unknown	294	160	38	70	26	
Spinal cord lesions, N (%)						**<0.001**
>3 lesions	72 (11%)	4 (2.2%)	19 (9.5%)	29 (16%)	20 (18%)	
0 lesions	410 (61%)	149 (83%)	125 (63%)	88 (50%)	48 (42%)	
1 lesion	123 (18%)	23 (13%)	35 (18%)	39 (22%)	26 (23%)	
2–3 lesions	62 (9.3%)	3 (1.7%)	20 (10%)	20 (11%)	19 (17%)	
Unknown	407	79	120	145	63	
Presence of spinal cord lesions, N (%)						**<0.001**
None	410 (61%)	149 (83%)	125 (63%)	88 (50%)	48 (42%)	
Some	257 (39%)	30 (17%)	74 (37%)	88 (50%)	65 (58%)	
Unknown	407	79	120	145	63	
Contrast enhancing lesions in the spinal cord, N (%)						**<0.001**
>0 CEL	76 (20%)	8 (11%)	19 (14%)	21 (20%)	28 (38%)	
0 CEL	312 (80%)	63 (89%)	120 (86%)	83 (80%)	46 (62%)	
Unknown	686	187	180	217	102	
Presence of infratentorial lesions, N (%)						**<0.001**
None	405 (46%)	198 (100%)	134 (48%)	59 (24%)	14 (9.3%)	
Some	472 (54%)	0 (0%)	143 (52%)	192 (76%)	137 (91%)	
Unknown	197	60	42	70	25	
Follow-up characteristics						
Follow-up time in years, median (IQR)	11.9 (6.1, 19.4)	10.0 (4.9, 14.5)	11.1 (6.2, 18.2)	14.7 (7.0, 21.4)	13.3 (7.8, 20.5)	**<0.001**
DMT exposure over the follow-up, N (%)	602/1074 (56.05%)	17/258 (6.59%)	197/319 (61.76%)	239/321 (74.45%)	149/176 (84.66%)	**<0.001**
Low/Moderate-efficacy DMT only exposure, N (%)	420/1074 (39.11%)	11/258 (4.26%)	162/319 (50.78%)	159/321 (49.53%)	88/176 (50%)	**<0.001**
High-efficacy DMT exposure, N (%)	182/1074 (16.95%)	6/258 (2.33%)	35/319 (10.97%)	80/321 (24.92%)	61/176 (34.66%)	

(Table 1 continues on next page)

Characteristic	Overall, N = 1074	Light green BRS, N = 258	Dark green BRS, N = 319	Orange BRS, N = 321	Red BRS, N = 176	p-value
(Continued from previous page)						
Time from first attack to first DMT in years, median (IQR)	0.78 (0.47, 2.75)	5.21 (3.15, 10.83)	0.96 (0.53, 2.92)	0.76 (0.46, 2.37)	0.64 (0.39, 1.80)	**<0.001**
Time from first attack to first HE DMT in years, median (IQR)	5.56 (1.49, 12.05)	6.07 (2.49, 7.85)	6.70 (1.25, 12.54)	5.90 (2.33, 11.70)	3.76 (0.59, 12.27)	0.5
McDonald MS 2017	765 (71%)	41 (16%)	258 (81%)	297 (93%)	169 (96%)	**<0.001**
Second demyelinating event	538 (50%)	34 (13%)	172 (54%)	209 (65%)	123 (70%)	**<0.001**
CSDW	384 (36%)	44 (17%)	115 (36%)	125 (39%)	100 (57%)	**<0.001**
Confirmed and sustained EDSS 3, N (%)	187 (17%)	9 (3.5%)	29 (9.1%)	72 (22%)	77 (44%)	**<0.001**
Confirmed EDSS 6, N (%)	53 (4.9%)	0 (0%)	5 (1.6%)	26 (8.1%)	22 (12%)	**<0.001**
Some RAW, N (%)	375 (35%)	41 (16%)	79 (25%)	150 (47%)	105 (60%)	**<0.001**
RAW after the first attack, N (%)	128 (12%)	5 (2%)	39 (12%)	52 (16%)	32 (18%)	**<0.001**
PIRA, N (%)	298 (28%)	49 (19%)	82 (26%)	93 (29%)	74 (42%)	**<0.001**
SPMS, N (%)	94 (8.8%)	1 (0.4%)	16 (5.0%)	43 (13%)	34 (19%)	**<0.001**
T2 rate >2 lesions/year, N (%)	400 (37%)	8 (3.1%)	136 (43%)	167 (52%)	89 (51%)	**<0.001**

Abbreviations: Barcelona-BRS: Barcelona Baseline Risk Score; CELs: Contrast-enhancing lesions; CSDW: Confirmed and sustained disability worsening; EDSS: Expanded Disability Status Scale; IQR: Interquartile Range; MRI: Magnetic Resonance Imaging; MS: multiple sclerosis; OBs: oligoclonal bands; PIRA: progression independent of relapse activity; RAW: relapse-associated worsening; SD: standard deviation; SPMS: secondary progressive multiple sclerosis.

Table 1: **Main characteristics of the deeply phenotyped Barcelona first-attack cohort at baseline and at follow-up.**

Association between DMT exposure and long-term outcomes

When DMT exposure was considered as a time-varying covariate in the predictive models (see methods), it was associated with a lower risk of developing the event immediately afterwards, for all clinical outcomes—although for 'some RAW' and 'yearly rate of new T2 lesions>2' the association did not reach statistical significance (Fig. 4). Significant HRs varied from 0.48 (95% CI 0330, 0.75), p = 0.001 (to predict SPMS) to 0.67 (95% CI 0.46, 0.99), p = 0.0419, (to predict first RAW after the first attack) (Supplementary Table S6). When the DMT exposure as time-dependent covariate was assessed in each one of the risk groups separately, we found that the estimated HRs were very similar across groups—and across outcomes (data not shown). Finally, the heatmaps built based on the Kaplan–Meier curves but considering only DMT exposed patients showed lower predicted risks when DMT exposure occurred early (within 12 months) than late (after 3 years) after the first attack (Fig. 3).

External validation analysis

We included 139 patients from the MS PATHS dataset who fulfilled our inclusion criteria (Supplementary Figure S5). Patient characteristics are shown in Supplementary Table S7. The median follow-up of the cohort was 1.0 years (IQR 0.6, 1.2). After applying the Barcelona-BRS model (model 0, based on data availability), we obtained the estimated log-times to EDSS 3.0 for each patient (Supplementary Figure S5) and then applied the data-driven thresholds obtained in the original cohort to classify patients into three BRS groups: Red BRS (N = 10 [7.2%]), Orange BRS (N = 25 [18%]), Dark Green BRS (N = 104 [74.8%]). Please note that there were no patients with 0 lesions, so no patients were classified as Light Green BRS.

All BRS groups behaved differently at the last MS PATHS assessment in terms of all clinical and MRI outcomes (Supplementary Figures S6–S10). These results are described in detail in the supplementary material (Supplementary Results).

Discussion

We present the Barcelona-BRS model, a risk assessment tool for patients after their first demyelinating attack. It classifies patients into risk groups for mid- and long-term clinical and MRI outcomes, confirming associations between worse prognosis and factors like older age, male sex, spinal cord or multifocal attack topographies, higher disability scores, greater T2 lesion burden, and contrast-enhancing lesions. We have also validated the model externally, confirming its robustness and generalisability in an independent cohort.

The Barcelona-BRS model consists of six Weibull models of increasing complexity, where key predictors like T2 lesion number and EDSS score are integral in all models, consistent with previous findings.[11,20] The model adapts to varying data availability, making it flexible for clinical use: its regression coefficients can be directly applied to individual patients to estimate patient-specific (log) times to EDSS 3.0 at symptom onset. These estimated times can then be used to classify patients into risk (BRS) groups, following the thresholds proposed in this paper.

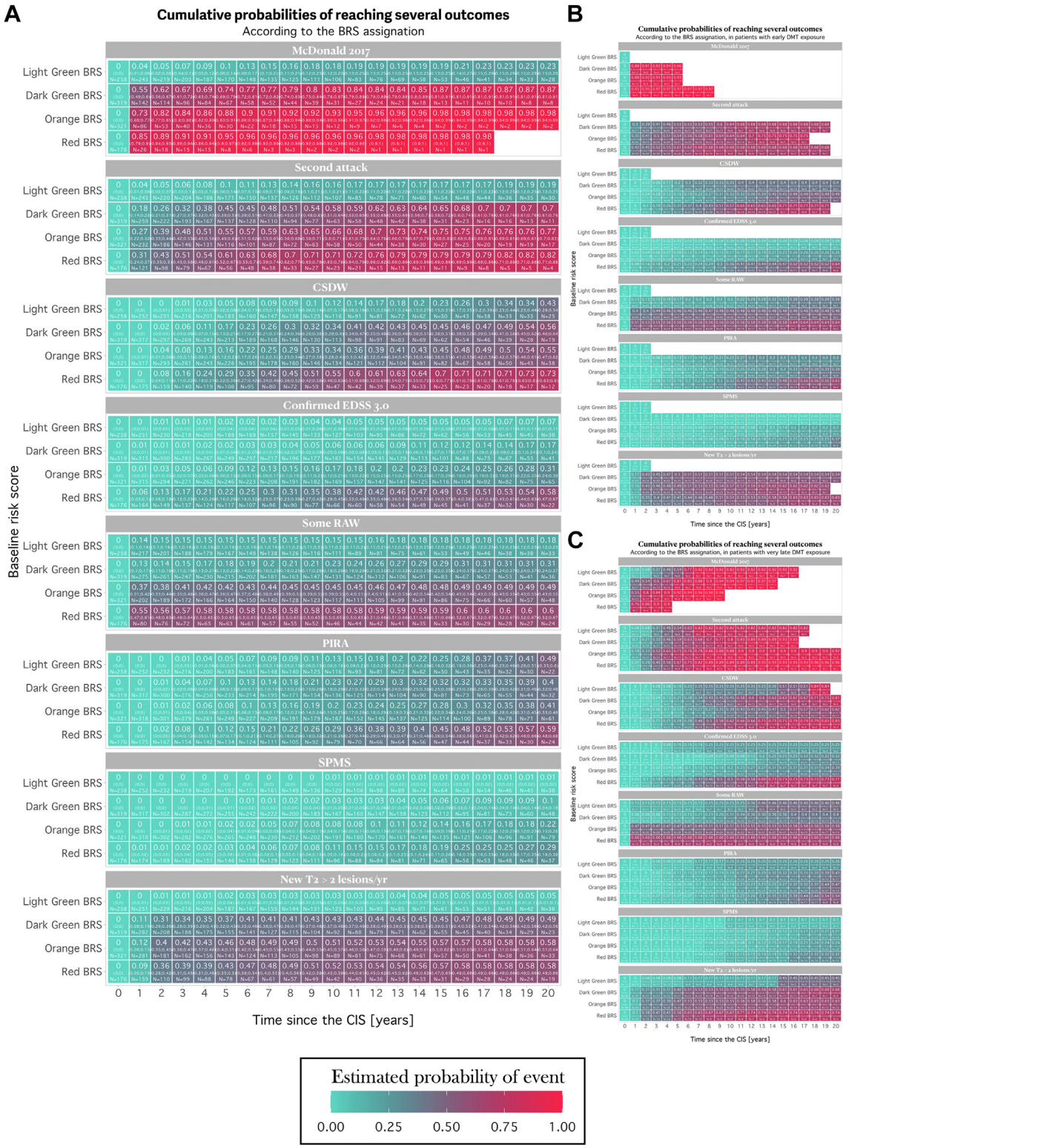

Fig. 3: Heatmaps of binary disease outcomes. This figure shows the estimated probabilities of reaching each one of the clinical outcomes based on the Kaplan Meier models, for each one of the Barcelona-BRS risk groups. **A.** Whole cohort; **B.** Patients with early DMT exposure (within 12 months of CIS); **C.** Patients with late DMT exposure (after 3 years of CIS). *Abbreviations:* BRS: baseline risk score; CIS: clinically isolated syndrome; CSDW: confirmed and sustained disability worsening; DMT: disease-modifying treatment; EDSS: Expanded Disability Status Scale; PIRA: progression independent of relapse activity; RAW: relapse-associated worsening; SPMS: secondary progressive multiple sclerosis.

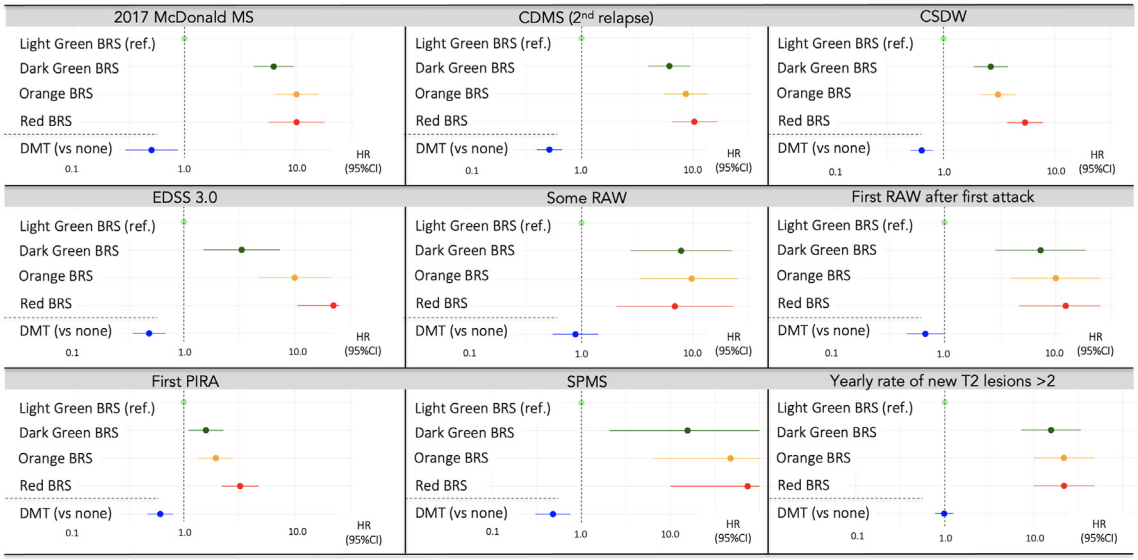

Fig. 4: Light green BRS group is used as reference. Please note the x-axis is displayed in logarithmic scale, for visualisation purposes. Also, the 95% CI have been truncated at 100 to avoid excessively large figure subpanels. Details on HR (95% CI) are shown in Supplementary Table S6. *Abbreviations:* BRS: Baseline Risk Score; CSDW: Confirmed and sustained disability worsening; EDSS: Expanded Disability Status Scale; PIRA: progression independent of relapse activity; RAW: relapse-associated worsening; SPMS: secondary progressive multiple sclerosis.

Previous predictive models for MS[10,11,20–22] often focused on specific aspects of the disease, such as clinical or MRI data, without integrating all possible predictors. Thus, while early models focused on clinical predictors and identified factors like male sex, older age, and non-optic-neuritis attacks as associated with faster disability progression,[9] more recent models have mainly focused on MRI alone or laboratory biomarkers, such as CSF OBs,[23] neurofilament light,[24] glial fibrillary acidic protein,[25] or chitinase 3-like 1,[26] lacking the comprehensiveness needed for practical application. In 2015, the first integrative predictive model for people with a first demyelinating attack was proposed.[11] It included all possible predictors of midterm disability described so far, as long as they could be obtained in the clinic.[11] Nonetheless, this model was not tested against long-term clinical or MRI outcomes and lacked the adaptive nature of the Barcelona-BRS model. More recently, other integrative models have been proposed, using data from the MS BASE dataset[27,28] or the Swedish MS registry,[29] providing important insights into the disease, although these models often lacked MRI information, key for MS prognosis. The Barcelona-BRS model overcomes these limitations by integrating multiple data types into a cohesive, adaptable tool for personalised patient risk assessment.

The BRS groups clearly differed in terms of long-term clinical and MRI outcomes, with the Light Green BRS group always showing the best results, followed by the dark-green, orange, and Red BRS groups. However, the performance of the Barcelona-BRS model seemed to be better for some outcomes than for others, highlighting the need for outcome-specific models in the future.

The BRS groups also differed in most quality of life scores, although for symptoms like anxiety and fatigue the differences were less clear, suggesting they might not be directly related to neurodegeneration or inflammation.[30] MRI results showed consistent differences across groups too, with the Red BRS group showing the most severe MRI outcomes, including T2 lesion volume and brain atrophy, and the Light Green BRS group having the best outcomes. Differences in brain volumetric measures, particularly deep grey matter fraction, align with existing literature linking grey matter atrophy to disability progression.[31]

Looking into the actual models and those factors associated with worse prognosis, we observed that older age at the first attack, male sex, a first attack focused on the spinal cord (or multifocal), a greater disability score at first visit, a greater T2 lesion burden at first attack, especially if there are infratentorial lesions, and the presence of contrast-enhancing lesions were all associated with shorter times to an EDSS of 3.0. So, all these factors were associated with greater chances of being classified as Red BRS, the group with the worst prognosis. These findings are in line with previous literature.[8,11,22,32] In this sense, this study does not reveal new prognostic biomarkers but instead offers a reliable and pragmatic way of combining them in the clinic after a first demyelinating attack. A note of caution must be made, though, since this BRS model should only be

applied to people with a first neurological episode suggestive of MS. Also, since up to a 10% of our patients with normal MRI might be positive for anti-MOG antibodies,[33] we should be aware that a non-deniable percentage of the Light Green BRS group might indeed have an alternative diagnosis to MS, with important therapeutic implications.

When we assessed the effects of DMT exposure (as a time-varying covariate) on long-term outcomes, we found that it was significantly associated with a lower risk of developing the event afterwards, for almost all outcomes, in line with previous studies.[34] Furthermore, when patients with an early DMT exposure were compared to patients with late exposure, those with an early exposure had better clinical outcomes, suggesting a protective effect of early DMT use, as also previously described.[7] Interestingly, the potentially protective effects of DMT exposure in relation to future outcomes appeared to be similar across all BRS groups (with HRs of around 0.5 in all analyses performed—data not shown). Also of importance, even when we accounted for DMT exposure, the differences across BRS groups were always maintained, being the Red BRS group the one with worst prognosis, independently of the outcome assessed. Nonetheless, we need to acknowledge that our cohort started almost 30 years ago, which means that many of our patients were only marginally exposed to DMTs and especially high-efficacy drugs.[5] Therefore, our estimations of DMT effects on long-term outcomes should be taken with caution.

Our external validation confirmed the applicability of the Barcelona-BRS model, even in clinical settings with limited data. The model effectively classified patients into risk groups, and outcomes varied as expected, supporting its generalisability.

The main strengths of our study are the uniquely large and rich cohort of people with a first demyelinating attack included in the analysis, and the innovative approach to build a flexible and generalisable risk assessment tool for patient management. However, there are some limitations which are worth mentioning. First, although our study accounted for DMT exposure through including a time-varying covariate in the models of long-term outcome predictions, other statistical approaches could have been considered. For instance, the inclusion of the interactions between DMT exposure and baseline covariates might have provided more accurate estimations of the BRS model. This and other, possibly more complex approaches able to account for the dynamic nature of DMT exposure and its likely drug-specific therapeutic lag[35] will need to be explored through future research.

Additionally, our model did not include laboratory or MRI biomarkers such as neurofilament light (NfL),[36,37] glial fibrillary acidic protein (GFAP),[36,38] or slowly expanding lesions (SELs),[39] which may become essential as they are incorporated into routine practice. This means that our model could not account for important aspects such as chronic inflammation and neurodegeneration, strongly related to progression and which may be present from disease onset.[1] We anticipate that future revisions of our work will need to integrate emerging biomarkers, potentially by adding a sixth predictive model to the Barcelona-BRS framework. This model would estimate the Weibull coefficients for NfL and GFAP at symptom onset alongside demographic, clinical, and MRI measures. Additionally, if future versions of the model can account for changes in DMT exposure and its long-term effects, the inclusion of dynamic neuroimaging biomarkers, such as SELs, should also be considered.

Another important point to acknowledge is that some of our outcome definitions are continuously subject to revision, especially as new scientific evidence emerges, which may have led to inconsistencies in outcome specifications. Nonetheless, we recognise the significant efforts being made to standardise outcome definitions, which is essential. Furthermore, as clinical practice enables more sensitive lesion detection, including optic nerve lesions, some patients who would have been classified as having a normal brain MRI in our study may instead be identified as having abnormal brain MRI findings. This may probably imply a benignisation of the BRS groups, potentially shifting their risk profile towards a more favorable prognosis. Consequently, further refinements of the BRS model will be necessary as it incorporates prospective data from clinical practice.

Finally, the external validation cohort had a relatively small sample size and limited follow-up. This restricted the assessment of the generalisability of our BRS model, which will need to be investigated in future, larger, and more modern validation cohorts. So, while the results remain promising, particularly due to the use of robust methods to estimate uncertainty in long-term predictions, further studies are warranted.

In conclusion, the Barcelona-BRS model is a flexible, generalisable tool for risk stratification after a first demyelinating attack, based on traditional risk factors such as age, sex, and lesion burden. It provides a ready-to-use solution for clinicians aiming to predict long-term outcomes and personalise treatment plans for MS patients.

Contributors

C.T.: conceptualisation, data curation, formal analysis, software, funding acquisition, investigation, methodology, supervision, validation, writing—original draft, and writing—review & editing. P.C.-M.: data curation, formal analysis, software, visualisation.

S.O.-R.: conceptualisation, investigation, methodology, validation, and writing—review & editing. Á.C.-C.: conceptualisation, investigation, methodology, validation, and writing—review & editing. M.J. A.: data curation and writing—review & editing. H.A.: data curation and writing—review & editing. G.A.: data curation and writing—review & editing. C.A.: data curation and writing—review & editing. R.C.: conceptualisation, investigation, methodology, and writing—review & editing. J.C.: data

curation and writing—review & editing. M.C.: data curation and writing—review & editing. I.G.: data curation and writing—review & editing. L.M.: data curation and writing—review & editing. C.N.: data curation and writing—review & editing. A.P.: data curation and writing—review & editing. D.P.: data curation and writing—review & editing. J.R.: data curation and writing—review & editing; B.R.: data curation and writing—review & editing. E.S.de C.: data curation and writing—review & editing. Á.V.-J.: data curation and writing—review & editing. A.Z.: data curation and writing—review & editing. J.S.G.: data curation and writing—review & editing. À.R.: data curation and writing—review & editing. X.M.: conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, and writing—review & editing. M.T.: conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, supervision, validation, and writing—review & editing.

Additionally, C.T., P.C., X.M., M.T. have directly accessed and verified the underlying data reported in the manuscript.

Data sharing statement

Data will be made available upon reasonable request to the corresponding and senior authors.

Declaration of interests

C. Tur is currently being funded by a Miguel Servet contract, awarded by the Instituto de Salud Carlos III (ISCIII), Ministerio de Ciencia e Innovación de España (CP23/00117). She has also received a 2020 Junior Leader La Caixa Fellowship (fellowship code: LCF/BQ/PI20/11760008), awarded by "la Caixa" Foundation (ID100010434), a 2021 Merck's Award for the Investigation in MS, awarded by Fundación Merck Salud (Spain), 2021 and 2024 Research Project Grants (PI21/01860 and PI24/01277) awarded by the ISCIII, Ministerio de Ciencia e Innovación de España, and a FORTALECE research grant (FORT23/00034) also by the ISCIII, Ministerio de Ciencia e Innovación de España. In 2015, she received an ECTRIMS Post-doctoral Research Fellowship and has received funding from the UK MS Society. She is a member of the Editorial Board of Neurology Journal and Multiple Sclerosis Journal. She has also received honoraria from Roche, Novartis, Merck, Sanofi, Immunic Therapeutics, and Bristol Myers Squibb. She is a steering committee member of the O'HAND trial (Roche) and of the Consensus group on Follow-on DMTs.

P. Carbonell-Mirabent has nothing to disclose.

S. Otero-Romero has received speaking and consulting honoraria from Genzyme, Biogen-Idec, Novartis, Roche, Excemed and MSD; as well as research support from Novartis.

A. Cobo-Calvo has received grant from Instituto de Salud Carlos III, Spain; JR19/00007.

M.J. Arévalo has nothing to disclose.

H. Ariño has nothing to disclose.

G. Arrambide has received compensation for consulting services, speaking honoraria or participation in advisory boards from Merck, Roche, and Horizon Therapeutics; and travel support for scientific meetings from Novartis, Roche, ECTRIMS and EAN. Has received speaking honoraria and consulting services or participation in advisory boards from Sanofi, Merck, Roche and Horizon Therapeutics; travel expenses for scientific meetings from Novartis, Roche, and ECTRIMS.

C. Auger has nothing to disclose.

R. Carvajal is currently being funded by a Research grand from the European Charcot Foundation. In 2023 he received a grand by "Vall d'Hebron Institut de Recerca". In 2021, he received an ECTRIMS Fellowship training performed during 2021–2022. He has also received speaking honoraria from Roche, Novartis, Biogen, Merck and Sanofi.

J. Castilló has nothing to disclose.

M. Comabella has received compensation for consulting services and speaking honoraria from Bayer Schering Pharma, Merck Serono, Biogen-Idec, Teva Pharmaceuticals, Sanofi-Aventis, Genzyme, and Novartis.

I. Galán has nothing to disclose.

L. Midaglia has nothing to disclose.

C. Nos has nothing to disclose.

A. Pappolla has received funding travel from Roche and speaking honoraria from Novartis. He performed an ECTRIMS Clinical Training Fellowship program during 2021 and is currently performing an MSIF-ARSEP Fellowship program.

D. Pareto holds a research grant from the Instituto Carlos III (PI22/01709), co-funded by the European Union.

J. Río has received speaking honoraria and personal compensation for participating on Advisory Boards from Biogen-Idec, Genzyme, Janssen, Merck- Serono, Novartis, Teva, Roche, and Sanofi-Aventis.

B. Rodríguez-Acevedo has received speaking honoraria from Merck and honoraria for consulting services from Novartis.

E. Saez de Gordoa has nothing to disclose.

A. Vidal-Jordana has received support has received support for contracts Juan Rodes (JR16/00024) and from Fondo de Investigación en Salud (PI17/02162 and PI22/01589) from Instituto de Salud Carlos III, Spain, and has engaged in consulting and/or participated as speaker in events organized by Roche, Novartis, and Merck.

A. Zabalza has received a Rio Hortega grant, from the Instituto de Salud Carlos III, Spain (CM22/00237) and received travel expenses for scientific meetings from Biogen-Idec, Merck Serono and Novartis; speaking honoraria from Eisai; and a study grant from Novartis.

J. Sastre-Garriga serves as co-Editor for Europe on the editorial board of Multiple Sclerosis Journal and as Editor-in-Chief in Revista de Neurología, receives research support from Fondo de Investigaciones Sanitarias (19/950) and has served as a consultant/speaker for Biogen, Celgene/Bristol Meyers Squibb, Sanofi, Novartis and Merck.

À. Rovira serves on scientific advisory boards for Novartis, Sanofi-Genzyme, Synthetic MR, Roche, Biogen, and OLEA Medical; has received speaker honoraria from Bayer, SanofiGenzyme, Merck-Serono, Teva Pharmaceutical Industries Ltd, Novartis, Roche, and Biogen; and is CMO and co-founder of TensorMedical.

X. Montalban has received speaking honoraria and travel expenses for participation in scientific meetings, has been a steering committee member of clinical trials or participated in advisory boards of clinical trials in the past years with Abbvie, Actelion, Alexion, Biogen, Bristol-Myers Squibb/Celgene, EMD Serono, Genzyme, Hoffmann-La Roche, Immunic, Janssen Pharmaceuticals, Medday, Merck, Mylan, Nervgen, Novartis, Sandoz, Sanofi-Genzyme, Teva Pharmaceutical, TG Therapeutics, Excemed, MSIF and NMSS.

M. Tintoré has received compensation for consulting services, speaking honoraria and research support from Almirall, Bayer Schering Pharma, Biogen-Idec, Genzyme, Immunic Therapeutics, Janssen, Merck-Serono, Novartis, Roche, Sanofi-Aventis, Viela Bioand Teva Pharmaceuticals. Data Safety Monitoring Board for Parexel and UCB Biopharma, Relapse Adjudication Committee for IMCYSE SA.

Acknowledgements

The authors acknowledge the use of MS PATHS data and would like to thank the MS PATHS network as a group. The MS PATHS project was funded by Biogen (Cambridge, MA, USA). The authors thank all of the patients who participated in the MS PATHS initiative as well as the patients seen at Cemcat throughout the years and their families. During the preparation of this work the authors used ChatGPT4o in order to improve the grammar and orthography of the paper. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication. The authors had full editorial control of the paper and provided their final approval of all content.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lanepe.2025.101302.

References

1 Kuhlmann T, Moccia M, Coetzee T, et al. Multiple sclerosis progression: time for a new mechanism-driven framework. *Lancet Neurol*. 2023;22:78–88.

2 Thompson AJ, Baranzini SE, Geurts J, Hemmer B, Ciccarelli O. Multiple sclerosis. *Lancet*. 2018;391:1622–1636.
3 Jakimovski D, Bittner S, Zivadinov R, et al. Multiple sclerosis. *Lancet*. 2024;403:183–202.
4 Kappos L, Wolinsky JS, Giovannoni G, et al. Contribution of relapse-independent progression vs relapse-associated worsening to overall confirmed disability accumulation in typical relapsing multiple sclerosis in a pooled analysis of 2 randomized clinical trials. *JAMA Neurol*. 2020;77:1132–1140.
5 Tur C, Carbonell-Mirabent P, Cobo-Calvo Á, et al. Association of early progression independent of relapse activity with long-term disability after a first demyelinating event in multiple sclerosis. *JAMA Neurol*. 2023;80:151–160.
6 Lublin FD, Häring DA, Ganjgahi H, et al. How patients with multiple sclerosis acquire disability. *Brain*. 2022;145:3147–3161.
7 Cobo-Calvo A, Tur C, Otero-Romero S, et al. Association of very early treatment initiation with the risk of long-term disability in patients with a first demyelinating event. *Neurology*. 2023;101:e1280–e1292.
8 Brex PA, Ciccarelli O, O'Riordan JI, Sailer M, Thompson AJ, Miller DH. A longitudinal study of abnormalities on MRI and disability from multiple sclerosis. *N Engl J Med*. 2002;346:158–164.
9 Confavreux C, Vukusic S, Adeleine P. Early clinical predictors and progression of irreversible disability in multiple sclerosis: an amnesic process. *Brain*. 2003;126:770–782.
10 Río J, Nos C, Tintoré M, et al. Defining the response to interferon-beta in relapsing-remitting multiple sclerosis patients. *Ann Neurol*. 2006;59:344–352.
11 Tintore M, Rovira À, Río J, et al. Defining high, medium and low impact prognostic factors for developing multiple sclerosis. *Brain*. 2015;138:1863–1874.
12 Malpas CB, Manouchehrinia A, Sharmin S, et al. Early clinical markers of aggressive multiple sclerosis. *Brain*. 2020;143:1400–1413.
13 Mallucci G, Ferraro OE, Trojano M, et al. Early prediction of unfavorable evolution after a first clinical episode suggestive of multiple sclerosis: the EUMUS score. *J Neurol*. 2024;271:3496–3505.
14 Mowry EM, Bermel RA, Williams JR, et al. Harnessing real-world data to inform decision-making: multiple sclerosis partners advancing technology and health solutions (MS PATHS). *Front Neurol*. 2020;11:632.
15 Kurtzke JF. Rating neurologic impairment in multiple sclerosis. *Neurology*. 1983;33:1444.
16 Learmonth YC, Motl RW, Sandroff BM, Pula JH, Cadavid D. Validation of patient determined disease steps (PDDS) scale scores in persons with multiple sclerosis. *BMC Neurol*. 2013;13:37.
17 Thompson AJ, Banwell BL, Barkhof F, et al. Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*. 2018;17:162–173.
18 Lorscheider J, Buzzard K, Jokubaitis V, et al. Defining secondary progressive multiple sclerosis. *Brain*. 2016;139:2395–2405.
19 Río J, Comabella M, Montalban X. Multiple sclerosis: current treatment algorithms. *Curr Opin Neurol*. 2011;24:230–237.
20 Rocca MA, Valsasina P, Meani A, et al. Spinal cord lesions and brain grey matter atrophy independently predict clinical worsening in definite multiple sclerosis: a 5-year, multicentre study. *J Neurol Neurosurg Psychiatry*. 2023;94:10–18.
21 Sormani MP, Gasperini C, Romeo M, et al. Assessing response to interferon-β in a multicenter dataset of patients with MS. *Neurology*. 2016;87:134–140.
22 Brownlee WJ, Altmann DR, Prados F, et al. Early imaging predictors of long-term outcomes in relapse-onset multiple sclerosis. *Brain*. 2019;142:2276–2287.
23 Arrambide G, Tintore M, Espejo C, et al. The value of oligoclonal bands in the multiple sclerosis diagnostic criteria. *Brain*. 2018;141:1075–1084.
24 Arrambide G, Espejo C, Eixarch H, et al. Neurofilament light chain level is a weak risk factor for the development of MS. *Neurology*. 2016;87:1076–1084.
25 Martínez MAM, Olsson B, Bau L, et al. Glial and neuronal markers in cerebrospinal fluid predict progression in multiple sclerosis. *Mult Scler*. 2015;21:550–561.
26 Cantó E, Tintoré M, Villar LM, et al. Chitinase 3-like 1: prognostic biomarker in clinically isolated syndromes. *Brain*. 2015;138:918–931.
27 Kalincik T, Manouchehrinia A, Sobisek L, et al. Towards personalized therapy for multiple sclerosis: prediction of individual treatment response. *Brain*. 2017;140:2426–2443.
28 Moradi N, Sharmin S, Malpas CB, et al. External validation of a clinical prediction model in multiple sclerosis. *Mult Scler*. 2023;29:261–269.
29 Hillert J, Stawiarz L. The Swedish MS registry – clinical support tool and scientific resource. *Acta Neurol Scand*. 2015;132:11–19.
30 Rossi S, Studer V, Motta C, et al. Neuroinflammation drives anxiety and depression in relapsing-remitting multiple sclerosis. *Neurology*. 2017;89:1338–1347.
31 Eshaghi A, Prados F, Brownlee WJ, et al. Deep gray matter volume loss drives disability worsening in multiple sclerosis. *Ann Neurol*. 2018;83:210–222.
32 Fisniku LK, Brex PA, Altmann DR, et al. Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain*. 2008;131:808–817.
33 Villacieros-Álvarez J, Espejo C, Arrambide G, et al. Myelin oligodendrocyte glycoprotein antibodies in adults with a first demyelinating event suggestive of multiple sclerosis. *Ann Neurol*. 2023. https://doi.org/10.1002/ana.26793.
34 Tur C, Moccia M, Barkhof F, et al. Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. *Nat Rev Neurol*. 2018;14:75–93.
35 Montobbio N, Bovis F, Signori A, Ponzano M, Schiavetti I, Sormani MP. Therapeutic lag: is treatment effect delayed in progressive MS? *Mult Scler*. 2024;30:843–846.
36 Benkert P, Maleska Maceski A, Schaedelin S, et al. Serum glial fibrillary acidic protein and neurofilament light chain levels reflect different mechanisms of disease progression under B-cell depleting treatment in multiple sclerosis. *Ann Neurol*. 2024. https://doi.org/10.1002/ana.27096.
37 Abdelhak A, Benkert P, Schaedelin S, et al. Neurofilament light chain elevation and disability progression in multiple sclerosis. *JAMA Neurol*. 2023;80:1317.
38 Monreal E, Fernández-Velasco JI, Álvarez-Lafuente R, et al. Serum biomarkers at disease onset for personalized therapy in multiple sclerosis. *Brain*. 2024. https://doi.org/10.1093/brain/awae260.
39 Bagnato F, Sati P, Hemond CC, et al. Imaging chronic active lesions in multiple sclerosis: a consensus statement. *Brain*. 2024;147:2913–2933.