doi:10.1093/brain/awaa081 BRAIN 2020: 143; 1400–1413 | 1400

BRAIN
A JOURNAL OF NEUROLOGY

Early clinical markers of aggressive multiple sclerosis

Charles B. Malpas,[1,2] Ali Manouchehrinia,[3] Sifat Sharmin,[1,2] Izanne Roos,[1,2] Dana Horakova,[4] Eva Kubala Havrdova,[4] Maria Trojano,[5] Guillermo Izquierdo,[6] Sara Eichau,[6] Roberto Bergamaschi,[7] Patrizia Sola,[8] Diana Ferraro,[8,9] Alessandra Lugaresi,[10,11] Alexandre Prat,[12] Marc Girard,[12] Pierre Duquette,[12] Pierre Grammond,[13] Francois Grand'Maison,[14] Serkan Ozakbas,[15] Vincent Van Pesch,[16,17] Franco Granella,[18] Raymond Hupperts,[19] Eugenio Pucci,[20] Cavit Boz,[21] Youssef Sidhom,[22] Riadh Gouider,[23] Daniele Spitaleri,[24] Aysun Soysal,[25] Thor Petersen,[26] Freek Verheul,[27] Rana Karabudak,[28] Recai Turkoglu,[29] Cristina Ramo-Tello,[30] Murat Terzi,[31] Edgardo Cristiano,[32] Mark Slee,[33] Pamela McCombe,[34,35] Richard Macdonell,[36] Yara Fragoso,[37] Javier Olascoaga,[38] Ayse Altintas,[39] Tomas Olsson,[40] Helmut Butzkueven,[41,42,43] Jan Hillert[40] and Tomas Kalincik[1,2] on behalf of the MSBase Study Group

Patients with the 'aggressive' form of multiple sclerosis accrue disability at an accelerated rate, typically reaching Expanded Disability Status Score (EDSS) $\geq$ 6 within 10 years of symptom onset. Several clinicodemographic factors have been associated with aggressive multiple sclerosis, but less research has focused on clinical markers that are present in the first year of disease. The development of early predictive models of aggressive multiple sclerosis is essential to optimize treatment in this multiple sclerosis subtype. We evaluated whether patients who will develop aggressive multiple sclerosis can be identified based on early clinical markers. We then replicated this analysis in an independent cohort. Patient data were obtained from the MSBase observational study. Inclusion criteria were (i) first recorded disability score (EDSS) within 12 months of symptom onset; (ii) at least two recorded EDSS scores; and (iii) at least 10 years of observation time, based on time of last recorded EDSS score. Patients were classified as having 'aggressive multiple sclerosis' if all of the following criteria were met: (i) EDSS $\geq$ 6 reached within 10 years of symptom onset; (ii) EDSS $\geq$ 6 confirmed and sustained over $\geqslant$6 months; and (iii) EDSS $\geq$ 6 sustained until the end of follow-up. Clinical predictors included patient variables (sex, age at onset, baseline EDSS, disease duration at first visit) and recorded relapses in the first 12 months since disease onset (count, pyramidal signs, bowel-bladder symptoms, cerebellar signs, incomplete relapse recovery, steroid administration, hospitalization). Predictors were evaluated using Bayesian model averaging. Independent validation was performed using data from the Swedish Multiple Sclerosis Registry. Of the 2403 patients identified, 145 were classified as having aggressive multiple sclerosis (6%). Bayesian model averaging identified three statistical predictors: age $>$ 35 at symptom onset, EDSS $\geq$ 3 in the first year, and the presence of pyramidal signs in the first year. This model significantly predicted aggressive multiple sclerosis [area under the curve (AUC) = 0.80, 95% confidence intervals (CIs): 0.75, 0.84, positive predictive value = 0.15, negative predictive value = 0.98]. The presence of all three signs was strongly predictive, with 32% of such patients meeting aggressive disease criteria. The absence of all three signs was associated with a 1.4% risk. Of the 556 eligible patients in the Swedish Multiple Sclerosis Registry cohort, 34 (6%) met criteria for aggressive multiple sclerosis. The combination of all three signs was also predictive in this cohort (AUC = 0.75, 95% CIs: 0.66, 0.84, positive predictive value = 0.15, negative predictive value = 0.97). Taken together, these findings suggest that older age at symptom onset, greater disability during the first year, and pyramidal signs in the first year are early indicators of aggressive multiple sclerosis.

Received August 22, 2019. Revised January 21, 2020. Accepted February 5, 2020. Advance access publication May 9, 2020

1 CORe Unit, Department of Medicine, University of Melbourne, Melbourne, Australia
2 Department of Neurology, Royal Melbourne Hospital, Melbourne, Australia
3 Centre for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden
4 Department of Neurology and Center of Clinical Neuroscience, First Faculty of Medicine, Charles University in Prague and General University Hospital, Prague, Czech Republic
5 Department of Basic Medical Sciences, Neuroscience and Sense Organs, University of Bari, Bari, Italy
6 Hospital Universitario Virgen Macarena, Sevilla, Spain
7 IRCCS Mondino Foundation, Pavia, Italy
8 Department of Neuroscience, Azienda Ospedaliera Universitaria, Modena, Italy
9 Department of Biomedical, Metabolic and Neurosciences, University of Modena and Reggio Emilia, Modena, Italy
10 Department of Biomedical and Neuromotor Science, University of Bologna, Bologna, Italy
11 IRCCS Istituto delle Scienze Neurologiche di Bologna, Bologna, Italy
12 CHUM and Universite de Montreal, Montreal, Canada
13 CISSS de Chaudière-Appalaches, Levis, Canada
14 Neuro Rive-Sud, Quebec, Canada
15 Dokuz Eylul University, Konak/Izmir, Turkey
16 Cliniques Universitaires Saint-Luc, Brussels, Belgium
17 Université Catholique de Louvain, Brussels, Belgium
18 Department of Medicine and Surgery, University of Parma, Parma, Italy
19 Zuyderland Ziekenhuis, Sittard, The Netherlands
20 UOC Neurologia, Azienda Sanitaria Unica Regionale Marche - AV3, Macerata, Italy
21 KTU Medical Faculty Farabi Hospital, Trabzon, Turkey
22 Department of Neurology, Razi Hospital, Manouba, Tunisia
23 Department of Neurology, Razi Hospital, LR 18SP03, Clinical Investigation Center Neurosciences and Mental Health, Faculty of Medicine University Tunis El Manar, Tunis, Tunisia
24 Azienda Ospedaliera di Rilievo Nazionale San Giuseppe Moscati Avellino, Avellino, Italy
25 Bakirkoy Education and Research Hospital for Psychiatric and Neurological Diseases, Istanbul, Turkey
26 Kommunehospitalet, Arhus C, Denmark
27 Groene Hart Ziekenhuis, Gouda, The Netherlands
28 Hacettepe University, Ankara, Turkey
29 Haydarpasa Numune Training and Research Hospital, Istanbul, Turkey
30 Hospital Germans Trias i Pujol, Badalona, Spain
31 Medical Faculty, 19 Mayis University, Samsun, Turkey
32 Hospital Italiano, Buenos Aires, Argentina
33 Flinders University, Adelaide, Australia
34 University of Queensland, Brisbane, Australia
35 Royal Brisbane and Women's Hospital, Brisbane, Australia
36 Austin Health, Melbourne, Australia
37 Universidade Metropolitana de Santos, Santos, Brazil
38 Instituto de Investigación Sanitaria Biodonostia, Hospital Universitario Donostia, San Sebastián, Spain
39 Koc University, School of Medicine, Department of Neurology, Istanbul, Turkey
40 Department of Clinical Neuroscience, Karolinska Institutet, Sweden
41 Central Clinical School, Monash University, Melbourne, Australia
42 Department of Neurology, The Alfred Hospital, Melbourne, Australia
43 Department of Neurology, Box Hill Hospital, Monash University, Melbourne, Australia

Correspondence to: Tomas Kalincik
Clinical Outcomes Research Unit (CORe)
Department of Medicine, Royal Melbourne Hospital
The University of Melbourne
Melbourne, Australia
E-mail: tomas.kalincik@unimelb.edu.au

Keywords: multiple sclerosis; prediction; disability; aggressive disease; precision medicine

Abbreviations: BMA = Bayesian model averaging; EDSS = Expanded Disability Status Score; PIP = posterior inclusion probability; ROC = receiver operating characteristic

Introduction

Relapsing remitting multiple sclerosis is associated with the accrual of disability, typically due to incomplete recovery from relapses (Jokubaitis *et al.*, 2016). A subset of patients accrue disability at an accelerated rate. While the rate of disability accrual in multiple sclerosis can represent a continuum, the European Committee for Treatment and Research in Multiple Sclerosis recognizes that the rapidly progressing phenotype warrants a distinct clinical definition (European Committee for Treatment and Research in Multiple Sclerosis, 2018). An early attempt to define this group of patients was 'malignant multiple sclerosis', which was defined as 'disease with a rapid progressive course, leading to significant disability in multiple neurologic systems or death in a relatively short time after disease onset' (Lublin and Reingold, 1996). It is estimated that malignant multiple sclerosis describes ~5–10% of patients (Gholipour *et al.*, 2011).

More recently, the term 'aggressive multiple sclerosis' has gained increased acceptance (Gholipour *et al.*, 2011). There is, however, no universally accepted definition of aggressive multiple sclerosis and a number of different criteria have been used. For example, in terms of the rate of disability accrual, Rush *et al.* (2015) suggest that reaching an Expanded Disability Status Scale (EDSS) of 4 within 5 years of symptom onset is a core feature of the phenotype. Some researchers used more severe criteria, requiring an EDSS of $\geqslant 6$ within 5 years (Menon *et al.*, 2017), while others have required a gain of two or more EDSS points over a 2-year period (Scott *et al.*, 2013). A recent workshop arranged by the European Committee for Treatment and Research in Multiple Sclerosis on the topic of aggressive multiple sclerosis was unable to come to a consensus definition (Iacobaeus, submitted for publication). The diverse definitions used to describe this phenotype render interpretation of the literature challenging. In this study, we chose to use a definition that was recently applied in the Barcelona cohort (Tintore *et al.*, 2019).

The determinants and clinical predictors of aggressive multiple sclerosis are not well understood; however, a number of risk factors of more rapid disability accrual have been identified (for a recent review see Rush *et al.*, 2015). Patient characteristics, such as being older at symptom onset or of male sex, have been recognized (Kantarci *et al.*, 1998; Held *et al.*, 2005; Tremlett *et al.*, 2006; Ribbons *et al.*, 2015). Several relapse characteristics have also been implicated, including: the presence of multifocal relapses (Kraft *et al.*, 1981; Wolfson and Confavreux, 1987; Amato *et al.*, 1999; Bergamaschi *et al.*, 2001); partial or incomplete recovery from relapse (Trojano *et al.*, 1995; Scott and Schramke, 2010); greater frequency of relapses (Weinshenker *et al.*, 1989; Confavreux *et al.*, 2003; Ebers, 2005); and reduced time between relapses (Weinshenker *et al.*, 1989; Trojano *et al.*, 1995). Relapses associated with motor, cerebellar, bowel/bladder, and cognitive signs have also been implicated (Citterio *et al.*, 1989; Weinshenker *et al.*, 1989, 1996; Phadke, 1990; Runmarker and Andersen, 1993; Amato *et al.*, 1999; Zarei *et al.*, 2003; Langer-Gould *et al.*, 2006; Kalincik, 2015). Neuroimaging signs have also been investigated, and the presence of high T_2 lesion burden (Filippi *et al.*, 1994; O'Riordan *et al.*, 1998; Brex *et al.*, 2002; Rudick *et al.*, 2006; Uher *et al.*, 2017), multiple gadolinium-enhancing lesions (Kappos *et al.*, 1999; Rudick *et al.*, 2006), T_1 lesions (Tomassini *et al.*, 2006), early atrophy (Kornek and Lassmann, 2003; Lukas *et al.*, 2010; Vaneckova *et al.*, 2012), and infratentorial lesions (Sailer *et al.*, 1999) have been shown to raise the risk of more rapid disability accrual.

While a number of risk factors have been described, the early clinical identification of patients at risk of aggressive multiple sclerosis presents a significant gap in current understanding. There is a paucity of large-scale observational studies that have interrogated multiple markers of aggressive multiple sclerosis simultaneously. These studies have shed light on the predictors of disease outcomes at the level of cohorts. Critically, new evidence has demonstrated that early second-line therapy markedly improves disability outcomes over the long-term (Fernández, 2017; Brown *et al.*, 2019). However, identification of individuals at high risk who are most likely to benefit from proactive treatment strategies remains an unmet need. The aim of this study was to define early clinical markers that identify patients who will later meet criteria for aggressive multiple sclerosis. To maximize clinical translation, we focused on markers that are accessible to clinicians within the first 12 months following symptom onset. We focused on estimating the risk of aggressive disease in individual patients, using Bayesian methods including Bayesian model averaging (BMA), and have conducted a validation in an independent cohort.

Materials and methods

Participants

For the primary analysis, patient data were obtained from the international MSBase cohort study (WHO ICTRP ID: ACTRN12605000455662) (Butzkueven *et al.*, 2006). The study was approved by the Melbourne Health Research Ethics Committee and site-specific ethics committees where relevant. All participants provided informed consent. Data from 58 589 patients attending 139 clinical centres in 34 countries were extracted from the MSBase registry in March 2018. Inclusion criteria were: (i) a diagnosis of clinically definite relapse-onset multiple sclerosis (Polman *et al.*, 2011); (ii) age at onset $\geqslant 18$ years; (iii) first EDSS recorded within 12 months of symptom onset; (iv) at least two recorded EDSS scores within 10 years of symptom onset; and (v) at least 10 years of observation time. All data were entered into the iMed patient record system or the MSBase online system as part of routine clinical practice. Data quality assurance procedures were performed as described elsewhere (Kalincik *et al.*, 2017).

A second cohort of patients from the Swedish Multiple Sclerosis Registry (Hillert and Stawiarz, 2015) was used as an independent validation sample. Data from 19 520 patients were

extracted from the registry in April 2018. The same inclusion criteria were applied as described above.

Definition of aggressive multiple sclerosis

Patients were classified as having aggressive multiple sclerosis if they met all three of the following criteria: (i) EDSS $\geqslant 6$ was reached within 10 years of symptom onset; (ii) EDSS $\geqslant 6$ was confirmed and sustained over $\geqslant 6$ months; and (iii) EDSS $\geqslant 6$ was sustained until the end of follow-up ($\geqslant 10$ years).

Early clinical markers

Based on a survey of the relevant literature described above, a set of clinically translatable predictor variables was chosen. Patient characteristics included sex and age at symptom onset. The following predictors were included based on observation over the first 12 months following symptom onset: median EDSS, number of relapses, number of severe relapses (relapses that significantly affected activities of daily living or required hospitalization), administration of steroids, hospitalization (both used as a proxy for severe relapses), incomplete recovery after a relapse (clinician-rated persistence of symptoms of a relapse), pyramidal signs, bowel-bladder signs, cognitive signs (presence of at least mild decrease in mentation as rated by the treating neurologist), and cerebellar signs. Several nuisance variables were included in the model, which were variables that were important for model adjustment but were not of primary interest in predicting the outcome because they would not be readily available to the clinician early in the disease. These nuisance variables were: disease duration at first visit (maximum of 365 days), total observation time (computed as the time from first visit with recorded EDSS score to the last visit with recorded EDSS score; minimum of 10 years), proportion of time spent on first-line (glatiramer acetate, interferon, dimethyl fumarate, teriflunomide) or second-line (mitoxantrone, natalizumab, cladribine, fingolimod, alemtuzumab, autologous stem cell transplantation, daclizumab, ocrelizumab, ofatumumab, rituximab, siponimod) therapy (Tramacere *et al.*, 2015) over the first year since symptom onset, and proportion of time spent on first- and second-line therapy over the first 10 years since symptom onset. MRI findings in the first year since symptom onset (presence of gadolinium enhancing, infratentorial, juxtatentorial, or periventricular lesions) and oligoclonal band results (positive or negative) were available for a subset of patients.

Statistical analyses

All statistical analyses were performed in R (Version 3.4.4) (R Core Team, 2019). Generalized linear models were used to determine the statistical predictors of aggressive disease status. Aggressive disease was entered as the binary dependent variable and a binomial response distribution with logistic link function was chosen. The following baseline variables were entered as predictors: gender, age at symptom onset. The following events within the first year of symptom onset were also entered: median EDSS, any hospitalization associated with a relapse, any treatment with steroids, number of severe relapses, number of any relapses, pyramidal signs, bowel/bladder signs, cerebellar signs, incomplete recovery from a relapse. The following nuisance variables were also entered: disease duration at first visit, disease duration at last visit (total observation time), proportion of time over the first year on first or second-line therapy, proportion of time over the 10-year observation period on first or second-line therapy. Only patients with complete data were included.

The use of these 17 predictors of interest resulted in $2^{17} = 131\,072$ possible statistical models (i.e. different combinations of predictors). From a practical standpoint, this raises the question of which clinical markers should be chosen for statistical modelling. While including all markers in a single model might seem a feasible solution, this assumes that all of these markers are included in the 'true' population model. Similarly, selecting a subset of variables also assumes that these are the only markers that are predictive of multiple sclerosis on the population. The selection of variables represents a well understood source of model selection bias, which can lead to misleading results, and it is difficult to know whether a particular set of predictors represents the 'true' model (Miller, 2002; Lukacs *et al.*, 2010). To mitigate this problem, we used BMA (Raftery *et al.*, 1997), a validated statistical technique that allows for all potential predictors to be examined in the model in parallel, while simultaneously considering all possible formulations of the model. The overall statistical solution is composed of all possible combinations of predictors weighted by model fit. As the best fitting combinations contribute more to the final model, and the worst fitting models contribute less, this naturally deals with the problem of model selection without having to impose model selection bias.

Briefly, for p predictors, BMA estimates all possible models corresponding to the 2^p possible combinations of predictors. The fit of each model was then evaluated using the log of the posterior odds. Parameter estimates were averaged over all models, weighted for the fit of each model, which inherently adjusted for the uncertainty associated with model selection bias.

The importance of each predictor was evaluated by the posterior inclusion probability (PIP) value, which was the probability that the parameter is not zero given the data. More formally, this is given as $P(B \neq 0|D)$ where B is the parameter estimate for the predictor and D are the data. This value was computed from the posterior distribution of B, which is given as:

$$P(B \neq 0|Y) = \sum_{k=1}^{k} P(B|M_k, D)\ P(M_k|D) \quad (1)$$

Where B is the parameter estimate for the predictor, D are the data, and mean_1 to mean_k are the models considered. This therefore represented the average of the posterior distribution for each parameter averaged over all models and weighted by the posterior probability (Hoeting *et al.*, 1999). Predictors with PIP values > 0.5 were considered 'important'. Predictions of aggressive disease status for each patient were computed using the posterior predictive distribution averaged across all models. Two sensitivity analyses were conducted. First, Bayesian generalized linear mixed models were computed to investigate the importance of modelling a random intercept across clinical sites. Second, all BMA analyses were repeated using a subgroup of patients who had an EDSS < 6 at first visit.

Receiver operating characteristic (ROC) curves were then computed for the combined model, the 'reduced' model containing important predictors only (also known as the median probability model), and for each of the important continuous predictors separately. The optimal cut-off probability was

determined using Youden's method (Youden, 1950). Key analyses were repeated using conventional frequentist generalized linear models, which were estimated using iteratively reweighted least squares. Robust standard errors were computed via heteroscedasticity-consistent sandwich estimation. Confidence intervals were computed based on these standard errors and are presented at the 95% level. Robust *P*-values were also computed, with the critical α set at 0.05.

BMA was performed using the Bayesian adaptive sampling package (Clyde *et al.*, 2011). The intercept was included in all models. The robust mixture of g-priors defined by Bayarri *et al.*. (2012) was placed on each model coefficient. A beta-binomial prior was placed on each model with parameters $\alpha = 1$ and $\beta = 1$. Bayesian adaptive sampling without replacement was used with uniform sampling probabilities specified across all parameters (Clyde *et al.*, 2011). Bayesian generalized linear mixed models were estimated using the *brms* package to evaluate the need for random terms nested within clinical site (Bürkner, 2017, 2018). ROC curves were computed using the *pROC* package (Robin *et al.*, 2011). Frequentist robust standard errors were computed using the *sandwich* package (Zeileis, 2004, 2006). The best model identified by BMA was replicated in the Swedish Multiple Sclerosis Registry cohort. Only the clinical predictors most strongly supported by the evidence were used for replication, with beta coefficients taken from the original model (rather than being re-estimated).

Data availability

Data were obtained from the international MSBase cohort study. Information regarding data availability can be obtained at https://www.msbase.org/.

Results

Sample characteristics

Data from 9946 patients who met the criteria of a minimum of 10 years observation time, as defined by recorded EDSS scores, were extracted from MSBase. Of these, 7438 patients were excluded because the first visit occurred >12 months after symptom onset. Of the remaining 2508 patients, 17 patients were excluded because of a diagnosis of neuromyelitis optica and a further 88 were excluded due to a diagnosis of primary progressive multiple sclerosis. The final sample comprised 2403 patients, of whom 145 (6%) met criteria for aggressive disease. The clinicodemographic characteristics of the final sample were comparable to the total group of patients who met the initial follow-up and EDSS record criteria (excluding neuromyelitis optica and primary progressive multiple sclerosis patients; shown in Supplementary Table 5). The average time from symptom onset to meeting aggressive disease criteria was 6.05 years [standard deviation (SD) = 2.79, range = 0–9.89 years]. As shown in Table 1, patients with aggressive disease tended to be older at symptom onset, have higher median EDSS during the first year, and a greater number of relapses in the first year. At a descriptive level of analysis, they were also more likely to experience hospitalization, incomplete recovery from relapse, cerebellar signs, bowel/bladder signs, and pyramidal signs within the first year of symptom onset. The time between symptom onset and initiation of first disease-modifying therapy was similar in both groups.

Predictors of aggressive disease

BMA revealed several important predictors of aggressive disease. The top 20 models, sorted by posterior probability, are showing in Fig. 1. The PIPs are shown in Fig. 2. The strongest evidence was observed for age at symptom onset and first year EDSS (PIP = 1). The presence of pyramidal signs was also strongly supported (PIP = 0.75), as was the proportion of time spent on second-line treatment over 10 years (PIP = 1.00). As shown in Table 2, the remaining predictors had PIPs < 0.5. When averaged across the whole model space, the standardized parameter estimates for age of symptom onset, first year EDSS, and time spent on second-line therapy were significant independent predictors of aggressive disease, with their respective confidence intervals not capturing zero. The estimate for pyramidal signs, however, was less precise and included 0 as a plausible value at the 95% level. Taken in the context of the PIP value > 0.5, this indicates that this predictor is important to include in any well-fitting model, but there is considerable uncertainty regarding the specific magnitude of this relationship.

MRI, oligoclonal bands, and initial presentation

The same Bayesian model averaging approach was repeated in the subgroup of patients with information about cerebral MRI ($n = 359$) and CSF oligoclonal bands ($n = 1169$) available. For the MRI data, the presence of gadolinium enhancing, infratentorial, juxtatentorial, or periventricular lesions were added to the full set of predictors as described above. For the oligoclonal band data, positive oligoclonal band status was added to the full set of predictors. All MRI markers had PIPs < 0.5 (range 0.01–0.04), which suggested they were not important predictors in the model, within the constraints of the limited cohort. The PIP for oligoclonal bands positive status was 0.07, which also indicates that this was not an important predictor of aggressive disease. The full BMA model was also run with topology of initial presentation included as predictors (supratentorial, optic pathways, brainstem, and spinal cord). None of these predictors emerged as important for predicting aggressive disease status, with all PIPs < 0.5 (range 0.01–0.03). Model output for these analyses is included in the Supplementary material, section A.

Sensitivity analyses and assumption checks

The first sensitivity analysis was performed to investigate whether the clustering of cases within clinical sites should

Table 1 Sample characteristics

Variable	Total (*n* = 2403)	Non-aggressive disease (*n* = 2258)	Aggressive disease (*n* = 145)
Baseline variables			
Age at symptom onset	31.72 (8.91)	31.3 (8.61)	38.4 (10.8)
Female, *n* (%)	1712 (71.2)	1613 (71.43)	99 (68.28)
Disease duration at first visit, days	119.90 (110.0)	118.00 (110.00)	143.00 (113.00)
First year EDSS	1.78 (1.26)	1.69 (1.16)	3.11 (1.83)
Events within first 12 months			
Number of relapses	0.74 (0.93)	0.72 (0.92)	1.06 (0.96)
Number of severe relapses	0.12 (0.40)	0.20 (0.40)	0.17 (0.44)
Treatment with steroids, *n* (%)	491 (20.4)	448 (19.84)	43 (29.66)
Hospitalization, *n* (%)	288 (12.0)	257 (11.38)	31 (21.38)
Incomplete recovery from relapse, *n* (%)	186 (7.7)	164 (7.26)	22 (15.17)
Pyramidal signs, *n* (%)	795 (33.1)	699 (30.96)	96 (66.21)
Bowel/bladder sign, *n* (%)	154 (6.4)	125 (5.54)	29 (20)
Cerebellar signs, *n* (%)	388 (16.1)	332 (14.70)	56 (38.62)
Treatment			
% Time on first-line, 10 years	46.00 (36.10)	46.30 (36.50)	40.80 (29.99)
% Time on second-line, 10 years	5.30 (14.1)	4.90 (13.7)	11.30 (18.2)
% Time on first-line, 1 year	17.1 (28.00)	17.2 (28.2)	14.80 (23.4)
% Time on second-line, 1 year	0.50 (5.1)	0.40 (4.80)	1.40 (7.70)
Time to treatment initiation, years	2.29 (2.50)	2.29 (2.51)	2.30 (2.28)

Values are presented as mean (SD) unless otherwise stated.

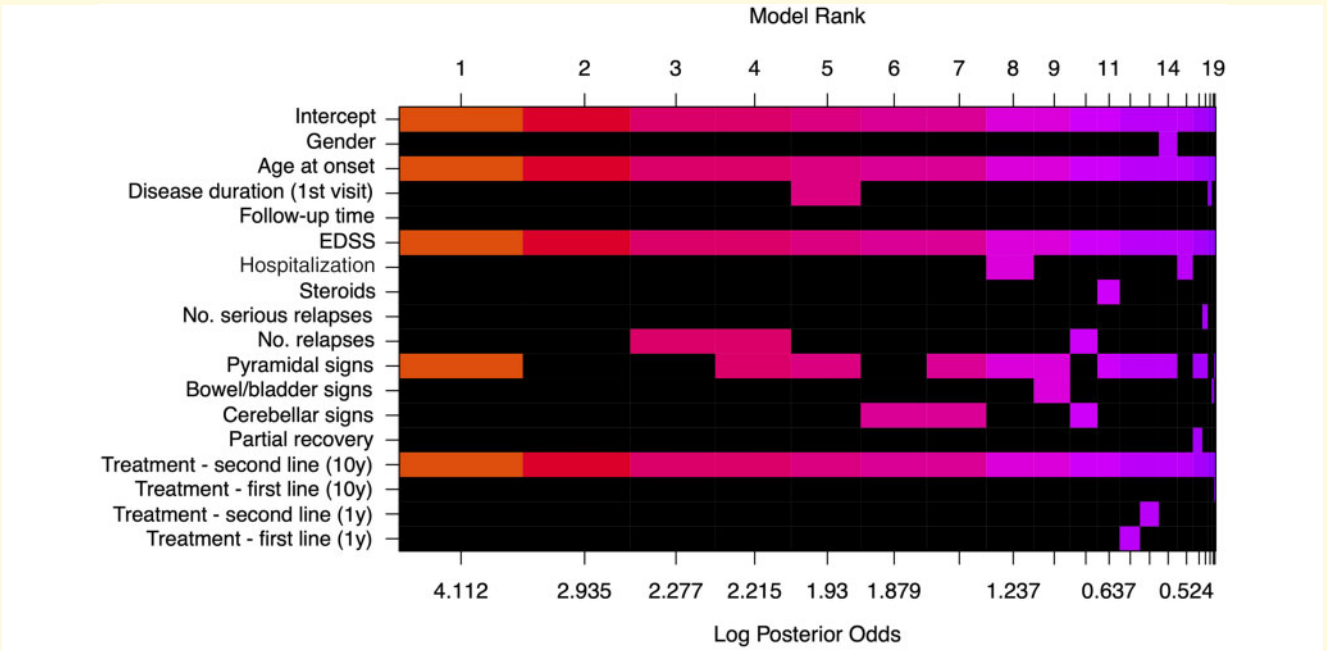

Figure 1 Visualization of the model space from the BMA analysis. The top 20 model is shown ranked by the log of the posterior odds. The colour indicates the posterior probability of the individual predictors (warmer = higher probability). Filled squares indicate that the predictor was not included in the model. As shown, age at onset, EDSS, and pyramidal signs were consistently included in the best fitting models. The intercept was included in all models.

be explicitly modelled via a random effects analysis. Two Bayesian generalized linear models were estimated; the first with the intercept fixed across all cases, and the second with a random intercept for each clinical site. The fit of each model was computed in the form of the leave-one out information criteria (LOOIC). As described in Supplementary material, section B, the fixed intercept model was a better fit to the data (LOOIC = 922.19) compared to the random intercept (LOOIC = 923.20). As such, the fixed intercept model, as implemented in the

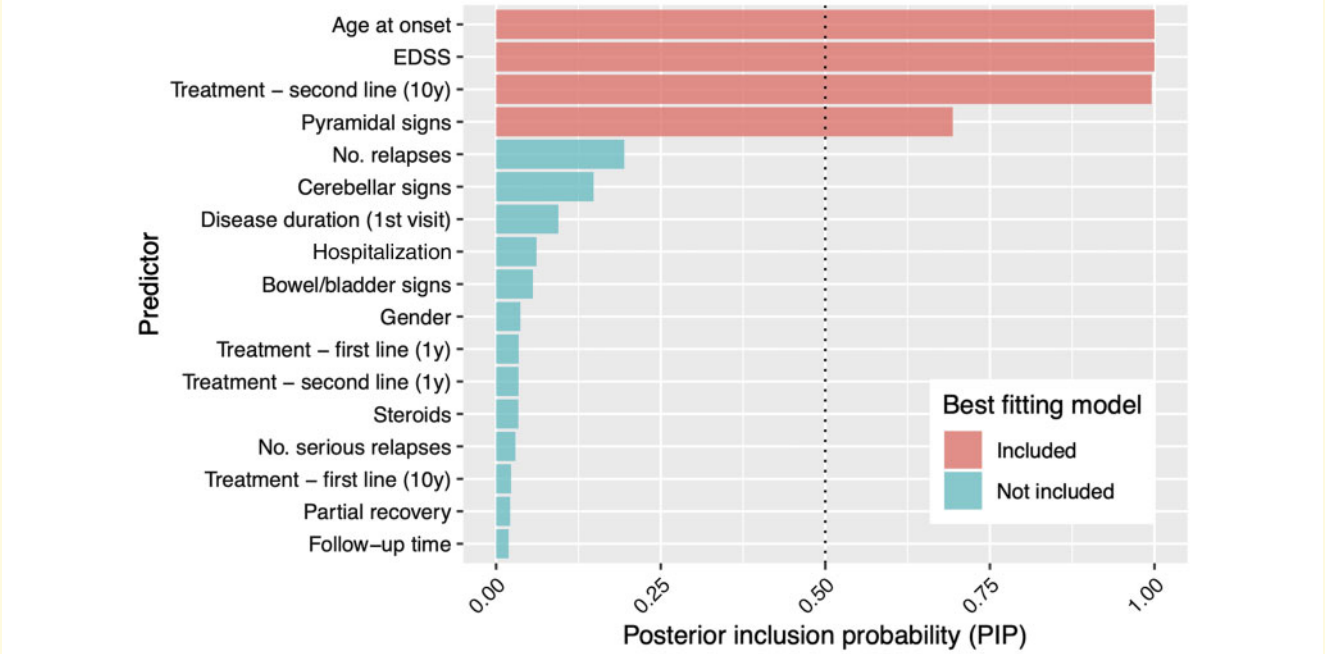

Figure 2 PIP for each predictor and model parameters averaged across the model space. EDSS, age at symptom onset, presence of pyramidal signs during the first year since symptom onset, and time spent on second-line therapy over 10 years were important predictors (PIPs > 0.5). All of these predictors were included in the single best fitting model. The evidence was not compelling for the other predictors.

BMA analysis, was deemed most appropriate for all analyses.

A second sensitivity analysis was performed in the subgroup of patients who had an EDSS < 6 at the time of first visit. After removing the 46 cases that had an EDSS $\geqslant 6$ at first visit (2% of cases), the sensitivity analysis was performed on the remaining 2357 cases. As shown in Supplementary material, section C, these results largely replicated the findings of the main analysis. Age at symptom onset (PIP = 1.00), EDSS at first visit (PIP = 1.00), and the presence of pyramidal signs in the first year (PIP = 0.74) were all important predictors. The proportion of time spend on a second-line therapy over the 10-year observation period was also an important predictor (PIP = 0.99).

Frequentist analysis

All analyses were repeated using a conventional frequentist approach. As shown in Table 3, the strongest evidence supported age and symptom onset and first year EDSS as significant predictors of aggressive multiple sclerosis ($P < 0.001$). Relatively weaker evidence was observed for the presence of pyramidal signs ($P = 0.017$). The proportion of time spent on second-line treatment over 10 years was statistically significance ($P < 0.001$). As shown in Supplementary material, section C, these findings were confirmed in the subgroup of patients with EDSS < 6 at first clinical visit.

Individual predictive accuracy

When averaged across the model space using BMA, the overall model was predictive of aggressive disease status. The area under the curve was 0.82 [95% confidence intervals (CI) = 0.78, 0.85] for the overall model, and 0.80 (95% CI = 0.75, 0.84) when using only the three important predictors (reduced model). Using the optimal cut-off of 0.06, the positive predictive value was 15%, compared to sample prevalence of aggressive disease of 6%. The negative predictive value was 98%. As shown in Table 4, ROC curves were also computed for each of the numerical predictors separately (age at symptom onset and median EDSS during the first year). Optimal thresholds were also computed, with patients classified as positive for each sign if their value was greater than the respective threshold (except for pyramidal signs, which were already binary). This procedure was repeated for a model containing all three of the important predictors identified above (age at symptom onset, median EDSS during the first year, and pyramidal signs).

The number of positive signs was then translated into the probability of developing aggressive disease. Of the patients with all three signs, 32% had aggressive disease. This was followed by any two signs (14%), any one sign (5%), and no signs (1%). The specific combinations of positive signs were also computed. As shown in Table 5, this revealed a more complex pattern of relationships. For example, being positive for EDSS and pyramidal signs was associated with slightly lower risk of aggressive disease (11%) compared to being positive for age and EDSS (21%).

External validity

The performance of the three important predictors was evaluated in an independent sample obtained from the Swedish Multiple Sclerosis Registry. The mean age at symptom onset (mean = 33.35, SD = 8.85) and median EDSS during the first year (mean = 1.51, SD = 1.28) were comparable to the MSBase cohort. The proportion of patients reported to show motor signs in the first year ($n = 95$, 17%) was lower than the proportion of MSBase patients with pyramidal signs in the first year ($n = 795$, 33%). Of the 556 patients, 34 met criteria for aggressive disease (6%). This prevalence was comparable to the MSBase cohort. The probability of meeting aggressive disease criteria was predicted in the Swedish Multiple Sclerosis Registry using three important predictors identified in the MSBase cohort only. The area under the curve was 0.75 (95% CI 0.66, 0.84), only marginally lower than in the MSBase cohort, and within the confidence limits. The ROC is visualized in the Supplementary material, section D. Using the optimal cut-off of 0.06 (determined from the primary analysis in the MSBase cohort), the positive predictive value was 15%, compared to sample prevalence of 6%. The negative predictive value was 97%.

Table 2 BMA analysis for prediction of aggressive disease status

Variable	B	SD	95% CI	Exp(B)	PIP
Baseline variables					
Age at symptom onset	0.07	0.01	0.05, 0.09	1.07	**1.00**
Male	0.01	0.06	0.00, 0.00	1.01	0.06
Disease duration at first visit	0.00	0.00	0.00, 0.00	1.00	0.15
First year EDSS	0.50	0.07	0.36, 0.64	1.65	**1.00**
Events within first 12 months					
Number of relapses	0.04	0.10	0.00, 0.29	1.04	0.25
Number of severe relapses	–0.01	0.05	0.00, 0.00	0.99	0.06
Treatment with steroids	0.01	0.07	0.00, 0.00	1.03	0.06
Hospitalization	0.02	0.12	0.00, 0.20	1.03	0.09
Incomplete recovery from relapse	0.00	0.05	0.00, 0.00	1.00	0.04
Pyramidal signs	0.47	0.36	0.00, 1.00	1.60	**0.75**
Bowel/bladder signs	0.02	0.12	0.00, 0.14	1.02	0.09
Cerebellar signs	0.07	0.19	0.00, 0.59	1.07	0.20
Treatment					
Time on first-line, 10 years	0.00	0.05	0.00, 0.00	1.00	0.04
Time on second-line, 10 years	2.32	0.52	1.37, 3.36	10.15	**1.00**
Time on first-line, 1 year	–0.01	0.10	0.00, 0.00	0.99	0.06
Time on second-line, 1 year	–0.05	0.37	0.00, 0.00	0.95	0.06

B = model coefficients computed as the mean of the posterior distribution; SE = standard deviation of the posterior distributions; CI = 95% credible intervals. Exp(B) = odds ratios of the model coefficients; $P(B! = 0 \mid Y)$ = posterior inclusion probability (PIP).

Discussion

In this study, we investigated a comprehensive set of early predictors of aggressive multiple sclerosis. Using data from the international MSBase registry (Butzkueven *et al.*, 2006), we identified early and clinically accessible makers that, if observed in the first year since symptom onset, convey increased risk of the patient meeting criteria for aggressive disease. Specifically, older age at symptom onset, greater disability in the first year, and the presence of pyramidal signs are associated with aggressive multiple sclerosis. Conversely, the absence of these signs conveyed lower probability of

Table 3 Frequentist general linear model analysis for prediction of aggressive disease status

Variable	B	SE	95% CIs	Exp(B)	Sig.
Baseline variables					
Age at symptom onset	0.07	0.01	0.05, 0.09	1.07	<**0.001**
Male	0.26	0.21	–0.15, 0.66	1.29	0.210
Disease duration at first visit	0.00	0.00	0.00, 0.00	1.00	0.223
First year EDSS	0.43	0.09	0.26, 0.60	1.54	<**0.001**
Events within first 12 months					
Number of relapses	0.18	0.12	–0.05, 0.41	1.20	0.122
Number of severe relapses	–0.33	0.23	–0.78, 0.12	0.72	0.147
Treatment with steroids	0.08	0.26	–0.42, 0.59	1.09	0.745
Hospitalization	0.25	0.32	–0.37, 0.88	1.29	0.429
Incomplete recovery from relapse	0.03	0.32	–0.59, 0.65	1.03	0.935
Pyramidal signs	0.56	0.24	0.10, 1.03	1.76	**0.017**
Bowel/bladder signs	0.32	0.29	–0.26, 0.89	1.38	0.276
Cerebellar signs	0.34	0.23	–0.10, 0.78	1.41	0.130
Treatment					
Time on first-line, 10 years	–0.19	0.32	–0.80, 0.43	0.83	0.554
Time on second-line, 10 years	2.27	0.58	1.14, 3.41	9.71	<**0.001**
Time on first-line, 1 year	–0.29	0.43	–1.14, 0.55	0.75	0.497
Time on second-line, 1 year	–1.36	1.63	–4.56, 1.84	0.26	0.404

B = model coefficients; SE = robust standard errors computed via heteroscedasticity-consistent sandwich estimation; CI = 95% confidence intervals computed using robust SEs; Exp(B) = odds ratios for the model coefficients; Sig = significance values were computed using the robust SEs. Bold significance values indicate $P < 0.05$. Assumption checks regarding multicollinearity, influential cases, autocorrelation, dispersion, and residual distribution were satisfied.

Table 4 Classification performance of the BMA analysis and specific variables

Variable	AUC (95% CIs)	Threshold	Sensitivity	Specificity	PPV	NPV
BMA (full)	0.82 (0.78, 0.85)	0.05	0.78	0.71	0.15	0.98
BMA (reduced)	0.80 (0.75, 0.84)	0.06	0.72	0.73	0.15	0.98
Age at onset	0.69 (0.64, 0.74)	35	0.60	0.71	0.12	0.96
First year EDSS	0.74 (0.70, 0.79)	2.5	0.53	0.85	0.19	0.97
Pyramidal signs	–	–	0.66	0.69	0.12	0.97

AUC = area under the curve; NPV = negative predictive value; PPV = positive predictive value. Threshold = the optimal cut-off determined using Youden's method (1955). As the presence pyramidal signs is a binary predictor, the AUC and threshold could not be computed. BMA (full) = the entire model with all 17 predictors; BMA (reduced) = the model with only age at onset, first year EDSS, and pyramidal signs.

Table 5 Prognostic value of combinations of positive signs

Age >35	EDSS ⩾3	Pyramidal signs	*n*	Non-aggressive disease/aggressive disease, *n*	% Aggressive disease	Relative risk
✓	✓	✓	125	85/40	32.00	22.85
✓	✓		52	41/11	21.15	15.11
✓		✓	177	154/23	12.99	9.28
	✓		88	78/10	11.36	8.11
	✓	✓	142	126/16	11.27	8.05
		✓	351	334/17	4.84	3.46
✓			397	384/13	3.27	2.34
			1071	1056/15	1.40	Reference

meeting aggressive multiple sclerosis criteria compared to the observed base rate. The predictive power of the overall model was replicated using an independent cohort of patients from the Swedish Multiple Sclerosis Registry. The Bayesian analytical approach enabled us to estimate posterior probabilities of aggressive disease which can be directly applied to individual patients who have recently presented with symptoms of relapse-onset multiple sclerosis.

The proportions of patients who develop aggressive multiple sclerosis form has been reported between 5.2% and 8.6% (Gholipour *et al.*, 2011; Menon *et al.*, 2013, 2017). The variability of the reported rates of aggressive multiple sclerosis is surprisingly small, given the highly variable definitions that were used by previous studies, ranging from EDSS ⩾ 6 at 5 years from disease onset to complex definitions consisting of patient age, time from first treatment, EDSS and relapse activity (Gholipour *et al.*, 2011; Saccardi *et al.*, 2012; Menon *et al.*, 2013). The definition of confirmed and sustained EDSS ⩾ 6 reached within 10 years from multiple sclerosis onset, which we used in the present study, is consistent with the definition used by the Barcelona cohort (Tintore *et al.*, 2019) and the EPIC cohort (Cree *et al.*, 2016). The three studies showed similar rates of aggressive multiple sclerosis, 6%, 7% and 4.7%, respectively. In the Barcelona cohort, MRI predictors were associated with 19% and 40% rate of aggressive disease (positive predictive value), depending on the thresholds for MRI activity. In our discovery MSBase cohort, a combination of three prognostic markers—age, EDSS and pyramidal signs—were associated with 32% rate of aggressive multiple sclerosis. In the validation cohort from Swedish Multiple Sclerosis Registry, the rate of aggressive multiple sclerosis was 6%, and the prognostic markers were associated with 15% rate of aggressive disease. The positive predictive value in the validation cohort was lower than in the discovery cohort, as expected, but it was still substantially higher than the mean population rate of aggressive multiple sclerosis.

Similarly, our identified early markers of aggressive disease are consistent with a number of previous studies. Several studies have reported general associations of older age at symptom onset (Confavreux *et al.*, 2003; Held *et al.*, 2005; Confavreux and Vukusic, 2006), early disability progression (Weinshenker *et al.*, 1989), and motor signs with accelerated worsening of disability (Weinshenker *et al.*, 1989; Atkins and Freedman, 2013; Correale *et al.*, 2014). In a number of areas, our results may seem to diverge from previously reported findings. For example, we did not find evidence for an independent, unequivocal effect of sex (Kantarci *et al.*, 1998; Tremlett *et al.*, 2006), incomplete recovery from relapse (Trojano *et al.*, 1995; Scott and Schramke, 2010), severity of relapse (Rush *et al.*, 2015), the number of relapses early in the disease course (Weinshenker *et al.*, 1989; Confavreux *et al.*, 2003; Ebers, 2005), and relapses presenting with cerebellar (Weinshenker *et al.*, 1989; Phadke, 1990; Amato *et al.*, 1999) or bowel/bladder (Citterio *et al.*, 1989; Runmarker and Andersen, 1993; Amato *et al.*, 1999; Langer-Gould *et al.*, 2006) signs on the risk of developing aggressive disease. Because of issues of multi-collinearity, we were unable to investigate the independent effect of cognitive impairment, which has previously been suggested to be a poor prognostic sign (Zarei *et al.*, 2003). The main conceptual difference between the present study and the previously published work is its focus on an objectively defined aggressive disease form rather than the overall risk of disability worsening used as the outcome in this study. Therefore, it is not surprising that several markers, which

are predictive of disability worsening in general, are not directly associated with the risk of developing aggressive disease—a disease form that is at an extreme of the continuum of disease severity. It is also not surprising that the proportion of time on second-line therapy was predictive of aggressive disease. Patients with early accrual of disability are more likely to be treated with more aggressive therapy. By including treatment variables in our analysis, we were able to adjust our model for this potential confounder.

Identifying patients who are at a high risk of developing an aggressive multiple sclerosis is of high clinical importance. Previously, we have demonstrated that choice of a second-line therapy, in particular early in the disease course, has the capacity to prevent or delay development of the secondary progressive multiple sclerosis form (Brown *et al.*, 2019) and mitigate the evolution of neurological disability (He *et al.*, in press). An objective guidance for treatment of potentially aggressive disease at early disease stages in a condition whose management revolves around effective prevention disability driven by episodes of CNS inflammation, is an area of unmet need. The Bayesian approach used in this study allows us not only to quantify, within a year from disease onset, the probability of developing significant restriction of gait within 10 years in individual patients, but also to quantify the uncertainty across competing models—a feature that cannot be assessed within frequentist frameworks (Raftery *et al.*, 1997). Reassuringly, the direction and the magnitude of the associations identified within the Bayesian and the frequentist frameworks are almost identical.

Many of these predictors share a significant proportion of variance. Our statistical approach of combining all predictors into a single analysis enables us to select only those predictors that add uniquely to the prediction. This redundancy allows us to focus the resulting predictive algorithm on a small number of accessible predictors that function similar to the full set of available variables (Table 4 and Fig. 3).

To the best of our knowledge, this is the largest study conducted to identify early clinical markers of aggressive multiple sclerosis. It enables direct translation of its results into clinical application. We used a strict set of inclusion criteria that required all patients to be observed within 12 months of symptom onset for a minimum of 10 years. This ensured that our control cases were truly negative in terms of aggressive multiple sclerosis. By requiring patients to meet criteria for aggressive multiple sclerosis from at least 6 months and until the end of follow-up, we ensured that the risk of false positives (i.e. patients who met aggressive disease criteria but then reverted to a negative status) was minimized. Our statistical approach was also a significant strength. BMA has a long history in the statistical literature but has not been applied extensively in multiple sclerosis research. By averaging the results over all possible models, BMA minimizes model selection bias (Raftery *et al.*, 1997). This approach advances the approaches used previously to predict the course of disability. In our approach, we have not relied on a single, arbitrarily selected model, but have reported the result that reflects the entire model space and combines possible statistical models.

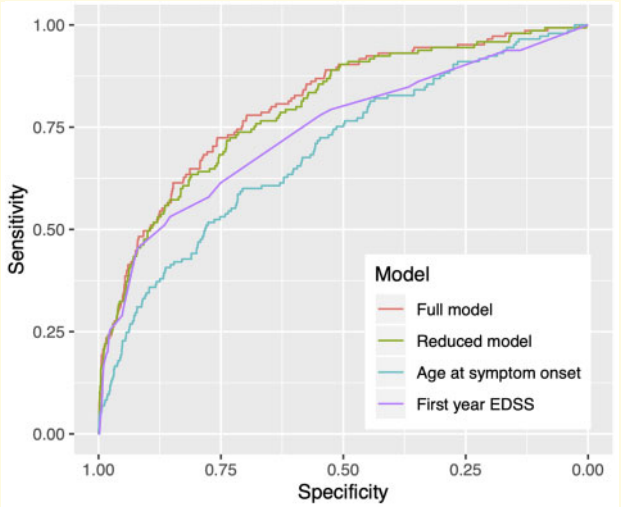

Figure 3 ROC curve. Full model = BMA analysis containing all predictors weighted by model fit. Reduced model = the model that contained only age at symptom onset, first year EDSS, and presence or pyramidal signs. Age at symptom onset = only the continuous predictor of age at symptom onset. First year EDSS = only the continuous predictor of median EDSS in the first year.

External validity of its results is imperative for any predictive algorithm (Copas, 1983; Harrell *et al.*, 1996). We therefore used an independent, population-based cohort, the Swedish Multiple Sclerosis Registry, which captures ~80% of the Swedish multiple sclerosis population, to validate our findings (Hillert and Stawiarz, 2015). This validation analysis confirmed that the overall model derived from the BMA analysis provided an accurate prediction in an independent multiple sclerosis cohort, and is thus generalizable to the prevalent multiple sclerosis population. At the level of the individual predictors contributing to the model, not all variables could be replicated directly. The definition of 'motor' signs in the Swedish Multiple Sclerosis Registry differs from the definition used by MSBase. Specifically, MSBase uses the pyramidal functional system score based on EDSS (Kurtzke, 1983) while the Swedish Multiple Sclerosis Registry reports motor symptoms as a separate entity. This difference is further supported by the observation of the relatively lower frequency of motor signs in the validation cohort compared to pyramidal signs in the discovery cohort. Further limitations are represented by the relative scarcity of formally reported MRI data and CSF information. We have conducted secondary analyses in sub-cohorts in whom this information was available; these analyses were not able to detect substantial contributions of qualitatively or quantitatively reported MRI and CSF oligoclonal bands during the first year to the prediction of aggressive disease based on clinical predictors. However, the size of the cohort with MRI data available was limited, with only 29 of 359 patients meeting criteria for aggressive multiple sclerosis and

no information about spinal cord imaging, and this result should be viewed with caution. In studies where MRI information was acquired prospectively from disease onset and with standardized MRI protocols, high lesion load was predictive of more severe disease course (Tintore *et al.*, 2015, 2019). Finally, the data used originate from observational registries based on clinical practice and are subjects to multiple sources of error and confounding. To mitigate these influences, we have applied an objective data quality assessment (Kalincik *et al.*, 2017) and have constructed large, inclusive multivariable models that have contributed to the final model through the BMA framework.

In summary, our findings suggest that older age at symptom onset, greater disability in the first year since symptom onset, and the presence of pyramidal signs indicate a higher risk of developing aggressive disease. Importantly, the absence of any of these signs is associated with a much lower risk of aggressive disease compared to the general multiple sclerosis population. For ease of clinical implementation, these criteria can be defined as: median EDSS $\geqslant 3$, any pyramidal signs on examination in the first year, and age >35 at symptom onset. An important direction for future research will be to determine whether aggressive multiple sclerosis can be prevented by the use of aggressive treatment strategies in patients who are predicted to develop aggressive disease.

Funding

This study was financially supported by the National Health and Medical Research Council of Australia (1129189, 1140766, 1157717). The MSBase Foundation is a not-for-profit organization that receives support from Roche, Merck, Biogen, Novartis, Bayer Schering, Sanofi Genzyme and Teva. The study was conducted separately and apart from the guidance of the sponsors.

Competing interests

A.M. is supported by Margaretha af Ugglas foundation. He has received consultancy fees from Biogen and speaker honoraria and/or travel expenses from Biogen and ECTRIMS. D.H. received speaker honoraria and consulting fees from Biogen, Merck, Teva and Novartis, as well as support for research activities from Biogen and research grants from Charles University in Prague [PRVOUK-P26/LF1/4], Czech Minsitry of Education [PROGRES Q27/LF1] and Czech Ministry of Health [NT13237-4/2012]. E.K.H. received speaker honoraria and consultant fees from Actelion, Biogen, Celgene, Merck, Novartis, Roche, Sanofi and Teva, and support for research activities from Czech Ministry of Education [project Progres Q27/LF1]. M.T. received speaker honoraria from Biogen-Idec, Bayer-Schering, Sanofi Aventis, Merck, Teva, Novartis and Almirall; has received research grants for her Institution from Biogen-Idec, Merck, and Novartis. G.I. received speaking honoraria from Biogen, Novartis, Sanofi, Merck, Roche, Almirall and Teva. S.E. received speaker honoraria and consultant fees from Biogen, Merck, Novartis, Almirall, Roche, Sanofi and Teva. R.B. received speaker honoraria from Bayer Schering, Biogen, Genzyme, Merck, Novartis, Sanofi-Aventis, Teva; research grants from Bayer Schering, Biogen, Merck, Novartis, Sanofi-Aventis, Teva; congress and travel/accommodation expense compensations by Almirall, Bayer Schering, Biogen, Genzyme, Merck, Novartis, Sanofi-Aventis, Teva. P.S. served on scientific advisory boards for Biogen Idec and TEVA, she has received funding for travel and speaker honoraria from Biogen Idec, Merck, Teva, Sanofi Genzyme, Novartis and Bayer and research grants for her Institution from Bayer, Biogen, Merck, Novartis, Sanofi, Teva. D.F. received travel grants and/or speaker honoraria from Merck, TEVA, Novartis, Biogen and Sanofi-Genzyme. A.L. is a Bayer, Biogen, Genzyme, Merck Advisory Board Member. She received travel grants and honoraria from Roche, Bayer, Biogen, Merck, Novartis, Sanofi, Teva and Fondazione Italiana Sclerosi Multipla (FISM). Her institution received research grants from Bayer, Biogen, Merck, Novartis, Sanofi, Teva and Fondazione Italiana Sclerosi Multipla (FISM). M.G. received consulting fees from Teva Canada Innovation, Biogen, Novartis and Genzyme Sanofi; lecture payments from Teva Canada Innovation, Novartis and EMD. He has also received a research grant from Canadian Institutes of Health Research. P.D. served on editorial boards and has been supported to attend meetings by EMD, Biogen, Novartis, Genzyme, and TEVA Neuroscience. He holds grants from the CIHR and the multiple sclerosis Society of Canada and has received funding for investigator-initiated trials from Biogen, Novartis, and Genzyme. P.G. is a Merck, Novartis, Teva-neuroscience, Biogen and Genzyme advisory board member, consultant for Merck, received payments for lectures by Merck, Teva-Neuroscience and Canadian Multiple sclerosis society, and received grants for travel from Teva-Neuroscience and Novartis. F.G. received honoraria or research funding from Biogen, Genzyme, Novartis, Teva Neurosciences, Mitsubishi and ONO Pharmaceuticals. V.V.P. received travel grants from Biogen, Bayer Schering, Genzyme, Merck, Teva and Novartis Pharma. His institution receives honoraria for consultancy and lectures from Biogen, Bayer Schering, Genzyme, Merck, Roche, Teva and Novartis Pharma as well as research grants from Novartis Pharma and Bayer Schering. F.G. received research grant from Biogen, served on scientific advisory boards for Biogen, Novartis, Merck, and Sanofi-Aventis and received funding for travel and speaker honoraria from Biogen, Merck, Sanofi-Aventis, and Almirall. R.H. received honoraria as consultant on scientific advisory boards from Merck, Biogen, Genzyme-Sanofi and Teva, research funding from Merck and Biogen, and speaker honoraria from Sanofi-Genzyme and Novartis. E.P. served on scientific advisory boards for Merck, Genzyme and Biogen; he has received honoraria and travel grants from Sanofi Aventis, Novartis, Biogen, Merck, Genzyme and Teva; he has

received travel grants and equipment from 'Associazione Marchigiana Sclerosi Multipla e altre malattie neurologiche'. T.K. served on scientific advisory boards for Roche, Genzyme-Sanofi, Novartis, Merck and Biogen, steering committee for Brain Atrophy Initiative by Genzyme, received conference travel support and/or speaker honoraria from WebMD Global, Novartis, Biogen, Genzyme-Sanofi, Teva, BioCSL and Merck and received research support from Biogen. C.B. received conference travel support from Biogen, Novartis, Bayer-Schering, Merck and Teva; has participated in clinical trials by Sanofi Aventis, Roche and Novartis. D.S. received honoraria as a consultant on scientific advisory boards by Bayer-Schering, Novartis and Sanofi-Aventis and compensation for travel from Novartis, Biogen, Sanofi Aventis, Teva and Merck. H.B. served on scientific advisory boards for Biogen, Novartis and Sanofi-Aventis and has received conference travel support from Novartis, Biogen and Sanofi Aventis. He serves on steering committees for trials conducted by Biogen and Novartis, and has received research support from Merck, Novartis and Biogen. T.P. received funding or speaker honoraria from Biogen, Merck, Novartis, Bayer Schering, Sanofi-Aventis, Roche, and Genzyme. F.V. is an advisory board member for Teva Biogen Merck and Novartis. C.R-T. received research funding, compensation for travel or speaker honoraria from Biogen, Novartis, Genzyme, Merck, Roche and Almirall. M.T. received travel grants from Novartis, Bayer-Schering, Merck and Teva; has participated in clinical trials by Sanofi Aventis, Roche and Novartis. E.C. received honoraria as consultant on scientific advisory boards by Biogen, Bayer-Schering, Merck, Genzyme and Novartis; has participated in clinical trials/other research projects by Merck, Roche and Novartis. M.S. has participated in, but not received honoraria for, advisory board activity for Biogen, Merck, Bayer Schering, Sanofi Aventis and Novartis. Y.F. received honoraria as a consultant on scientific advisory boards by Novartis, Teva, Roche and Sanofi-Aventis and compensation for travel from Novartis, Biogen, Sanofi Aventis, Teva, Roche and Merck. J.O. serves on scientific advisory boards for Biogen, Novartis, Sanofi and Roche; has received speaker honoraria from Almirall, Biogen, Bayer, Sanofi, Merck, Novartis and Roche and research grants from Biogen, Merck, Novartis and Teva. A.A. received personal fees and speaker honoraria from Teva, Merck, Novartis, Sanofi-Genzyme; received travel and registration grants from Merck, Roche, Sanofi-Genzyme. T.O. has received honoraria for lectures/advisory boards and/or unrestricted multiple sclerosis research grants from Biogen, Novartis, Sanofi, Merck and Roche. C.B.M., S.S., A.P., S.O., Y.S., R.G., A.S., R.K., R.T., P.M., and R.M. report no competing interests.

Supplementary material

Supplementary material is available at *Brain* online.

References

Amato MP, Ponziani G, Bartolozzi ML, Siracusa G. A prospective study on the natural history of multiple sclerosis: clues to the conduct and interpretation of clinical trials. J Neurol Sci 1999; 168: 96–106.

Atkins HL, Freedman MS. Hematopoietic stem cell therapy for multiple sclerosis: top 10 lessons learned. Neurotherapeutics 2013; 10: 68–76.

Bayarri MJ, Berger JO, Forte A, García-Donato G. Criteria for Bayesian model choice with application to variable selection. Ann Stat 2012; 40: 1550–77.

Bergamaschi R, Berzuini C, Romani A, Cosi V. Predicting secondary progression in relapsing–remitting multiple sclerosis: a Bayesian analysis. J Neurol Sci 2001; 189: 13–21.

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: 158–64.

Brown JWL, Coles A, Horakova D, Havrdova E, Izquierdo G, Prat A, et al. Association of initial disease-modifying therapy with later conversion to secondary progressive multiple sclerosis. JAMA 2019; 321: 175–87.

Bürkner P-C. brms: an R package for Bayesian multilevel models using Stan. J Stat Soft 2017; 80: 1–28.

Bürkner P-C. Advanced Bayesian multilevel modeling with the R package brms. R J 2018; 10: 395–411.

Butzkueven H, Chapman J, Cristiano E, Grand'Maison F, Hoffmann M, Izquierdo G, et al. MSBase: an international, online registry and platform for collaborative outcomes research in multiple sclerosis. Mult Scler 2006; 12: 769–74.

Citterio A, Azan G, Bergamaschi R, Erbetta A, Cosi V. Multiple sclerosis: disability and mortality in a cohort of clinically diagnosed patients. Neuroepidemiology 1989; 8: 249–53.

Clyde MA, Ghosh J, Littman ML. Bayesian adaptive sampling for variable selection and model averaging. J Comput Graph Stat 2011; 20: 80–101.

Confavreux C, Vukusic S. Age at disability milestones in multiple sclerosis. Brain 2006; 129: 595–605.

Confavreux C, Vukusic S, Adeleine P. Early clinical predictors and progression of irreversible disability in multiple sclerosis: an amnesic process. Brain 2003; 126: 770–82.

Copas JB. Regression, prediction and shrinkage. J R Stat Soc Ser B Methodol 1983; 45: 311–35.

Correale J, Abad P, Alvarenga R, Alves-Leon S, Armas E, Barahona J, et al. Management of relapsing–remitting multiple sclerosis in Latin America: practical recommendations for treatment optimization. J Neurol Sci 2014; 339: 196–206.

Cree BA, Gourraud PA, Oksenberg JR, Bevan C, Crabtree-Hartman E, Gelfand JM, et al. Long-term evolution of multiple sclerosis disability in the treatment era. Ann Neurol 2016; 80: 499–510.

Ebers GC. Prognostic factors for multiple sclerosis: the importance of natural history studies. J Neurol 2005; 252: iii15–iii20.

European Committee for Treatment and Research in Multiple Sclerosis. 6th Focused Workshop of the European Committee for Treatment and Research in Multiple Sclerosis. 1-2 March, 2018; Brussels, Belgium.

Fernández Ó. Is there a change of paradigm towards more effective treatment early in the course of apparent high-risk MS? Mult Scler Relat Disord 2017; 17: 75–83.

Filippi M, Horsfield M, Morrissey S, MacManus D, Rudge P, McDonald W, et al. Quantitative brain MRI lesion load predicts the course of clinically isolated syndromes suggestive of multiple sclerosis. Neurology 1994; 44: 635.

Gholipour T, Healy B, Baruch N, Weiner HL, Chitnis T. Demographic and clinical characteristics of malignant multiple sclerosis. Neurology 2011; 76: 1996–2001.

Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996; 15: 361–87.

Held U, Heigenhauser L, Shang C, Kappos L, Polman C. Predictors of relapse rate in MS clinical trials. Neurology 2005; 65: 1769–73.

Hillert J, Stawiarz L. The Swedish MS registry–clinical support tool and scientific resource. Acta Neurol Scand 2015; 132: 11–9.

Hoeting JA, Madigan D, Raftery AE, Volinsky CT. Bayesian model averaging: a tutorial. Stat Sci 1999; 14: 382–401.

Jokubaitis VG, Spelman T, Kalincik T, Lorscheider J, Havrdova E, Horakova D, et al. Predictors of long-term disability accrual in relapse-onset multiple sclerosis. Ann Neurol 2016; 80: 89–100.

Kalincik T. Multiple sclerosis relapses: epidemiology, outcomes and management. A systematic review. Neuroepidemiology 2015; 44: 199–214.

Kalincik T, Kuhle J, Pucci E, Rojas JI, Tsolaki M, Sirbu C-A, et al. Data quality evaluation for observational multiple sclerosis registries. Mult Scler 2017; 23: 647–55.

Kantarci O, Siva A, Eraksoy M, Karabudak R, Sütlaş N, Ağaoğlu J, et al. Survival and predictors of disability in Turkish MS patients. Neurology 1998; 51: 765–72.

Kappos L, Moeri D, Radue EW, Schoetzau A, Schweikert K, Barkhof F, et al. Predictive value of gadolinium-enhanced magnetic resonance imaging for relapse rate and changes in disability or impairment in multiple sclerosis: a meta-analysis. Lancet 1999; 353: 964–9.

Kornek B, Lassmann H. Neuropathology of multiple sclerosis—new concepts. Brain Res Bull 2003; 61: 321–6.

Kraft G, Freal J, Coryell J, Hanan C, Chitnis N. Multiple sclerosis: early prognostic guidelines. Arch Phys Med Rehabil 1981; 62: 54–8.

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

Langer-Gould A, Popat RA, Huang SM, Cobb K, Fontoura P, Gould MK, et al. Clinical and demographic predictors of long-term disability in patients with relapsing-remitting multiple sclerosis: a systematic review. Arch Neurol 2006; 63: 1686–91.

Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. Neurology 1996; 46: 907–11.

Lukacs PM, Burnham KP, Anderson DR. Model selection bias and Freedman's paradox. Ann Inst Stat Math 2010; 62: 117.

Lukas C, Minneboo A, De Groot V, Moraal B, Knol DL, Polman CH, et al. Early central atrophy rate predicts 5 year clinical outcome in multiple sclerosis. J Neurol Neurosurg Psychiatry 2010; 81: 1351–6.

Menon S, Shirani A, Zhao Y, Oger J, Traboulsee A, Freedman MS, et al. Characterising aggressive multiple sclerosis. J Neurol Neurosurg Psychiatry 2013; 84: 1192–8.

Menon S, Zhu F, Shirani A, Oger J, Freedman MS, Tremlett H. Disability progression in aggressive multiple sclerosis. Mult Scler 2017; 23: 456–63.

Miller A. Subset selection in regression. Boca Raton: Chapman and Hall/CRC; 2002.

O'Riordan JI, Thompson A, Kingsley DP, MacManus D, Kendall BE, Rudge P, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS. A 10-year follow-up. Brain 1998; 121: 495–503.

Phadke JG. Clinical aspects of multiple sclerosis in north-east Scotland with particular reference to its course and prognosis. Brain 1990; 113: 1597–628.

Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 2011; 69: 292–302.

R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2019.

Raftery AE, Madigan D, Hoeting JA. Bayesian model averaging for linear regression models. J Am Stat Assoc 1997; 92: 179–91.

Ribbons KA, McElduff P, Boz C, Trojano M, Izquierdo G, Duquette P, et al. Male sex is independently associated with faster disability accumulation in relapse-onset MS but not in primary progressive MS. PloS One 2015; 10: e0122686.

Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics 2011; 12: 77.

Rudick RA, Lee JC, Simon J, Fisher E. Significance of T2 lesions in multiple sclerosis: a 13-year longitudinal study. Ann Neurol 2006; 60: 236–42.

Runmarker B, Andersen O. Prognostic factors in a multiple sclerosis incidence cohort with twenty-five years of follow-up. Brain 1993; 116: 117–34.

Rush CA, MacLean HJ, Freedman MS. Aggressive multiple sclerosis: proposed definition and treatment algorithm. Nat Rev Neurol 2015; 11: 379.

Saccardi R, Freedman MS, Sormani MP, Atkins H, Farge D, Griffith LM, et al. A prospective, randomized, controlled trial of autologous haematopoietic stem cell transplantation for aggressive multiple sclerosis: a position paper. Mult Scler J 2012; 18: 825–34.

Sailer M, O'Riordan J, Thompson AJ, Kingsley DP, MacManus DG, McDonald WI, et al. Quantitative MRI in patients with clinically isolated syndromes suggestive of demyelination. Neurology 1999; 52: 599–606.

Scott TF, Laforet G, You X. Aggressive relapsing multiple sclerosis characterized by rapid disability progression. Mult Scler Relat Disord 2013; 2: 370–6.

Scott TF, Schramke CJ. Poor recovery after the first two attacks of multiple sclerosis is associated with poor outcome five years later. J Neurol Sci 2010; 292: 52–6.

Tintore M, Arrambide G, Otero-Romero S, Carbonell-Mirabent P, Río J, Tur C, et al. The long-term outcomes of CIS patients in the Barcelona inception cohort: Looking back to recognize aggressive MS. Mult Scler J 2019: doi: 10.1177/1352458519877810.

Tintore M, Rovira A, Río J, Otero-Romero S, Arrambide G, Tur C, et al. Defining high, medium and low impact prognostic factors for developing multiple sclerosis. Brain 2015; 138: 1863–74.

Tomassini V, Paolillo A, Russo P, Giugni E, Prosperini L, Gasperini C, et al. Predictors of long–term clinical response to interferon beta therapy in relapsing multiple sclerosis. J Neurol 2006; 253: 287–93.

Tramacere I, Del Giovane C, Salanti G, D'Amico R, Filippini G. Immunomodulators and immunosuppressants for relapsing-remitting multiple sclerosis: a network meta-analysis. Cochrane Database Syst Rev 2015: 9. doi: 10.1002/14651858.CD011381.pub2.

Tremlett H, Paty D, Devonshire V. Disability progression in multiple sclerosis is slower than previously reported. Neurology 2006; 66: 172–7.

Trojano M, Avolio C, Manzari C, Calo A, De Robertis F, Serio G, et al. Multivariate analysis of predictive factors of multiple sclerosis course with a validated method to assess clinical events. J Neurol Neurosurg Psychiatry 1995; 58: 300–6.

Uher T, Vaneckova M, Sobisek L, Tyblova M, Seidl Z, Krasensky J, et al. Combining clinical and magnetic resonance imaging markers enhances prediction of 12-year disability in multiple sclerosis. Mult Scler 2017; 23: 51–61.

Vaneckova M, Kalincik T, Krasensky J, Horakova D, Havrdova E, Hrebikova T, et al. Corpus callosum atrophy–a simple predictor of multiple sclerosis progression: a longitudinal 9-year study. Eur Neurol 2012; 68: 23–7.

Weinshenker BG, Bass B, Rice G, Noseworthy J, Carriere W, Baskerville J, et al. The natural history of multiple sclerosis: a geographically based study: 2 predictive value of the early clinical course. Brain 1989; 112: 1419–28.

Weinshenker BG, Issa M, Baskerville J. Long-term and short-term outcome of multiple sclerosis: a 3-year follow-up study. Arch Neurol 1996; 53: 353–8.

Wolfson C, Confavreux C. Improvements to a simple Markov model of the natural history of multiple sclerosis. Neuroepidemiology 1987; 6: 101–15.

Youden WJ. Index for rating diagnostic tests. Cancer 1950; 3: 32–5.

Zarei M, Chandran S, Compston A, Hodges J. Cognitive presentation of multiple sclerosis: evidence for a cortical variant. J Neurol Neurosurg Psychiatry 2003; 74: 872–7.

Zeileis A. Econometric computing with HC and HAC covariance matrix estimators. Research Report Series / Department of Statistics and Mathematics 2004: 10. Available from: https://epub.wu.ac.at/id/eprint/520 (10 February 2020, date last accessed).

Zeileis A. Object-oriented computation of sandwich estimators. J Stat Soft 2006; 16: 1–16.