Long-Term Evolution of Multiple Sclerosis Disability in the Treatment Era

University of California, San Francisco MS-EPIC Team:
Bruce A. C. Cree, MD, PhD, MAS,[1] Pierre-Antoine Gourraud, PhD, MPH,[1]
Jorge R. Oksenberg, PhD,[1] Carolyn Bevan, MD, MS,[1]
Elizabeth Crabtree-Hartman, MD,[1] Jeffrey M. Gelfand, MD, MAS,[1]
Douglas S. Goodin, MD,[1] Jennifer Graves, MD, PhD, MAS,[1] Ari J. Green, MD,[1]
Ellen Mowry, MD, MAS,[1] Darin T. Okuda, MD,[2] Daniel Pelletier, MD,[3]
H.- Christian von Büdingen, MD,[1] Scott S. Zamvil, MD, PhD,[1] Alisha Agrawal,[1]
Stacy Caillier,[1] Caroline Ciocca,[1] Refujia Gomez,[1] Rachel Kanner,[1] Robin Lincoln,[1]
Antoine Lizee, MSc,[1] Pamela Qualley,[1] Adam Santaniello,[1] Leena Suleiman,[1]
Monica Bucci, MD,[1] Valentina Panara, MD,[1] Nico Papinutto, PhD,[1]
William A. Stern,[1] Alyssa H. Zhu, MSc,[1] Gary R. Cutter, PhD,[4]
Sergio Baranzini, PhD,[1] Roland G. Henry, PhD,[1] and Stephen L. Hauser, MD[1]

Objective: To characterize the accrual of long-term disability in a cohort of actively treated multiple sclerosis (MS) patients and to assess whether clinical and magnetic resonance imaging (MRI) data used in clinical trials have long-term prognostic value.
Methods: This is a prospective study of 517 actively managed MS patients enrolled at a single center.
Results: More than 91% of patients were retained, with data ascertained up to 10 years after the baseline visit. At this last assessment, neurologic disability as measured by the Expanded Disability Status Scale (EDSS) was stable or improved compared to baseline in 41% of patients. Subjects with no evidence of disease activity (NEDA) by clinical and MRI criteria during the first 2 years had long-term outcomes that were no different from those of the cohort as a whole. 25-OH vitamin D serum levels were inversely associated with short-term MS disease activity; however, these levels had no association with long-term disability. At a median time of 16.8 years after disease onset, 10.7% (95% confidence interval [CI] = 7.2–14%) of patients reached an EDSS $\geq$ 6, and 18.1% (95% CI = 13.5–22.5%) evolved from relapsing MS to secondary progressive MS (SPMS).
Interpretation: Rates of worsening and evolution to SPMS were substantially lower when compared to earlier natural history studies. Notably, the NEDA 2-year endpoint was not a predictor of long-term stability. Finally, the data call into question the utility of annual MRI assessments as a treat-to-target approach for MS care.

ANN NEUROL 2016;80:499–510

View this article online at wileyonlinelibrary.com. DOI: 10.1002/ana.24747

This article was published online on 13 August 2016. An error was subsequently identified. This notice is included in the online and print versions to indicate that both have been corrected on 03 September 2016.

Received Jun 11, 2015, and in revised form Jul 12, 2016. Accepted for publication Jul 24, 2016.

Address correspondence to Dr Cree, MS Center at UCSF, Neurology, 675 Nelson Rising Lane, Box 3206, San Francisco, CA 94158. E-mail: Bruce.Cree@ucsf.edu

From the [1]Department of Neurology, University of California, San Francisco, San Francisco, CA; [2]Department of Neurology, University of Texas Southwestern Medical Center, Dallas, TX; [3]Department of Neurology, University of Southern California, Los Angeles, CA; and [4]Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL

Additional Supporting Information may be found in the online version of this article.

Disease-modifying therapies (DMTs) for multiple sclerosis (MS) introduced over the past 2 decades have led to improved outcomes over the short term; however, whether the long-term prognosis has changed is not known. Natural history studies from the pretreatment era suggest that between one-third and one-half of patients will experience an insidious worsening (progression) of neurological disability approximately 15 years after onset.[1–5] Although MS relapses may produce permanent neurological impairments,[6] severe disability generally occurs in patients with progressive forms of MS (PMS) that typically develop either after an earlier relapsing phase (ie, secondary progressive MS [SPMS]) or less commonly from disease onset (ie, primary progressive MS [PPMS]).[7]

Because the goal of therapy in MS is to prevent or postpone long-term disability, short-term outcomes that are commonly used in both daily practice and in clinical trials require validation in prospective observational studies as surrogates for long-term disability. Magnetic resonance imaging (MRI) sequences that visualize and quantify focal inflammation and white matter scarring associated with relapsing–remitting MS (RRMS) are highly sensitive markers for clinical events,[8] and quantitation of brain atrophy has shown promise for monitoring the neurodegeneration associated with PMS.[9] However, the relationship between short-term MRI measurements and long-term disability is not well established.

The EPIC (expression/genomics, proteomics, imaging, and clinical) study comprises a single-center prospective observational cohort of MS patients who have been evaluated annually since July 2004. In this report, we characterize the long-term disease course in this contemporary, actively treated MS cohort. We also assess whether clinical and radiologic features at baseline and their change over 2 years, measures commonly used as outcomes in randomized clinical trials, have predictive value for long-term MS disability.

Patients and Methods

Subject Enrollment and Retention

Patients (age = 18–65 years) evaluated at the Multiple Sclerosis Center at the University of California, San Francisco (UCSF) between July 2004 and September 2005 were invited to participate. Ambulatory subjects and those with a recent onset of clinically definite MS (2001 International Panel Diagnostic Criteria)[10] or clinically isolated syndrome (CIS) were preferentially recruited, although individuals with all clinical subtypes of the disease participated. CIS was defined as an initial clinical demyelinating event with findings on brain MRI consistent with MS.[8] Subjects were excluded if they were unable to tolerate MRI scans, had poor venous access, or had other significant medical illnesses that might interfere with the goals of the study. Enrollment of subjects who had experienced a clinical relapse or received treatment with glucocorticoids within the previous month was delayed by 30 days so that the baseline MRI scans were not reflective of recent disease activity or influenced by glucocorticoid use. The use of DMTs for MS was permitted. To simplify the analysis, we grouped together patients with CIS and RRMS as a single group (RMS). During the course of the study, diagnostic criteria for MS evolved so that many patients initially designated as CIS would now be classified as MS.[11] Similarly, we grouped subjects with PPMS and SPMS together in a single category of PMS, reflecting the potential common histopathologic and genetic basis for these subtypes.[12,13] At baseline, a comprehensive neurological assessment with brain MRI and blood sample acquisition for biomarkers and genomics was performed.[14,15] Thereafter, subjects were followed annually for 5 years, and underwent reevaluation at extended time points up to 10 years after baseline. Annual study visits occurred within a $\pm$3-month window. The Committee on Human Research at UCSF approved the protocol, and informed consent was obtained from all participants.

Clinical Assessments

Disability progression was defined by clinically significant worsening in the Expanded Disability Status Scale (EDSS), the timed 25-foot walk (T25W), the 9-hole peg test (9HPT), and the paced serial auditory addition test (PASAT-3).[15,16] A clinically significant change in EDSS was defined according to the baseline EDSS score. To reduce the inherent noise for transitions between EDSS scores of 0 and 1.0, a 1.5-point or greater increase in the EDSS score was required for subjects with a baseline EDSS score of 0, a 1.0-point or greater increase for scores between 1.0 to 5.0, and a 0.5-point or greater increase for scores greater than 5.0.[17]

A 20% increase or greater in the T25W[18] and 9HPT[19] were considered to be clinically significant. The reliable change index was used for determining clinically significant worsening in PASAT-3 scores.[20] Although these outcomes are normally distributed, we chose to model their worsening based on the more stringent thresholds for clinically impactful worsening. A combined metric of any clinically significant change on the T25W, the 9HPT, or the PASAT-3 was also used. SPMS was defined by development of irreversible worsening of MS disability (increase in EDSS over at least a 1-year duration) independent of relapses in the subgroup of patients with RMS.[21] To control for possible transient worsening of disability, we also confirmed worsening of disability outcomes at year 5 that persisted to the most recent follow-up visit.

DMTs

Subjects were treated with US Food and Drug Administration (FDA)-approved therapies. In some cases, off-label therapies were also used. To reduce the complexity of treatment data, we defined 2 treatment tiers, grouping therapies together based on their relative perceived effectiveness using data from clinical trials.

The first tier was referred to as "platform therapy" that included: interferon (IFN) beta-1b, IFN beta-1a intramuscularly,

IFN beta-1a 3 times per week, and glatiramer acetate. Also included in this group were several off-label therapies: monthly pulsed dose glucocorticoids, azathioprine, and mycophenolate mofetil, which were used in a few study subjects despite these agents being of unproven benefit.[22–24] A noninferiority clinical trial showed that azathioprine had comparable efficacy to IFN therapy.[22] There are fewer available efficacy data for mycophenolate mofetil; however, we estimated the impact of this antimetabolite based on open-label observations[23] as well as that of teriflunomide, which has a related mechanism of action, and is an FDA-approved MS therapy with efficacy similar to that of IFN beta-1a.[25]

The second tier was referred to as high-potency therapy and included natalizumab, rituximab, mitoxantrone, and cyclophosphamide. Limited comparative data suggest that these treatments are more efficacious than platform therapies.[26–29] Therapeutic escalation was defined as changing treatment between baseline and the year 2 follow-up point, either from no treatment to platform therapy or from platform therapy to high-potency therapy; patients whose treatment was escalated during year 3 based on the year 2 data were also included in this group. Because fingolimod, dimethyl fumarate, and teriflunomide only became available in the past few years, these therapies did not contribute to the analysis of therapeutic escalation over the first 2 years of the study.

Brain MRI Scans

Annual brain MRI scans were acquired on the same 3T GE scanner (GE Medical Systems, Milwaukee, WI) with standardized head positioning and pulse sequences that included: high-resolution T1-weighted volume (inversion recovery spoiled gradient-echo, repetition time [TR]/echo time [TE]/inversion time [TI] = 7/2/400 milliseconds, flip angle = 8°, resolution = 0.94 × 0.94 × 1mm) with and without gadolinium–diethylenetriamine pentaacetic acid (DPTA); and T2-weighted volume (fast-recovery fast spin-echo [FRFSE], TR/TE = 2,000/81 milliseconds, resolution = 0.47 × 0.47 × 3mm). Proton density–weighted images were acquired from baseline to year 4 (FRFSE, TR/TE = 2,000/20 milliseconds, resolution = 0.47 × 0.47 × 3mm), and fluid-attenuated inversion recovery images (fast spin-echo, TR/TE/TI = 9,000/126/2,200 milliseconds, resolution = 0.47 × 0.47 × 3mm) were acquired thereafter. The T2- and T1-weighted images were used to determine MS lesion borders using semiautomated lesion segmentation software (Amira [FEI, Hillsboro, OR] and Lesion Segmentation Toolbox [Structural Brain Mapping Group, Jena, Germany]). Lesion masks for each time point were created. The lesion masks were then used to subtract MRI lesions from the T1-acquired images. The masked T1-weighted images were used to segment gray matter and white matter structures for volumetric analyses (FreeSurfer). The MS lesion masks were also used to determine the T2 lesion volume (the radiologic burden of disease). Gadolinium-DPTA was administered for the T1 plus contrast-enhanced scans, and a neuroradiologist determined the number of gadolinium-enhanced lesions and interpreted all MRI scans to insure safety.

Composite Predictors

The composite predictor, no evidence of disease activity (NEDA), was defined as no relapses, no clinically significant increase in EDSS, no new or enlarging T2 lesions, and no gadolinium-enhanced lesions on brain MRI examinations from baseline through the second year of the study. This measure is similar to that used to assess therapeutic efficacy in randomized controlled trials; however, several features are notably different.[30] First, as described above, a clinically significant change in the EDSS was defined to limit inherent noise in the EDSS. In addition, once the change in EDSS had occurred, it had to be maintained throughout the remainder of the 2-year evaluation period (in contrast to the 3- or 6-month sustained changes that are used in most clinical trials). Lastly, patient-reported relapses were included.[31]

Laboratory Studies

Blood samples were banked for biomarker studies at each visit. DNA from peripheral blood mononuclear cells was used for genome-wide association studies and for high-resolution sequence-based typing of the *HLA-DRB1* gene.[12] The MS genetic burden was determined using single nucleotide polymorphisms from 88 validated MS susceptibility loci as previously described.[32] Vitamin D levels were assessed in batch from stored samples using the DiaSorin LIAISON total 25-OH vitamin D chemiluminescence assay (Heartland Assays, Ames, IA). 25-OH vitamin D levels assessed at baseline, year 1, and year 2 were deseasonalized[33] and averaged to determine the mean 25-OH vitamin D during the first 2 years of the study.

Statistics

All statistical analyses were computed using code written in R (r-project.org). Survival analysis was used to generate Kaplan–Meier estimates for time to EDSS = 6 and SPMS in the subgroup of patients meeting criteria for clinically definite MS (CDMS).[34] Logistic regression was used to determine whether baseline clinical, radiologic, and genetic features of this cohort correlated with long-term disability, and to model clinical and MRI changes from baseline to year 2 and escalation therapy as predictor variables for long-term disability outcomes. To maintain homogeneity of cohort time for the logistic regression analysis, subjects without year 10 follow-up visits were removed. Propensity scores for treatment at baseline were developed using the following clinical and MRI variables: gender, age of onset, baseline disease duration, baseline EDSS, prestudy annualized relapse rate, prestudy medication possession ratio (the proportion of time of treatment with a DMT from clinical onset to baseline), baseline T2 lesion volume (T2LV), and baseline brain volume. The MRI variables of T2LV and brain volume loss were used as a proxy for imaging severity that neurologists might have evaluated prior to study entry. In contrast, the number of gadolinium-DPTA lesions at the baseline scan would not be known prior to enrollment and therefore was not included in the propensity score. A propensity score for the baseline treatment tier was included as a covariate for all analyses that assessed the impact of baseline to year 2 predictors on long-

term outcomes. A multivariate analysis was also developed using variables selected for by lasso (least absolute shrinkage and selection operator, an L1-constrained shrinkage and selection method) and cross-validation. Cross-validation was performed 50 times, and the variables selected for at least 45 times were included. Thus, our models adjusted for both baseline treatment and therapeutic escalation. Scores were computed separately on RMS and PMS patients (stratification by clinical course).

Results

Demographic Characteristics

A total of 517 subjects were enrolled: 366 had RMS, 48 SPMS, 21 PPMS, and 82 CIS (Supplementary Table 1). A total of 489 subjects completed year 2, and year 10 follow-up data were available on 471 of 517 subjects (91%). Of those patients with long-term follow-up, the median time in the cohort was 9.8 years since enrollment (9.9 years excluding subjects who died) and was 16.8 years since disease onset. The baseline characteristics of subjects for whom a recent clinical evaluation was available are summarized in Table 1 (n = 471), including individuals with RMS (n = 407) and PMS (n = 64). Data for the entire cohort are summarized in Supplementary Table 1. The baseline characteristics of subjects who were retained in the study compared to those lost to follow-up were generally similar; the annualized relapse rate was slightly higher and disease duration was somewhat shorter in the lost to follow-up group (Supplementary Tables 2–4). More than half of the cohort (246 patients [52%]) had a low EDSS (0–1.5) at baseline. The EDSS score distribution was bimodal, with RMS patients having lower EDSS scores than PMS patients. Baseline characteristics of RMS subjects including gender, disease duration, age of onset, and EDSS were not significantly different from RMS patients who received care at UCSF during the same time period but did not participate in the study (Supplementary Table 5).

Clinical Outcomes

Over the 10 years of follow-up, 225 (55.3%) RMS patients experienced a clinically significant increase in the EDSS score (Fig 1A and Supplementary Table 6). Clinically significant worsening in the T25W, 9HPT, and PASAT-3 occurred less commonly for each of these individual outcomes than change in the EDSS.

RMS subjects at all levels of baseline EDSS score exhibited a roughly equivalent risk for clinically significant worsening during the subsequent 10 years (see Fig 1A). In contrast, for PMS subjects worsening occurred for >75% of subjects, and for 100% of those with baseline EDSS scores < 3 (see Fig 1B and Supplementary Tables 6 and 7). During the study period, 46 of the 407 patients (10.1%) with RMS at baseline transitioned to SPMS. Female sex was modestly associated with a lower risk of developing SPMS (odds ratio [OR] = 0.61, 95% confidence interval [CI] = 0.40–0.94, $p = 0.02$). A later age of onset of MS was also associated with an increased risk of developing SPMS (OR = 1.04, 95% CI = 1.02–1.07, $p = 0.001$ for each 10-year increase in the age of onset). At a median time of 16.8 years after disease onset, 10.7% (95% CI = 7.2–14%) of patients reached an EDSS ≥ 6 and 18.1% (95% CI = 13.5–22.5%) evolved from RMS to SPMS. We estimated that only 4.7% (95% CI = 2.6–6.8%) of relapse-onset patients reached an EDSS ≥ 6 at 10 years after disease onset and 16.2% (95% CI = 11.5–20.7%) after 20 years (Fig 2A). The risk of transition to SPMS was 6.4% (95% CI = 4–8.8%) 10 years after onset and 24.2% after 20 years (95% CI = 18.5–29.6%; see Fig 2B).

Clinical and Radiologic Predictors of Disability Progression

Associations between baseline characteristics, treatment escalation, and other variables on the development of long-term disability are presented in Tables 2 and 3. The baseline to year 2 predictor analyses are presented with and without propensity score adjustment in Supplementary Tables 8 and 9. The great majority of RMS patients (n = 334; 82.1%) experienced clinical and/or MRI disease activity during the first 2 years of the study (Supplementary Tables 10 and 11). Only 73 RMS patients (17.9%) satisfied combined clinical and radiologic criteria of NEDA. NEDA at year 2 was not associated with statistically significant EDSS outcomes at year 10 and, contrary to expectations, the NEDA group showed a trend toward more, rather than less, worsening in EDSS score over the long term (OR = 1.42, $p = 0.189$; see Supplementary Tables 8 and 9).

We also determined whether disease activity over a 2-year period measured solely by radiologic criteria had effects on 10-year outcomes. The development of new or enlarging T2 lesions from baseline to year 2 was not associated with subsequent clinical worsening as measured by EDSS, T25W, 9HPT, or PASAT-3. Importantly, this was also true in the subgroup of RMS patients (n = 67) who were clinically inactive but had had new or enlarging T2 lesions during years 0 to 2 (EDSS OR = 1.55, 95% CI = 0.91–2.65, $p = 0.104$). Thus, we were unable to identify any consequential effect of early MRI disease activity on 10-year clinical outcomes.

In terms of clinical variables, an increase in EDSS during years 0 to 2 was paradoxically associated with a lower, rather than a higher, risk of subsequent worsening ($p = 2.62 \times 10^{-5}$; see Supplementary Tables 8 and 9).

TABLE 1. Baseline Clinical and MRI Features of Subjects Completing Long-Term Follow-up

Characteristic	All, n = 471	RMS, n = 407	PMS, n = 64	*p*
Demographic				
Age at exam, mean ± SD	42.7 ± 9.9	41.7 ± 9.7	48.6 ± 8.7	1.22e-07
Sex, No. (%)				
Women	318 (67.5)	280 (68.8)	38 (59.4)	0.151
Men	153 (32.5)	127 (31.2)	26 (40.6)	0.151
Years of follow-up, MIR[a]	9.8 [8.6, 10.2] {1–11.5}	9.9 [8.6, 10.2] {1–11.5}	9.3 [8.6, 10.2] {1–11.2}	0.051
Clinical				
Age of onset, mean ± SD	33.3 ± 9.3	33.4 ± 9.2	32.6 ± 10.2	0.548
Disease duration, MIR	7 [2, 13.5] {0–46}	6 [2, 12] {0–46}	15 [7, 22.2] {1–45}	3.53e-10
Disease course, No. (%)				
CIS	70 (14.9)	70 (17.2)		
RR	337 (71.5)	337 (82.8)		
SP	45 (9.6)		45 (70.3)	
PP	19 (4)		19 (29.7)	
EDSS score, MIR	1.5 [1, 3] {0–7}	1.5 [1, 2] {0–6.5}	4.5 [3.5, 6] {1.5–7}	9.19e-28
MSSS, MIR	2.4 [0.9, 4.3] {0–9.8}	2.1 [0.7, 3.7] {0–9.5}	5.2 [3.4, 7.2] {0.8–9.8}	9.43e-15
Relapse history				
Annualized relapse rate, MIR	0.5 [0.2, 1] {0–7.3}	0.5 [0.3, 1.1] {0–7.3}	0.2 [0.1, 0.4] {0–1.1}	4.9e-09
Vitamin D level, ng/ml, mean ± SD	24.4 ± 8.8	24.4 ± 8.7	24.1 ± 9.5	0.808
Treatment				
Treatment history, No. (%)				
No treatment	183 (38.9)	155 (38.1)	28 (43.8)	0.41
Platform therapy	281 (59.7)	247 (60.7)	34 (53.1)	0.274
High potency	7 (1.5)	5 (1.2)	2 (3.1)	0.244
Years to first treatment from diagnosis, MIR	3.1 [0.8, 8.7] {0–43.9}	2.8 [0.7, 7.6] {0–43.9}	6.4 [3.3, 13.6] {0–36.4}	3.86e-05
Medication possession ratio, prestudy, MIR	0.2 [0, 0.6] {0–1}	0.2 [0, 0.6] {0–1}	0.3 [0, 0.5] {0–0.9}	0.979
MRI				
T2 lesion volume, ml, MIR	2.7 [0.8, 6.8] {0–103.9}	2.4 [0.7, 5.7] {0–103.9}	7 [2.1, 12] {0–71.7}	1.42e-05
Number of gad enhancing lesions, MIR	0 [0, 0] {0–10}	0 [0, 0] {0–9}	0 [0, 0] {0–10}	0.561
Total brain volume, ml, mean ± SD	1,460.1 ± 87.7	1,470.1 ± 83	1,396.2 ± 90.7	4.02e-08
Gray matter volume, ml, mean ± SD	790.8 ± 58.9	797.5 ± 56.5	748 ± 56	5.14e-09
White matter volume, ml, mean ± SD	669.3 ± 42.5	672.6 ± 41	648.2 ± 45.9	1.49e-04
Ventricular CSF volume, ml, MIR	41 [30, 55] {10–172}	39 [29.5, 51] {10–172}	55 [39, 71] {15–134}	1.43e-07
Cortical gray matter volume, ml, mean ± SD	626.6 ± 48.4	632.2 ± 46.3	591 ± 46.3	4.15e-09

Subjects completing long-term follow-up include subjects with a year 10 visit and deceased subjects. 25-OH vitamin D levels are deseasonalized. Probability values compare RMS and PMS subjects. For normally distributed data, mean and SD are shown and Student *t* test was used. For data that are not normally distributed, median, interquartile, and range are shown and a Wilcoxon test was used. For qualitative data, counts and percentages are shown and Fisher exact test was used.

[a]These ranges include deceased subjects. MIR = median [IQR] {range}.

CIS = clinically isolated syndrome; CSF = cerebrospinal fluid; EDSS = Expanded Disability Status Scale; MRI = magnetic resonance imaging; MSSS = Multiple Sclerosis Severity Score; PMS = progressive multiple sclerosis; PP = primary progressive; RMS = CIS and RRMS as a single group; RR = relapsing–remitting; SD = standard deviation; SP = secondary progressive.

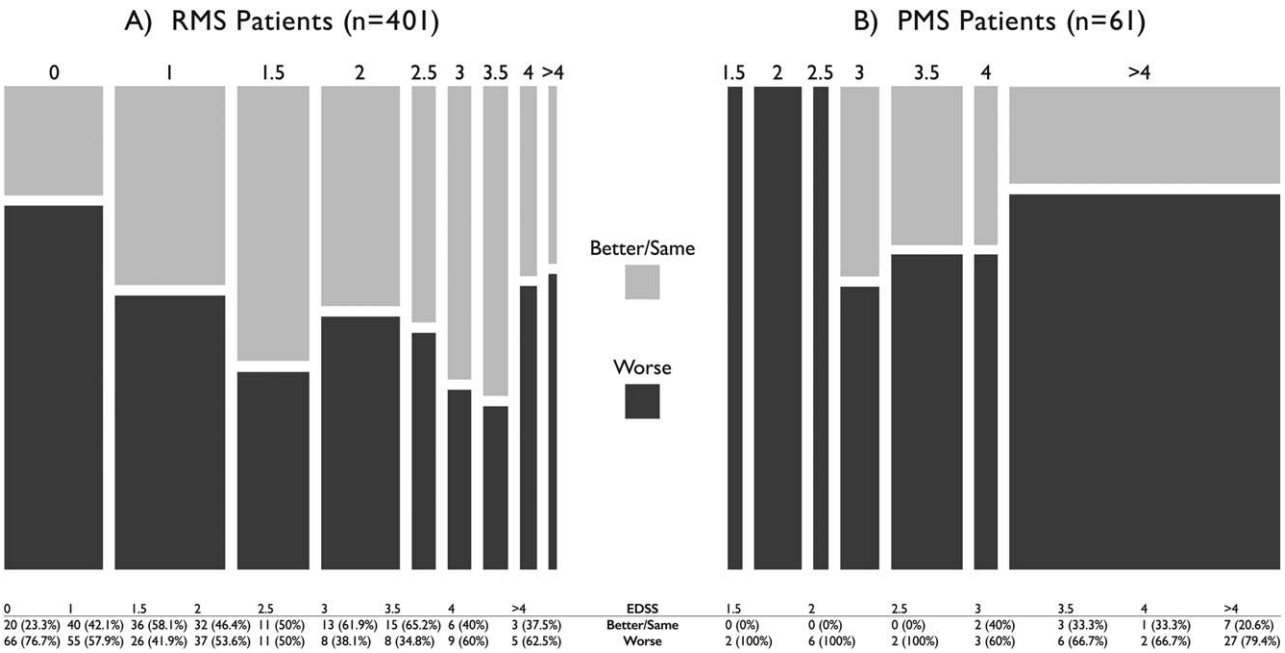

FIGURE 1: Increase in Expanded Disability Status Scale (EDSS) at last visit by baseline EDSS score. For the subjects who died of multiple sclerosis (MS), an EDSS score of 10 was assigned. For subjects who became disabled due to non-MS causes, the last EDSS score was carried forward. There is 1 patient with a visit at year 10 for whom an EDSS score is missing. For the 8 subjects who died due to non-MS causes, year 10 EDSS scores were not available. The width of the bar represents the proportion of subjects in the cohort with each baseline EDSS score. The height of the black and gray bars represents the relative proportion of subjects with worse scores (black) versus stable or improved scores (gray). PMS = progressive MS; RMS = clinically isolated syndrome and relapsing–remitting multiple sclerosis as a single group.

This association was accounted for in part by prolonged recovery from relapses (n = 49, OR = 0.355) and, possibly, a response to treatment escalation (n = 6, OR = 0.532). For relapsing patients who developed SPMS, worsening during years 0 to 2 was associated, as expected, with further worsening (n = 16, OR = 2.59). Therefore, at least in this data set, understanding changes in EDSS scores as a predictor of long-term further worsening required contextualizing these changes with respect to relapses and disease course.

Serum levels of 25-OH vitamin D were associated with risk of focal disease activity (Supplementary Table 12)[35,36]; however, the average 25-OH vitamin D levels over the first 2 years of observation had no association with long-term disability outcomes (see Supplementary Tables 8 and 9).[37] These results are consistent with the clinical and radiologic findings indicating that disease activity during years 0 to 2 did not measurably impact clinical outcomes at year 10.

Influence of CIS Subjects

Because we grouped CIS together with RRMS subjects, it is possible that at least some of the CIS subjects did not have MS and therefore could bias the cohort toward benign disease. Of the 82 CIS subjects, 34 experienced a second clinical attack or developed SPMS. An additional 22 subjects developed new MRI lesions and fulfilled International Panel criteria for radiographic dissemination over time.[11] Ten CIS subjects were lost to follow-up and did not contribute to the long-term outcome data. Sixteen CIS subjects did not experience clinical or radiographic dissemination over time. We performed a sensitivity analysis that excluded these 16 stable CIS subjects (with and without propensity score adjustment in Supplementary Tables 13 and 14) and concluded that our observations were not biased by including CIS subjects.

Potential Impact of Informative Censoring

The lost to follow-up RMS patients had a shorter disease duration, higher Multiple Sclerosis Severity Score (MSSS) scores, and a slightly higher annualized relapse rate (see Supplementary Table 3). These factors might contribute to disease worsening, resulting in informative censoring, thereby slightly biasing the retained cohort to have milder disease. Twelve subjects (9 RMS and 3 PMS) experienced worsening in function postbaseline but did not complete a year 10 visit. Because excluding these subjects might bias the cohort in favor of a more benign prognosis, we reanalyzed the data set under the assumption that these subjects experienced sustained worsening at year 10 (see Supplementary Tables 7, 9, 11, and 14). Excluding these subjects did not influence our conclusions, with the exception of reducing the potential

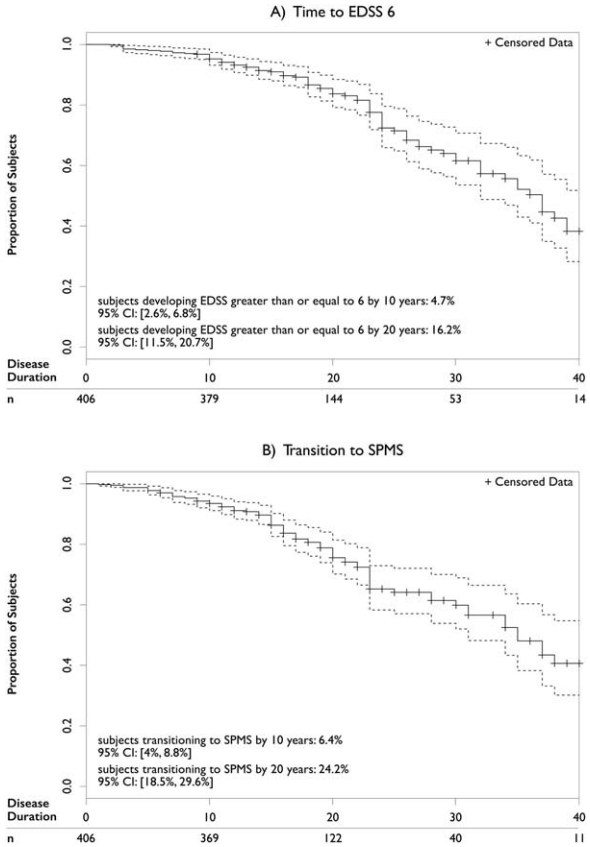

FIGURE 2: Kaplan–Meier curves for time to Expanded Disability Status Scale (EDSS) = 6 and time to transition to secondary progressive multiple sclerosis (SPMS) for clinically isolated syndrome (CIS), relapsing–remitting, and secondary progressive subjects with long-term follow-up who met criteria for clinically definite MS either at baseline or during the study (n = 406). Kaplan–Meier survival curves for time to EDSS = 6 (ambulatory impairment requiring a cane to walk 100m) and time to onset of secondary progressive MS (insidious deterioration of neurological function that is not relapse related) are shown. Disease duration in years is graphed on the x-axis and the percentage of the cohort at each time point that has not met the failure events is depicted on the y-axis. For this analysis, subjects in our cohort were limited to those meeting criteria for clinically definite MS.[34] Subjects who had 1 relapse and who met International Panel criteria for dissemination over time by magnetic resonance imaging criteria at baseline (n = 19) but did not experience further clinical events over the course of follow-up were excluded (n = 10). Similarly, CIS subjects who did not experience further clinical relapses or disease progression (n = 48) were excluded. If subjects had already met the endpoint at the time of entry into the study (ie, subjects with EDSS = 6 or subjects with SPMS at baseline), time to EDSS = 6 or time to SPMS was obtained through systematic review of medical records and a standardized questionnaire administered to all subjects at study entry. CI = confidence interval.

impact of baseline to year 2 changes in gray and white matter volumes on 9HPT worsening. We also performed additional sensitivity analyses assuming that the relapsing subjects who were lost to follow-up, and had EDSS scores < 6 at the time of their last documented visit, worsened such that at the next hypothetical visit all these subjects reached EDSS = 6. This worst-case scenario resulted in a change in the median time to EDSS = 6 of only 2 years (from 37 years to 35 years, log-rank test, $p = 0.036$). A similar sensitivity analysis was performed for SPMS, revealing that the median time to SPMS was decreased from 35 years to 34 years (log-rank test, $p =$ not significant).

Effect of Treatment Escalation

Subjects who experienced treatment escalation were more likely to have had clinical relapses in the prior year (OR = 1.9, 95% CI = 1.03–3.56, $p = 0.04$) and to have experienced brain parenchymal volume loss (OR = 2.63, 95% CI = 1.38–5.33, $p = 0.005$). Long-term outcomes for patients whose treatment was escalated were not different from those for patients whose treatment was not escalated (see Supplementary Tables 8 and 9). Multivariate analysis that accounted for treatment at baseline, therapeutic escalation, and treatment tier (as a time-dependent covariate during the course of the study), using a propensity score adjusted–model, did not reveal additional significant associations, interactions, or potential sources of confounding.

Discussion

Long-term disability worsening was measured in a large, prospectively followed cohort of MS patients at a single tertiary referral center over a 10-year period. Nearly half experienced no clinically significant disability worsening throughout the duration of the study, as measured by the global disability measure EDSS or in tests of walking (T25W), upper limb (9HPT), and cognitive (PASAT-3) function. Reflecting a real-life population of early MS patients, most individuals were actively treated. Escalation to higher potency therapies was a common occurrence during the course of the study. Although we did not find that treatment escalation reduced the risk of further disability progression, it is possible that active management with DMTs influenced the overall favorable outcomes. Patients experiencing clinical relapses or radiologic worsening were more likely to undergo treatment escalation. Nonetheless, clinically significant disability accrued in 59% of subjects, illustrating a remaining unmet need for more effective DMTs in RMS, and any effective therapy for progressive MS.

Earlier natural history studies found that up to 54% of RMS patients transitioned to SPMS after a median time of 19 years.[38–40] With a median time of 16.8 years since disease onset, we would have anticipated that between 36% and 50% of the RMS patients would have developed SPMS, whereas only 11.3% of this

TABLE 2. Univariate Analysis of Clinical Outcomes from Baseline to Last Visit: All RMS Subjects with Long-Term Follow-up, n = 407

Response	Predictor	RMS Subjects		
		OR	95% CI	*p*
EDSS worsening	Baseline ARR	0.75	0.6, 0.93	0.011
	Baseline EDSS	0.76	0.64, 0.89	9.5e-04
	Baseline white matter volume, dl	0.54	0.33, 0.88	0.015
PASAT worsening	Baseline EDSS	1.38	1.04, 1.81	0.021
	Baseline brain volume, dl	0.5	0.32, 0.77	0.002
	Baseline gray matter volume, dl	0.43	0.22, 0.85	0.016
	Baseline white matter volume, dl	0.25	0.1, 0.63	0.003
	Baseline cortical gray matter volume, dl	0.39	0.17, 0.89	0.027
	Baseline gadolinium lesion count	1.31	0.98, 1.7	0.045
T25W worsening	Baseline age at exam	1.04	1.01, 1.06	0.01
	Baseline EDSS	1.64	1.33, 2.03	4e-06
	Baseline disease duration	1.04	1.02, 1.07	0.002
	Baseline T25W	1.3	1.07, 1.61	0.014
	Baseline 9HPT	1.11	1.05, 1.17	1.84e-04
	Baseline brain volume, dl	0.62	0.45, 0.84	0.003
	Baseline gray matter volume, dl	0.49	0.3, 0.79	0.004
	Baseline white matter volume, dl	0.51	0.27, 0.95	0.036
	Baseline cortical gray matter volume, dl	0.4	0.22, 0.7	0.002
	Baseline CSF fluid volume, dl	4.68	1.3, 17.27	0.018
9HPT worsening	Baseline EDSS	1.43	1.12, 1.83	0.004
	Baseline gray matter volume, dl	0.53	0.29, 0.97	0.043
	Baseline cortical gray matter volume, dl	0.4	0.19, 0.84	0.017

Univariate regression on 4 clinical outcomes for RMS subjects: EDSS worsening from baseline to year 10, PASAT worsening from baseline to year 10, T25W worsening from baseline to year 10, and 9HPT worsening from baseline to year 10. Predictors were tested for each of these outcomes. The predictors tested include gender, age of multiple sclerosis onset, baseline age at examination, baseline ARR, medication possession ratio from disease onset to baseline, baseline EDSS score, baseline disease duration, baseline PASAT score, baseline T25W, baseline 9HPT, baseline brain volume, baseline gray matter volume, baseline white matter volume, baseline cortical gray matter volume, baseline CSF volume, baseline T2 lesion volume, and baseline gadolinium-enhanced lesion count. Associations with probability values < 0.05 are shown.
9HPT = 9-hole peg test; ARR = annualized relapse rate; CI = confidence interval; CSF = cerebrospinal fluid; EDSS = Expanded Disability Status Scale; OR = odds ratio; PASAT = paced serial auditory addition test; RMS = clinically isolated syndrome and relapsing–remitting multiple sclerosis as a single group; T25W = 25-foot walk.

cohort transitioned to SPMS during the course of the study. This transition rate of 1% annually is lower than reported from natural history studies but similar to a more recent retrospective analysis in an IFN-treated population.[41] Similarly, evolution of sustained disability in this cohort was slower than expected. At 16.8 years after onset, 10.7% (95% CI = 7.2–14%) of patients had reached an EDSS ≥ 6, whereas in some natural history studies 50% of the cohort had reached an EDSS = 6 by 15 to 16 years, albeit it with wide confidence intervals.[5,42]

We also assessed the predictive value of MRI metrics commonly measured in MS clinical trials over 2-year intervals, and found no association between new T2 lesions or gadolinium-DPTA–enhanced lesions and worse long-term outcomes. Earlier reports suggested that 2 or

TABLE 3. Univariate Analysis of Clinical Outcomes from Baseline to Last Visit: Subjects with Long-Term Follow-up Excluding Clinically and Radiographically Stable CIS Subjects, n = 391

Response	Predictor	RMS Subjects		
		OR	95% CI	*p*
EDSS worsening	Baseline ARR	0.75	0.6, 0.94	0.013
	Baseline EDSS	0.76	0.64, 0.89	0.001
	Baseline white matter volume, dl	0.58	0.34, 0.97	0.038
PASAT worsening	Baseline EDSS	1.34	1.01, 1.77	0.037
	Baseline brain volume, dl	0.52	0.32, 0.81	0.005
	Baseline gray matter volume, dl	0.49	0.24, 0.97	0.044
	Baseline white matter volume, dl	0.23	0.09, 0.6	0.003
	Baseline gadolinium lesion count	1.32	0.99, 1.72	0.039
T25W worsening	Baseline age at exam	1.04	1.01, 1.07	0.01
	Baseline EDSS	1.65	1.34, 2.05	3.99e-06
	Baseline disease duration	1.04	1.02, 1.07	0.002
	Baseline T25W	1.3	1.07, 1.62	0.015
	Baseline 9HPT	1.11	1.05, 1.18	1.43e-04
	Baseline brain volume, dl	0.64	0.46, 0.88	0.006
	Baseline white matter volume, dl	0.51	0.31, 0.82	0.006
	Baseline cortical gray matter volume, dl	0.41	0.22, 0.73	0.003
	Baseline CSF fluid volume, dl	5	1.37, 18.75	0.015
9HPT worsening	Baseline EDSS	1.45	1.13, 1.87	0.004
	Baseline cortical gray matter volume, dl	0.45	0.2, 0.94	0.037

Univariate regression on 4 clinical outcomes for RMS subjects: EDSS worsening from baseline to year 10, PASAT worsening from baseline to year 10, T25W worsening from baseline to year 10, and 9HPT worsening from baseline to year 10. Predictors were tested for each of these outcomes. The predictors tested include gender, age of multiple sclerosis onset, baseline age at examination, baseline ARR, medication possession ratio from disease onset to baseline, baseline EDSS score, baseline disease duration, baseline PASAT score, baseline T25W, baseline 9HPT, baseline brain volume, baseline gray matter volume, baseline white matter volume, baseline cortical gray matter volume, baseline CSF volume, baseline T2 lesion volume, and baseline gadolinium-enhanced lesion count. Associations with probability values < 0.05 are shown.
9HPT = 9-hole peg test; ARR = annualized relapse rate; CI = confidence interval; CIS = clinically isolated syndrome; CSF = cerebrospinal fluid; EDSS = Expanded Disability Status Scale; OR = odds ratio; PASAT = paced serial auditory addition test; RMS = clinically isolated syndrome and relapsing–remitting multiple sclerosis as a single group; T25W = 25-foot walk.

more new gadolinium-DPTA–enhanced lesions or new T2 lesions without any clinical correlate carry a negative prognostic value in actively treated MS patients.[43–45] These conclusions are not supported by the present study. Attrition bias might have contributed to the earlier results, and we believe that the current study with follow-up data available on >91% of subjects is largely free from this confounder. Our observations are similar to those of the long-term follow-up study from the pivotal IFN beta-1b study, a study also largely free of attrition bias,[46] and a population-based, contemporaneous study conducted in the treatment era.[47] Our observations suggest that radiographic markers of focal inflammation (gadolinium-DPTA–enhanced lesions or new or enlarging T2 lesions) may only carry short-term associations with clinical events. However, the effectiveness of therapy may contribute to misclassification and/or reduction in predictive validity.

Similarly, an increase in the EDSS score over 2 years was not associated with worse long-term prognosis. Thus, short-term increases in EDSS do not necessarily predict future accumulation of disability in RMS patients over the longer term, a conclusion also reached in a pooled analysis of patients randomized to placebo arms

in 31 clinical trials.[48] Because neither clinical nor radiographic features over 2 years had predictive value, it is not surprising that the combined measure of these variables, NEDA, was also not associated with long-term disability risk. Although this observation must be interpreted with caution because of the relatively small number of NEDA patients in our cohort, another recently published observational study also found that the proportion of patients meeting a NEDA definition declined substantially over time.[49] These observations challenge the concept that NEDA represents remission. Although NEDA may be a useful measure for assessing relative therapeutic efficacy, many patients who meet NEDA criteria over 2 years go on to develop clinically significant disability. Worsening in patients who meet the 2-year NEDA endpoint could result from active spinal cord disease not captured with brain MRI, progressive axonal or neuronal degeneration, or an escape from a true but transient remission state. A recent study that incorporated thresholds for acceptable brain volume loss for NEDA found that one-third of NEDA patients treated with fingolimod still experienced significant brain volume loss during the NEDA interval, indicating that ongoing tissue injury occurs in NEDA patients.[50]

It is possible that escalation to high-potency therapy might have reduced disability that otherwise would have accrued in the RMS cohort, although we did not observe a favorable impact of escalation therapy itself on long-term disability. More likely, an interval of 2 years is too short to have any significant long-term predictive value. In this regard, we found that brain volume at baseline was predictive of long-term PASAT-3 performance yet the change in brain volume over 2 years was not. This suggests that measuring brain volume change over longer periods of time, perhaps 3 or 4 years, might have predictive value for long-term cognitive function, a testable hypothesis. Some clinical outcomes, such as the T25W and 9HPT, are insensitive to change over a 2-year interval, whereas change in other outcomes such as the EDSS is not predictive because of a well-recognized inherent variability. Our study of 404 relapsing MS patients could also be underpowered to detect weak effects of these clinical measures on long-term MS disability. Despite all these caveats, our observations call into question the prevailing assumption that commonly used clinical and MRI markers of MS activity as measured over the 2-year duration of many MS clinical trials are a sufficient proxy for long-term disability.

Two-year observations also indicated that clinically silent MRI activity was not associated with worse outcomes over the long term, a finding consistent with results from a large meta-analysis of placebo-treated patients enrolled in MS clinical trials.[51] These data argue that the common practice of obtaining routine surveillance MRI scans may have limited added value in the setting of otherwise quiescent MS.[52] Our observations challenge the notion that one should use MRI in a treat-to-target paradigm. Newer MRI sequences, including spinal cord measures not routinely measured in clinical practice, might provide a more robust measure of disability risk. For example, cross-sectional data demonstrated that EDSS is highly correlated with gray matter volume in the cervical and thoracic cord, independent of brain volume.[53,54]

The finding that levels of 25-OH vitamin D measured during the first 2 years of the study were associated with new focal MS lesions as expected,[35,36] but not with long-term disability, provides additional support for the conclusion that short-term changes in MS disease activity do not necessarily associate with favorable long-term outcomes.

Our data have limitations that must be acknowledged. This single-center observation cohort design is fundamentally different from population-based epidemiological studies. Unlike studies of MS natural history, all subjects in this study provided written informed consent and underwent a variety of a clinical, imaging, and biological assessments. Therefore, the participants in our study are inherently different in that they agreed to participate in research and therefore may experience a somewhat different evolution of their disease. Although the characteristics of the cohort are similar to those of participants in MS clinical trials, it is possible that this large group of subjects, recruited by multiple practitioners at a single center who interact with each other on a daily basis, might have been unintentionally biased toward enrolling individuals with a milder disease. This seems unlikely, for several reasons. Enrollment was encouraged for all interested patients in our clinics, and baseline characteristics of the EPIC subjects were identical to patients who chose not to participate in the study but received care at the UCSF MS Center during the recruitment period. Furthermore, during the course of the study, nearly all patients experienced active disease and PMS patients recruited during the same interval from the same patient pool relentlessly worsened as expected. Patients with tumefactive presentations, as well as patients with rapidly progressive MS, participated in this study. Several patients who had rapidly progressive MS, unfortunately, died from MS during the observation period. Thus, the population of MS patients from which this data set was drawn appears to be representative of MS patients receiving care, certainly at our institution, and could perhaps reflect a changing face of MS in general.

Another limitation is that the data set is only moderate in size, and that analysis of subgroups of interest

could be limited by insufficient statistical power. Replication will be required before changes to current clinical practice can be recommended. Nine percent of subjects were lost to follow-up, and although generally similar to those retained in the study, those lost to follow-up had a shorter disease duration, had higher MSSS scores, and had a slightly higher annualized relapse rate (see Supplementary Table 3). These factors might contribute to disease worsening, resulting in informative censoring thereby biasing the retained cohort to have milder disease. A sensitivity analysis showed that the impact of those lost to follow-up could have an influence on the median time to EDSS (shortened by 2 years) but not on the median time to SPMS. We therefore conclude that our observations regarding the evolution of major disability milestones and secondary progression cannot be solely accounted for by a bias introduced through informative censoring. A final potential source of confounding is the variable disease duration at the time of entry into the study. Although the duration of in-study follow-up was relatively uniform, subjects entered EPIC at different times with respect to the onset of MS (as is also the case for all MS clinical trials). Adjusting for disease duration only partially accounts for this source of possible exposure time bias.

There are several important implications of these data for management of patients with MS. First, treating to target with 2-year NEDA as the goal may not result in protection against long-term disability. Second, neurological disability appears to evolve more slowly than estimated from older natural history cohorts. The availability of DMTs and escalation to higher potency therapies might account, at least in part, for the clinically important lower rates of disability accumulation and evolution to SPMS observed here. However, more than half of RMS patients treated with platform therapies still worsen over a decade of observation irrespective of short-term MRI or clinical changes. Thus, long-term studies are urgently needed to determine if high-intensity therapy, initiated at the time of diagnosis or used in patients with seemingly inactive disease, is superior to the escalation approach employed in this cohort.

Acknowledgment

The NIH - National Institute of Neurological Diseases and Stroke (RO1NS26799, S.L.H., J.R.O.; K23 NS048869, B.A.C.C.; K23 NS067055, E.M.), the Valhalla Foundation, and gifts from Friends of the Multiple Sclerosis Research Group at UCSF supported this study.

We thank the patients for participating in this demanding study.

Author Contributions

Study concept and design: S.B., B.A.C.C., P.-A.G., S.L.H., R.G.H., and J.R.O. Data acquisition and analysis: A.A., C.B., M.B., S.C., C.C., B.A.C.C., G.R.C., E.C.-H., J.M.G., R.G., D.S.G., P.-A.G., J.G., A.J.G., R.G.H., R.K., R.L., A.L., E.M., D.T.O., V.P., N.P., D.P., P.Q., A.S., W.A.S., L.S., H.-C.v.B., S.S.Z., and A.H.Z. Drafting the manuscript and figures: A.A., S.B., C.B., B.A.C.C., G.R.C., R.G., J.M.G., D.S.G., P.-A.G., J.G., A.J.G., S.L.H., R.G.H., J.R.O., W.A.S., H.-C.v.B., and S.S.Z. All authors edited and approved the final version of the manuscript.

Potential Conflicts of Interest

Companies that make MS DMTs described in this article include: Bayer, Biogen, EMD Serono, Pfizer, and Teva. The following authors disclosed financial relationships with these companies. S.B.: consultancy, EMD Serono, Teva. B.A.C.C.: consultancy, Biogen, EMD Serono, Teva. G.R.C.: consultancy, Biogen, Teva, EMD Serono, Pfizer. D.S.G.: speaking fees, EMD Serono, Teva. E.C.-H.: consultancy, Biogen, Teva. D.T.O.: consultancy and speaking fees, Teva. D.P.: consultancy, Biogen. S.S.Z.: speaking fees, Biogen.

References

1. Confavreux C, Vukusic S, Moreau T, et al. Relapses and progression of disability in multiple sclerosis. N Engl J Med 2000;343:1430–1438.
2. 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.
3. Tremlett H, Paty D, Devonshire V. The natural history of primary progressive MS in British Columbia, Canada. Neurology 2005;65:1919–1923.
4. Tremlett H, Paty D, Devonshire V. Disability progression in multiple sclerosis is slower than previously reported. Neurology 2006; 66:172–177.
5. Debouverie M, Pittion-Vouyovitch S, Louis S, et al. Natural history of multiple sclerosis in a population-based cohort. Eur J Neurol 2008;15:916–921.
6. Lublin FD, Baier M, Cutter G. Effect of relapses on development of residual deficit in multiple sclerosis. Neurology 2003;61:1528–1532.
7. Lublin FD, Reingold SC, Cohen JA, et al. Defining the clinical course of multiple sclerosis: the 2013 revisions. Neurology 2014;83:278–286.
8. O'Riordan JI, Thompson AJ, Kingsley DP, 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.
9. Rudick RA, Lee JC, Nakamura K, et al. Gray matter atrophy correlates with MS disability progression measured with MSFC but not EDSS. J Neurol Sci 2009;282:106–111.
10. McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. Ann Neurol 2001;50:121–127.
11. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 2011;69:292–302.
12. Lassmann H, van Horssen J, Mahad D Progressive multiple sclerosis: pathology and pathogenesis. Nat Rev Neurol 2012;8:647–656.

13. International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium 2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 2011;476:214–219.
14. Nickles D, Chen HP, Li MM, et al. Blood RNA profiling in a large cohort of multiple sclerosis patients and healthy controls. Hum Mol Genet 2013;22:4194–4205.
15. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33:1444–1452.
16. Rudick RA, Cutter G, Reingold S. The multiple sclerosis functional composite: a new clinical outcome measure for multiple sclerosis trials. Mult Scler 2002;8:359–365.
17. Calabresi PA, Radue EW, Goodin D, et al. Safety and efficacy of fingolimod in patients with relapsing-remitting multiple sclerosis (FREEDOMS II): a double-blind, randomised, placebo-controlled, phase 3 trial. Lancet Neurol 2014;13:545–556.
18. Kaufman M, Moyer D, Norton J. The significant change for the Timed 25-foot Walk in the multiple sclerosis functional composite. Mult Scler 2000;6:286–290.
19. Ytterberg C, Johansson S, Andersson M, et al. Variations in functioning and disability in multiple sclerosis. A two-year prospective study. J Neurol 2008;255:967–973.
20. Barker-Collo SL, Purdy SC. Determining the presence of reliable change over time in multiple sclerosis: evidence from the PASAT, Adjusting-PSAT, and Stroop Test. Int J MS Care 2013;15:170–178.
21. Lublin FD, Reingold SC. Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Neurology 1996;46:907–911.
22. Massacesi L, Tramacere I, Amoroso S, et al. Azathioprine versus beta interferons for relapsing-remitting multiple sclerosis: a multicentre randomized non-inferiority trial. PLoS One 2014;9:e113371.
23. Michel L, Vukusic S, De Seze J, et al. Mycophenolate mofetil in multiple sclerosis: a multicentre retrospective study on 344 patients. J Neurol Neurosurg Psychiatry 2014;85:279–283.
24. Zivadinov R, Rudick RA, De Masi R, et al. Effects of IV methylprednisolone on brain atrophy in relapsing-remitting MS. Neurology 2000;57:1239–1247.
25. Oh J, O'Connor PW. Teriflunomide in the treatment of multiple sclerosis: current evidence and future prospects. Ther Adv Neurol Disord 2014;7:239–252.
26. Rudick RA, Stuart WH, Calabresi PA, et al. Natalizumab plus interferon beta-1a for relapsing multiple sclerosis. N Engl J Med 2006;354:911–923.
27. Kappos L, Li D, Calabresi PA, et al. Ocrelizumab in relapsing-remitting multiple sclerosis: a phase 2, randomised, placebo-controlled, multicentre trial. Lancet 2011;378:1779–1787.
28. Vollmer T, Panitch H, Bar-Or A, et al. Glatiramer acetate after induction therapy with mitoxantrone in relapsing multiple sclerosis. Mult Scler 2008;14:663–670.
29. Castillo-Trivino T, Mowry EM, Gajofatto A, et al. Switching multiple sclerosis patients with breakthrough disease to second-line therapy. PLoS One 2011;6:e16664.
30. Bevan CJ, Cree BA. Disease activity free status: a new end point for a new era in multiple sclerosis clinical research? JAMA Neurol 2014;7:269–270.
31. Lublin FD, Cofield SS, Cutter GR, et al. Randomized study combining interferon and glatiramer acetate in multiple sclerosis. Ann Neurol 2013;73:327–340.
32. Gourraud PA, McElroy JP, Caillier SJ, et al. Aggregation of multiple sclerosis genetic risk variants in multiple and single case families. Ann Neurol 2011;69:65–74.
33. Munger KL, Levin LI, Hollis BW, et al. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. JAMA 2006;296:2832–2838.
34. Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. Ann Neurol 1983;13:227–231.
35. Mowry EM, Waubant E, McCulloch CE, et al. Vitamin D status predicts new brain magnetic resonance imaging activity in multiple sclerosis. Ann Neurol 2012;72:234–240.
36. Ascherio A, Munger KL, White R, et al. Vitamin D as an early predictor of multiple sclerosis activity and progression. JAMA Neurol 2014;71:306–314.
37. Gelfand JM, Cree BA, McElroy J, et al. Vitamin D in African Americans with multiple sclerosis. Neurology 2011;76:1824–1830.
38. Tremlett H, Zhao Y, Devonshire V. Natural history of secondary-progressive multiple sclerosis. Mult Scler 2008;14:314–324.
39. Debouverie M, Laforest L, Van Ganse E, et al. Earlier disability of the patients followed in multiple sclerosis centers compared to outpatients. Mult Scler 2009;15:251–257.
40. Tedeholm H, Skoog B, Lisovskaja V, et al. The outcome spectrum of multiple sclerosis: disability, mortality, and a cluster of predictors from onset. J Neurol 2015;262:1148–1163.
41. Trojano M, Pellegrini F, Fuiani A, et al. New natural history of interferon-beta-treated relapsing multiple sclerosis. Ann Neurol 2007;6:300–306.
42. Weinshenker BG, Bass B, Rice GP, et al. The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. Brain 1989;112(pt 1):133–146.
43. Bermel RA, You X, Foulds P, et al. Predictors of long-term outcome in multiple sclerosis patients treated with interferon β. Ann Neurol 2013;73:95–103.
44. Prosperini L, Gallo V, Petsas N, et al. One-year MRI scan predicts clinical response to interferon beta in multiple sclerosis. Eur J Neurol 2009;16:1202–1209.
45. Romeo M, Martinelli-Boneschi F, Rodegher M, et al. Clinical and MRI predictors of response to interferon-beta and glatiramer acetate in relapsing-remitting multiple sclerosis patients. Eur J Neurol 2013;20:1060–1067.
46. Goodin DS, Traboulsee A, Knappertz V, et al. Relationship between early clinical characteristics and long-term disability outcomes: 16 year cohort study (follow-up) of the pivotal interferon β-1b trial in multiple sclerosis. J Neurol Neurosurg Psychiatry 2012;83:282–287.
47. Brown MG, Asbridge M, Hicks V, et al. Estimating typical multiple sclerosis disability progression speed from clinical observations. PLoS One 2014;9:e105123.
48. Ebers GC, Heigenhauser L, Daumer M, et al. Disability as an outcome in MS clinical trials. Neurology 2008;71:624–631.
49. Rotstein DL, Healy BC, Malik MT, et al. Evaluation of no evidence of disease activity in a 7-year longitudinal multiple sclerosis cohort. JAMA Neurol 2015;72:152–158.
50. Kappos L, De Stefano N, Freedman MS, et al. Inclusion of brain volume loss in a revised measure of 'no evidence of disease activity' (NEDA-4) in relapsing-remitting multiple sclerosis. Mult Scler 2015 [Epub ahead of print].
51. Daumer M, Neuhaus A, Morrissey S, et al. MRI as an outcome in multiple sclerosis clinical trials. Neurology 2009;72:705–711.
52. Cook SD, Dhib-Jalbut S, Dowling P, et al. Use of magnetic resonance imaging as well as clinical disease activity in the clinical classification of multiple sclerosis and assessment of its course: a report from an International CMSC Consensus Conference, March 5–7, 2010. Int J MS Care 2012;14:105–114.
53. Schlaeger R, Papinutto N, Panara V, et al. Spinal cord gray matter atrophy correlates with multiple sclerosis disability. Ann Neurol 2014;76:568–580.
54. Schlaeger R, Papinutto N, Zhu AH, et al. Association between thoracic spinal cord gray matter atrophy and disability in multiple sclerosis. JAMA Neurol 2015;72:897–904.