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

University of California, San Francisco MS-EPIC Team, Cree BAC et al. Annals of Neurology 2016

Supplementary Tables

Supplementary Table 1: Baseline clinical and MRI features of the entire EPIC cohort

Characteristic	All (n = 517)	RMS (n = 448)	PMS (n = 69)	*p*-value
Demographic				
Age at exam, mean ± sd	42.5 ± 9.8	41.6 ± 9.6	48.6 ± 8.7	1.57e-08
Sex				
Women, n (%)	355 (68.7%)	314 (70.1%)	41 (59.4%)	0.094
Men, n (%)	162 (31.3%)	134 (29.9%)	28 (40.6%)	0.094
Years of follow-up, MIR	9.4 (8.4, 10.2) (0-11.5)	9.6 (8.4, 10.2) (0-11.5)	8.5 (8.5, 10.1) (1-11.2)	0.14
Clinical				
Age of onset, mean ± sd	33.4 ± 9.3	33.5 ± 9.2	32.8 ± 10.2	0.576
Disease duration, MIR	6 (2, 13) (0-46)	6 (2, 12) (0-46)	15 (7, 21) (1-45)	1.96e-11
Disease course				
CIS, n (%)	82 (15.9%)	82 (18.3%)		
RR, n (%)	366 (70.8%)	366 (81.7%)		
SP, n (%)	48 (9.3%)		49 (69.6%)	
PP, n (%)	21 (4.1%)		21 (30.4%)	
EDSS score, MIR	1.5 (1, 3) (0-7)	1.5 (1, 2) (0-6.5)	4.5 (3.5, 6) (1.5-7)	8.3e-30
MSSS, MIR	2.4 (0.9, 4.3) (0-9.8)	2.2 (0.7, 4) (0-9.5)	5.2 (3.4, 7.2) (0.8-9.8)	7.2e-15
Relapse history				
Annualized relapse rate, MIR	0.5 (0.3, 1) (0-8.3)	0.5 (0.3, 1.1) (0-8.3)	0.2 (0.1, 0.5) (0-1.1)	4.89e-10
Vitamin D level (ng/mL), mean ± sd	24.3 ± 9	24.3 ± 9	24.1 ± 9.4	0.86
Treatment				
Treatment history				
No treatment, n (%)	209 (40.4%)	177 (39.5%)	32 (46.4%)	0.294
Platform therapy, n (%)	301 (58.2%)	266 (59.4%)	35 (50.7%)	0.191
High potency, n (%)	7 (1.4%)	5 (1.1%)	2 (2.9%)	0.237
Years to first treatment from diagnosis, MIR	3.1 (0.8, 8.7) (0-43.9)	2.7 (0.7, 7.3) (0-43.9)	6.5 (3.3, 14.6) (0-36.4)	4.25e-06
Medication possession ratio (prestudy), MIR	0.2 (0, 0.6) (0-1)	0.2 (0, 0.6) (0-1)	0.2 (0, 0.5) (0-0.9)	0.865
MRI				
T2 lesion volume (mL), MIR	2.6 (0.7, 6.6) (0-103.9)	2.3 (0.7, 5.6) (0-103.9)	6.1 (1.6, 11.9) (0-71.7)	8.19e-06
Number of gad enhancing lesions, MIR	0 (0, 0) (0-10)	0 (0, 0) (0-9)	0 (0, 0) (0-10)	0.786
Total brain volume (mL), mean ± sd	1461.5 ± 89.3	1471.3 ± 85.3	1397 ± 88.5	5.34e-09
Grey matter volume (mL), mean ± sd	792.1 ± 60.1	798.7 ± 58.1	748.9 ± 55.6	8.61e-10
White matter volume (mL), mean ± sd	669.4 ± 42.4	672.6 ± 41.1	648.1 ± 44.7	5.31e-05
Ventricular CSF volume (mL), MIR	40 (30, 54.2) (10-172)	38.5 (29, 50.2) (10-172)	55 (39, 69) (15-134)	2.08e-07
Cortical grey matter volume (mL), mean ± sd	627.5 ± 49.4	633 ± 47.7	591.9 ± 45.7	8e-10
Genetic				
HLA-DRB1*1501				
1 allele, n (%)	197 (39%)	172 (39.3%)	25 (37.3%)	0.79
2 alleles, n (%)	34 (6.7%)	30 (6.8%)	4 (6%)	1
MSGB score, mean ± sd	9.9 ± 0.6	9.9 ± 0.6	10 ± 0.7	0.353

All subjects with a baseline visit are included. Of the 82 CIS subjects, 34 experienced a second clinical attack or developed disease progression. An additional 22 subjects developed new MRI lesions and would fulfill International Panel criteria for radiographic dissemination over time. 12 CIS subjects were lost to follow up and did not contribute to the long-term outcome data. 16 CIS subjects did not fulfill International Panel Criteria for dissemination over time. These 16 CIS subjects who remained clinically and radiographically stable over the course of the study continue to be followed in the cohort and could 'convert' to MS. 7 of the 82 CIS subjects had an asymptomatic gadolinium enhanced lesion at baseline brain MRI and at least one other brain lesion. Of these 7 subjects who would be classified as having MS by 2010 International Panel criteria, 2 developed new T2 lesions on brain MRI and 5 did not. All CIS subjects had to have abnormal brain MRI scans to be included in the study. Only a single CIS subject with a spinal cord presentation and a solitary brain lesion consistent with a demyelinating plaque was included in the study and was lost to follow after the first year of the study. 17% of control subjects have 1 HLA-DRB1*1501 allele. 1% of control subjects have 2 HLA-DRB1*1501 alleles. Multiple Sclerosis Genetic Burden[33] scores are based on 88 SNPs. Control subjects have a mean MSGB score of 9.6 and a standard deviation of 0.6. *P*-values are shown to compare RMS and PMS subjects. For normally distributed data, mean and standard deviation are shown and a t-test was used to generate p-values. 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's exact test was used.

MIR = median (IQR) (range); CSF = cerebrospinal fluid.

Supplementary Table 2: Comparison of baseline characteristics of subjects retained in study versus those lost to follow-up

Characteristic	All (n = 517)	In study (n = 471)	Lost to follow-up (n = 46)	*p*-value
Demographic				
Age at exam, mean ± sd	42.5 ± 9.8	42.7 ± 9.9	41.2 ± 8.7	0.286
Sex				
Women, n (%)	355 (68.7%)	317 (67.5%)	37 (80.4%)	0.095
Men, n (%)	162 (31.3%)	153 (32.5%)	9 (19.6%)	0.095
Years of follow-up, MIR	9.4 (8.4, 10.2) (0-11.5)	9.8 (8.6, 10.2) (1-11.5)	1.7 (1, 4.4) (0-6.2)	1.84e-28
Clinical				
Age of onset, mean ± sd	33.4 ± 9.3	33.3 ± 9.3	34.8 ± 9.2	0.302
Disease duration, MIR	6 (2, 13) (0-46)	7 (2, 13.5) (0-46)	4.5 (2, 10) (0-23)	0.031
Disease course				
CIS, n (%)	82 (15.9%)	70 (14.9%)	12 (26.1%)	0.057
RR, n (%)	366 (70.8%)	337 (71.5%)	29 (63%)	0.237
SP, n (%)	48 (9.3%)	44 (9.6%)	3 (6.5%)	0.789
PP, n (%)	21 (4.1%)	19 (4%)	2 (4.3%)	0.709
EDSS score, MIR	1.5 (1, 3) (0-7)	1.5 (1, 3) (0-7)	2 (1.5, 3) (0-6.5)	0.269
MSSS, MIR	2.4 (0.9, 4.3) (0-9.8)	2.4 (0.9, 4.3) (0-9.8)	3.6 (2, 5.9) (0-9.1)	0.019
Relapse history				
Annualized relapse rate, MIR	0.5 (0.3, 1) (0-8.3)	0.5 (0.2, 1) (0-7.3)	0.6 (0.4, 1.2) (0.1-8.3)	0.03
Vitamin D level (ng/mL), mean ± sd	24.3 ± 9	24.4 ± 8.8	23.2 ± 11.1	0.492
Treatment				
Treatment history				
No treatment, n (%)	209 (40.4%)	183 (38.9%)	26 (56.5%)	0.027
Platform therapy, n (%)	301 (58.2%)	281 (59.7%)	20 (43.5%)	0.041
High potency, n (%)	7 (1.4%)	7 (1.5%)		
Years to first treatment from diagnosis, MIR	3.1 (0.8, 8.5) (0-43.9)	3.1 (0.8, 8.7) (0-43.9)	3.4 (0.8, 9) (0-19.5)	0.832
Medication possession ratio (prestudy), MIR	0.2 (0, 0.6) (0-1)	0.2 (0, 0.6) (0-1)	0 (0, 0.5) (0-1)	0.043
MRI				
T2 lesion volume (mL), MIR	2.6 (0.7, 6.6) (0-103.9)	2.7 (0.8, 6.8) (0-103.9)	1 (0.3, 4) (0-23.3)	0.01
Number of gad enhancing lesions, MIR	0 (0, 0) (0-10)	0 (0, 0) (0-10)	0 (0, 0) (0-6)	0.077
Total brain volume (mL), mean ± sd	1461.5 ± 89.3	1460.1 ± 87.7	1475.5 ± 103.7	0.334
Grey matter volume (mL), mean ± sd	792.1 ± 60.1	790.8 ± 58.9	805.5 ± 71.3	0.18
White matter volume (mL), mean ± sd	669.4 ± 42.4	669.3 ± 42.5	669.9 ± 41.7	0.932
Ventricular CSF volume (mL), MIR	40 (30, 54.2) (10-172)	41 (30, 55) (10-172)	33.5 (28, 45) (17-105)	0.027
Cortical grey matter volume (mL), mean ± sd	627.5 ± 49.4	626.3 ± 48.2	636.6± 58.8	0.272
Genetic				
HLA-DRB1*1501				
1 allele, n (%)	197 (39%)	182 (39.2%)	15 (36.6%)	0.868
2 alleles, n (%)	34 (6.7%)	34 (7.3%)		
MSGB score, mean ± sd	9.9 ± 0.6	9.9 ± 0.6	9.9 ± 0.6	0.572

In-study subjects include subjects with 10-year follow up and deceased subjects. Lost to follow-up subjects include all other subjects with a baseline visit. *P*-values are shown to compare in study and lost to follow-up subjects. For normally distributed data, mean and standard deviation are shown and a 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's exact test was used.
MIR = median (IQR) (range); CSF = cerebrospinal fluid.

Supplementary Table 3: Comparison of baseline characteristics of RMS subjects retained in study versus those lost to follow-up

Characteristic	All (n=448)	In study (n=407)	Lost to follow-up (n=41)	*p*-value
Demographic				
Age at exam, mean ± sd	41.6 ± 9.6	41.7 ± 9.7	40.3 ± 8.5	0.321
Sex				
Women, n (%)	314 (70.1%)	280 (68.8%)	34 (82.9%)	0.073
Men, n (%)	134 (29.9%)	127 (31.2%)	7 (17.1%)	0.073
Years of follow-up, MIR	9.6 (8.4, 10.2) (0-11.5)	9.9 (8.6, 10.2) (1-11.5)	1.7 (1, 4) (0-6.2)	7.83e-26
Clinical				
Age of onset, mean ± sd	35.5 ± 9.2	33.4 ± 9.2	34.7 ± 9.1	0.388
Disease duration, MIR	6 (2, 12) (0-46)	6 (2, 12) (0-46)	3 (1, 8) (0-23)	0.03
Disease course				
CIS, n (%)	82 (18.3%)	70 (17.2%)	12 (29.3%)	0.087
RR, n (%)	366 (81.7%)	337 (82.2%)	29 (70.7%)	0.087
EDSS score, MIR	1.5 (1, 2) (0-6.5)	1.5 (1, 2) (0-6.5)	1.5 (1, 2.5) (0-4.5)	0.112
MSSS, MIR	2.1 (0.3, 1.1) (0-8.3)	2.1 (0.7, 3.7) (0-9.5)	3.3 (1.3, 5.2) (0-9.1)	0.009
Relapse history				
Annualized relapse rate, MIR	0.5 (0.3, 1.1) (0-8.3)	0.5 (0.3, 1.1) (0-7.3)	0.7 (0.4, 1.3) (0.2-8.3)	0.047
Vitamin D level (ng/mL), mean ± sd	24.3 ± 9	24.4 ± 8.7	23.2 ± 11.4	0.488
Treatment				
Treatment history				
No treatment, n (%)	177 (39.5%)	155 (38.1%)	22 (53.7%)	0.065
Platform therapy, n (%)	266 (59.4%)	247 (60.7%)	19 (46.3%)	0.095
High potency, n (%)	5 (1.1%)	5 (1.2%)		
Years to first treatment from diagnosis, MIR	2.8 (0.7, 7.4) (0-43.9)	2.8 (0.7, 7.6) (0-43.9)	3 (0.6, 6.3) (0-19.5)	0.525
Medication possession ratio (prestudy), MIR	0.2 (0, 0.6) (0-1)	0.2 (0, 0.6) (0-1)	0 (0, 0.5) (0-1)	0.123
MRI				
T2 lesion volume (mL), MIR	2.3 (0.7, 5.6) (0-103.9)	2.4 (0.7, 5.7) (0-103.9)	1 (0.2, 4) (0-23.3)	0.014
Number of gad enhancing lesions, MIR	0 (0, 0) (0-9)	0 (0, 0) (0-9)	0 (0, 0) (0-6)	0.042
Total brain volume (mL), mean ± sd	1471.5 ± 85.3	1470.1 ± 83	1483.8 ± 105.4	0.423
Grey matter volume (mL), mean ± sd	798.7 ± 58.1	797.5 ± 56.5	811 ± 71.6	0.245
White matter volume (mL), mean ± sd	672.6 ± 41.1	672.6 ± 41	672.6 ± 42.4	0.998
Ventricular CSF volume (mL), MIR	38.5 (29, 50.2) (10-172)	39 (29.5, 51) (10-172)	33 (26, 45) (17-105)	0.088
Cortical grey matter volume (mL), mean ± sd	633 ± 47.7	632.2 ± 46.3	640.6± 59.8	0.385
Genetic				
HLA-DRB1*1501				
1 allele, n (%)	172 (39.3%)	159 (39.7%)	15 (36.1%)	0.725
2 alleles, n (%)	30 (6.8%)	30 (7.5%)		
MSGB score, mean ± sd	9.9 ± 0.6	9.9 ± 0.6	9.9 ± 0.6	0.748

In-study subjects include subjects with RMS at baseline with 10-year follow up and deceased subjects. Lost to follow-up subjects include all other RMS subjects with a baseline visit. *P*-values are shown to compare in study and lost to follow-up subjects. For normally distributed data, mean and standard deviation are shown and a 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's exact test was used.
MIR = median (IQR) (range); CSF = cerebrospinal fluid.

Supplementary Table 4: Comparison of baseline characteristics of PMS subjects retained in study versus those lost to follow-up

Characteristic	All (n=69)	In study (n=64)	Lost to follow-up (n=5)	*p*-value
Demographic				
Age at exam, mean ± sd	48.6 ± 8.6	48.6 ± 8.7	48.4 ± 7.4	0.965
Sex				
Women, n (%)	41 (59.4%)	38 (59.4%)	3 (60%)	1
Men, n (%)	28 (40.6%)	26 (40.6%)	2 (40%)	1
Years of follow-up, MIR	9.2 (8.5, 10.1) (0-11.2)	9.3 (8.6, 10.2) (1-11.2)	4.4 (4.1, 5) (1-5.2)	6.4e-04
Clinical				
Age of onset, mean ± sd	32.8 ± 10.2	32.6 ± 10.2	35.4 ± 11.2	0.613
Disease duration, MIR	15 (7, 21) (1-46)	15 (7, 22.2) (1-45)	15 (12, 15) (4-19)	0.627
Disease course				
SP, n (%)	48 (69.6%)	45 (70.3%)	3 (60%)	0.636
PP, n (%)	21 (34.4%)	19 (29.7%)	2 (40%)	0.636
EDSS score, MIR	4.5 (3.5, 6) (1.5-7)	4.5 (3.5, 6) (1.5-7)	3.5 (2.5, 6) (2.5-6.5)	1
MSSS, MIR	5.2 (3.4, 7.2) (0.8-9.8)	5.2 (3.4, 7.2) (0.8-9.8)	6.6 (3.4, 7.2) (2.3-7.7)	0.862
Relapse history				
Annualized relapse rate, MIR	0.2 (0.1, 0.5) (0-1.1)	0.2 (0.1, 0.4) (0-1.1)	0.5 (0.2, 0.5) (0.1-0.7)	0.495
Vitamin D level (ng/mL), mean ± sd	24.1 ± 9.4	24.1 ± 9.5	23.9 ± 8.5	0.952
Treatment				
Treatment history				
No treatment, n (%)	32 (46.4%)	28 (38.9%)	4 (80%)	0.175
Platform therapy, n (%)	35 (50.7%)	34 (59.7%)	1 (20%)	0.198
High potency, n (%)	2 (2.9%)	2 (1.5%)		
Years to first treatment from diagnosis, MIR	6.5 (3.1, 14.6) (0-36.4)	6.4 (3, 13.6) (0-36.4)	12.4 (11.8, 15.3) (2.5-17)	0.242
Medication possession ratio (prestudy), MIR	0.2 (0, 0.5) (0-0.9)	0.3 (0, 0.5) (0-0.9)	0 (0, 0.2) (0-3)	0.066
MRI				
T2 lesion volume (mL), MIR	6.1 (1.6, 11.9) (0-71.7)	7 (2.1, 12) (0-71.7)	4 (1.6, 4.8) (0.4-11.1)	0.312
Number of gad enhancing lesions, MIR	0 (0, 0) (0-10)	0 (0, 0) (0-10)	0 (0, 0) (0-3)	0.664
Total brain volume (mL), mean ± sd	1397 ± 88.5	1396.2 ± 90.7	1407.6 ± 59.6	0.708
Grey matter volume (mL), mean ± sd	748.9 ± 55.6	748 ± 56	760.2 ± 54.8	0.653
White matter volume (mL), mean ± sd	648.1 ± 44.7	648.2 ± 45.9	647.2 ± 29	0.948
Ventricular CSF volume (mL), MIR	55 (39, 69) (15-134)	55 (39, 71) (15-134)	34 (33, 43) (33-57)	0.043
Cortical grey matter volume (mL), mean ± sd	591.9 ± 45.7	591 ± 46.3	603.2 ± 40.2	0.547
Genetic				
HLA-DRB1*1501				
1 allele, n (%)	35 (37.3%)	23 (36.5%)	2 (50%)	0.626
2 alleles, n (%)	4 (6%)	4 (6.3%)		
MSGB score, mean ± sd	10 ± 0.7	10 ± 0.7	9.8 ± 0.6	0.534

In-study subjects include subjects with PMS at baseline with 10-year follow up and deceased subjects. Lost to follow-up PMS subjects include all other subjects with a baseline visit. *P*-values are shown to compare in study and lost to follow-up subjects acknowledging that these values should be interpreted with caution because only 5 PMS patients who were lost to follow up. For normally distributed data, mean and standard deviation are shown and a 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's exact test was used.

MIR = median (IQR) (range); CSF = cerebrospinal fluid.

Supplementary Table 5: Comparison of baseline characteristics of EPIC versus Non-EPIC RR subjects

Characteristic	EPIC RR (n=366)	non-EPIC RR (n=234)	*p*-value
Women, n (%)	261 (71.3%)	177 (75.6%)	0.259
Men, n (%)	105 (28.7%)	57 (24.4%)	0.259
Clinical Age of onset, mean ± sd	32.2 ± 8.5	32 ± 10.1	0.882
Disease duration, MIR	7 (3, 13) (0-46)	7.7 (3.8, 14.2) (0-45.5)	0.074
EDSS score, MIR	1.5 (1, 2.5) (0-6.5)	2 (1, 2.5) (0-6.5)	0.304
MSSS, MIR	2.1 (0.7, 3.7) (0-9.5)	2.1 (1, 3.7) (0-9.1)	0.411

P-values compare EPIC and non-EPIC RR subjects. For normally distributed data, mean and standard deviation are shown and Student's t-test was used. For data that are not normally distributed, median, interquartile, and range are shown and a Kolmogorov-Smirnov test was used. For qualitative data, counts and percentages are shown and Fisher's exact test was used.
MIR = median (IQR) (range)

Supplementary Table 6: Summary of clinical outcomes for subjects completing 10-year follow-up

	Baseline to last visit RMS (n = 407)	PMS (n = 64)	*p*-value	Year 2 to last visit RMS (n = 407)	PMS (n = 64)	*p*-value
EDSS worse, n (%)	225 (55.3%)	48 (75%)	7.33e-04	144 (35.4%)	38 (59.4%)	1.25e-05
T25W worse, n (%)	80 (19.7%)	24 (37.5%)	5.05e-06	82 (20.1%)	23 (35.9%)	7.57e-06
9HPT worse, n (%)	43 (10.6%)	16 (25%)	8.19e-05	44 (10.8%)	14 (21.9%)	0.002
PASAT worse, n (%)	33 (8.1%)	9 (14.1%)	0.029	27 (6.6%)	11 (17.2%)	8e-04
Composite worse, n (%)	112 (27.5%)	30 (46.9%)	7.37e-07	110 (27%)	29 (45.3%)	9.8e-07

43 subjects missed the year 2 visit and EDSS scores were not obtained on 2 other subjects. For the 4 subjects who died of non-MS causes by year 2, EDSS scores were not available. The 4 subjects that died from MS are scored with an EDSS of 10. For 8 subjects who died of non-MS causes, year 10 EDSS scores were not available. For subjects who became disabled due to non-MS causes, the last EDSS recorded was carried forward. An EDSS score was not determined for 1 subject with a visit at year 10. Subjects with less than 10 years of follow-up are not included since their 10-year outcome is unknown. Composite worsening is defined as a worsening on either PASAT or 9HPT or T25W. *P*-values compare RMS to PMS subjects with Fisher's exact test.

Supplementary Table 7: Summary of clinical outcomes for subjects completing 10-year follow-up including those subjects who worsened after baseline but did not complete the year 10 visit

	Baseline to last visit RMS (n = 416)	PMS (n = 67)	*p*-value	Year 2 to last visit RMS (n = 416)	PMS (n = 67)	*p*-value
EDSS worse, n (%)	232 (55.8%)	50 (74.6%)	9.62e-04	145 (34.9%)	38 (58.2%)	1.63e-05
T25W worse, n (%)	84 (20.2%)	27 (20.3%)	3.84e-07	83 (20%)	25 (37.3%)	1.94e-06
9HPT worse, n (%)	46 (11.1%)	18 (26.9%)	8.19e-05	45 (10.8%)	15 (22.4%)	0.001
PASAT worse, n (%)	35 (8.4%)	10 (14.9%)	0.02	27 (6.5%)	12 (17.9%)	3.27e-04
Composite worse, n (%)	112 (26.9%)	30 (44.8%)	7.37e-07	110 (26.4%)	29 (43.3%)	9.8e-07

10 subjects experienced clinical worsening after their baseline but did not complete the year 10 visit. These subjects were included in a sensitivity analysis to prevent possible bias introduced by their exclusion (see also Supplementary Tables 5b, 6b and 8b). 43 subjects missed the year 2 visit and EDSS scores were not obtained on 2 other subjects. For the 4 subjects who died of non-MS causes by year 2, EDSS scores were not available. The 4 subjects that died from MS are scored with an EDSS of 10. For 8 subjects who died of non-MS causes, year 10 EDSS scores were not available. For subjects who became disabled due to non-MS causes, the last EDSS recorded was carried forward. An EDSS score was not determined for 1 subject with a visit at year 10. Subjects with less than 10 years of follow-up are not included since their 10-year outcome is unknown. Composite worsening is defined as a worsening on either PASAT or 9HPT or T25W. *P*-values compare RMS to PMS subjects with Fisher's exact test.

Supplementary Table 8: Analysis on clinical outcomes from year 2 to 10 – all RMS subjects with long-term follow up (n = 407)

Response and Predictors	wo/PS adjustment			w/PS adjustment					
	OR	95% CI	*p*-value	OR	95% CI	*p*-value	$\beta 0$	$\beta 1$	$\beta 2$
EDSS Worsening									
Baseline to year 2 increase in EDSS	0.35	[0.22, 0.56]	1.12e-05	0.34	[0.21, 0.55]	1.01e-05	-0.3	-1.06	0.39
Baseline to year 2 new T2 lesions	0.98	[0.92, 1.03]	0.56	0.98	[0.91, 1.03]	0.475	-0.66	-0.02	0.41
Baseline to year 2 NEDA	1.42	[0.84, 2.39]	0.189	1.48	[0.87, 2.5]	0.148	-0.8	0.39	0.44
Baseline to year 3 therapeutic escalation	1.04	[0.58, 1.84]	0.894	1.11	[0.61, 1.98]	0.729	-0.7	0.1	0.41
Baseline to year 2 mean vitamin D level	0.99	[0.96, 1.02]	0.493	0.99	[0.96, 1.02]	0.474	-0.44	-0.01	0.42
PASAT Worsening									
Baseline to year 2 % BVL	1.4	[1.03, 1.9]	0.025	1.38	[1, 1.86]	0.037	-3.18	0.32	0.65
Baseline to year 2 new T2 lesions	0.85	[0.62, 1.02]	0.242	0.83	[0.59, 1.02]	0.194	-3.02	-0.18	1.13
Baseline to year 2 NEDA	0.66	[0.19, 1.79]	0.452	0.7	[0.2, 1.92]	0.525	-2.94	-0.36	0.85
Baseline to year 3 therapeutic escalation	1.98	[0.74, 4.75]	0.144	2.2	[0.81, 5.42]	0.098	-3.14	0.79	0.87
Baseline to year 2 mean vitamin D level	1.04	[0.99, 1.09]	0.103	1.04	[0.99, 1.09]	0.121	-3.8	0.04	0.68
T25W Worsening									
Baseline to year 2 new T2 lesions	0.99	[0.9, 1.04]	0.7	0.98	[0.9, 1.04]	0.618	-1.38	-0.02	0.43
Baseline to year 2 NEDA	0.65	[0.33, 1.24]	0.209	0.67	[0.33, 1.27]	0.237	-1.29	-0.4	0.38
Baseline to year 3 therapeutic escalation	1.95	[1.03, 3.63]	0.037	2.04	[1.07, 3.84]	0.028	-1.55	0.71	0.49
Baseline to year 2 mean vitamin D level	1.01	[0.97, 1.04]	0.725	1	[0.97, 1.04]	0.781	-1.46	0	0.37
9HPT Worsening									
Baseline to year 2 delta grey matter volume (dL)	0.24	[0.06, 0.98]	0.048	0.24	[0.06, 1]	0.049	-2.39	-1.42	0.41
Baseline to year 2 delta white matter volume (dL)	0.15	[0.03, 0.82]	0.027	0.14	[0.03, 0.81]	0.026	-2.33	-1.93	0.43
Baseline to year 2 new T2 lesions	1.01	[0.92, 1.06]	0.799	1	[0.92, 1.06]	0.878	-2.18	0	0.4
Baseline to year 2 NEDA	0.35	[0.1, 0.92]	0.055	0.37	[0.11, 0.97]	0.069	-1.99	-0.99	0.33
Baseline to year 3 therapeutic escalation	2.26	[1.05, 4.64]	0.031	2.46	[1.12, 5.15]	0.019	-2.41	0.9	0.48
Baseline to year 2 mean vitamin D level	1	[0.96, 1.04]	0.949	1	[0.96, 1.04]	0.955	-2.15	0	0.33

Regression analysis with and without propensity score adjustment on 4 clinical outcomes for RMS subjects: EDSS worsening from year 2 to year 10, PASAT worsening from year 2 to year 10, T25W worsening from year 2 to year 10, and 9HPT worsening from year 2 to year 10 (responses are underlined in the first column). A propensity score for baseline treatment was included as a covariate in the adjusted model: $ln(p_{worse}/1\text{-}p_{worse}) = \beta_0 + \beta_{1(predictor)}X_1 + \beta_{2(PS)}X_2$. The baseline to year 2 predictors that were tested for each outcome include ARR, 1 or more new T2 lesions, % BVL, grey matter volume change, white matter volume change, cortical grey matter volume change, cerebrospinal fluid volume change, change in T2 volume, EDSS worsening, EDSS worsening confirmed at year 3, worsening PASAT, worsening T25W, worsening 9HPT, 1 or more new gad+ lesions, NEDA, average vitamin D level, and baseline to year 3 increase in tier of therapy. % BVL is the percent brain volume loss between baseline and year 2 multiplied by -1. Associations with p-values < 0.05 and clinically relevant variables are shown. Mean 25-OH vitamin D levels are deseasonalized.

Supplementary Table 9: Analysis on clinical outcomes from year 2 to 10 – all RMS subjects with long-term follow up including those who worsened after baseline but did complete the year 10 visit (n = 416)

Response and Predictors	wo/PS adjustment			w/PS adjustment					
	OR	95% CI	*p*-value	OR	95% CI	*p*-value	$\beta 0$	$\beta 1$	$\beta 2$
EDSS Worsening									
Baseline to year 2 increase in EDSS	0.35	[0.22, 0.56]	1.12e-05	0.35	[0.22, 0.55]	1.03e-05	-0.25	-1.05	0.3
Baseline to year 2 new T2 lesions	0.99	[0.92, 1.03]	0.599	0.98	[0.91, 1.03]	0.526	-0.61	-0.02	0.32
Baseline to year 2 NEDA	1.42	[0.84, 2.39]	0.186	1.47	[0.87, 2.49]	0.151	-0.75	0.38	0.37
Baseline to year 3 therapeutic escalation	1.04	[0.58, 1.85]	0.888	1.1	[0.61, 1.96]	0.751	-0.65	0.09	0.33
Baseline to year 2 mean vitamin D level	0.99	[0.96, 1.02]	0.449	0.99	[0.96, 1.02]	0.437	-0.38	-0.01	0.34
PASAT Worsening									
Baseline to year 2 % BVL	1.39	[1.02, 1.87]	0.029	1.36	[0.99, 1.84]	0.042	-3.21	0.31	0.7
Baseline to year 2 new T2 lesions	0.85	[0.61, 1.02]	0.234	0.83	[0.59, 1.02]	0.19	-3.02	-0.19	1.11
Baseline to year 2 NEDA	0.66	[0.19, 1.81]	0.465	0.71	[0.2, 1.96]	0.546	-2.98	-0.34	0.88
Baseline to year 3 therapeutic escalation	2	[0.75, 4.81]	0.138	2.24	[0.83, 5.53]	0.091	-3.17	0.81	0.9
Baseline to year 2 mean vitamin D level	1.04	[0.99, 1.09]	0.095	1.04	[0.99, 1.09]	0.115	-3.83	0.04	0.69
T25W Worsening									
Baseline to year 2 new T2 lesions	0.99	[0.91, 1.04]	0.758	0.99	[0.9, 1.04]	0.685	-1.32	-0.01	0.34
Baseline to year 2 NEDA	0.65	[0.32, 1.23]	0.204	0.66	[0.33, 1.26]	0.225	-1.23	-0.41	0.31
Baseline to year 3 therapeutic escalation	1.94	[1.02, 3.6]	0.038	2.01	[1.05, 3.77]	0.031	-1.49	0.7	0.41
Baseline to year 2 mean vitamin D level	1	[0.97, 1.04]	0.842	1	[0.97, 1.03]	0.898	-1.35	0	0.31
9HPT Worsening									
Baseline to year 2 new T2 lesions	1.01	[0.92, 1.06]	0.825	1	[0.92, 1.06]	0.896	-2.13	0	0.35
Baseline to year 2 NEDA	0.35	[0.1, 0.9]	0.051	0.36	[0.11, 0.95]	0.063	-1.93	-1.02	0.27
Baseline to year 3 therapeutic escalation	2.21	[1.02, 4.52]	0.035	2.38	[1.09, 4.97]	0.024	-2.34	0.87	0.42
Baseline to year 2 mean vitamin D level	1	[0.96, 1.04]	0.881	1	[0.95, 1.04]	0.873	-2	0	0.28

Regression analysis with and without propensity score adjustment on 4 clinical outcomes for RMS subjects: EDSS worsening from year 2 to year 10, PASAT worsening from year 2 to year 10, T25W worsening from year 2 to year 10, and 9HPT worsening from year 2 to year 10 (responses are underlined in the first column). A propensity score for baseline treatment was included as a covariate in the adjusted model: $ln(p_{worse}/1\text{-}p_{worse}) = \beta_0 + \beta_{1(\text{predictor})}X_1 + \beta_{2(PS)}X_2$. The baseline to year 2 predictors that were tested for each outcome include ARR, 1 or more new T2 lesions, % BVL, grey matter volume change, white matter volume change, cortical grey matter volume change, cerebrospinal fluid volume change, change in T2 volume, EDSS worsening, EDSS worsening confirmed at year 3, worsening PASAT, worsening T25W, worsening 9HPT, 1 or more new gad+ lesions, NEDA, average vitamin D level, and baseline to year 3 increase in tier of therapy. % BVL is the percent brain volume loss between baseline and year 2 multiplied by -1. Associations with p-values < 0.05 and clinically relevant variables are shown. Mean 25-OH vitamin D levels are deseasonalized.

Supplementary Table 10: Summary of clinical and MRI changes from baseline to year 2

Change	All (n = 471)	RMS (n = 407)	PMS (n = 64)
Clinical metrics			
EDSS worse, n (%)	166 (35.2%)	139 (34.2%)	27 (42.2%)
T25W worse, n (%)	33 (7%)	21 (5.2%)	12 (18.8%)
9HPT worse, n (%)	8 (1.7%)	2 (0.5%)	6 (9.4%)
PASAT worse, n (%)	36 (7.6%)	29 (7.1%)	7 (10.9%)
Composite worse, n (%)	68 (14.4%)	47 (11.5%)	20 (32.8%)
ARR, MIR	0 (0, 0.5) (0-2.9)	0 (0, 0.5) (0-2.9)	0 (0, 0) (0-2)
MRI metrics			
T2LV (mL), MIR	-0.3 (-1.2, 0) (-23.9-13.1)	-0.2 (-1.2, 0) (-23.9-9.7)	-0.4 (-1.7, 0) (-4.6-13.1)
Subjects with new T2 lesions, n (%)	201 (42.7%)	172 (42.3%)	29 (45.3%)
Subjects with new gad lesions, n (%)	59 (12.5%)	52 (12.8%)	7 (10.9%)
% BVL, MIR	-0.7 (-1.3, -0.4) (-8.6-2)	-0.7 (-1.3, -0.4) (-8.6-2)	-0.6 (-1.4, -0.4) (-6.2-0.8)
Combined MRI and Clinical metric			
NEDA, n (%)	82 (17.6%)	73 (17.9%)	10 (15.6%)
Laboratory values			
Average 25-OH vitamin D level (ng/mL), mean ± sd	24.3 ± 8.1	24.3 ± 8	24.7 ± 8.8

Subjects completing long-term follow-up include subjects with a year 10 visit and deceased subjects. 43 subjects missed the year 2 visit and EDSS scores were not obtained on 2 other subjects. For the 4 subjects who died of non-MS causes by year 2, EDSS scores were not available. Percentages are of the total number of subjects completing long-term follow-up. %BVL is the percent brain volume loss between baseline and year 2.
MIR = median (IQR) (range).

Supplementary Table 11: Summary of clinical and MRI changes from baseline to year 2 including those who worsened after baseline but did complete the year 10 visit

Change	All (n = 483)	RMS (n = 416)	PMS (n = 67)
Clinical metrics			
EDSS worse, n (%)	169 (35%)	142 (34.1%)	27 (40.3%)
T25W worse, n (%)	35 (.27%)	23 (5.5%)	12 (17.9%)
9HPT worse, n (%)	9 (1.9%)	2 (0.5%)	7 (10.4%)
PASAT worse, n (%)	38 (7.9%)	31 (7.5%)	7 (10.4%)
Composite worse, n (%)	72 (14.9%)	50 (12%)	22 (32.8%)
ARR, MIR	0 (0, 0.5) (0-2.9)	0 (0, 0.5) (0-2.9)	0 (0, 0) (0-2)
MRI metrics			
T2LV (mL), MIR	-0.2 (-1.2, 0) (-23.9-13.1)	-0.2 (-1.2, 0) (-23.9-9.7)	-0.3 (-1.5, 0.2) (-4.6-13.1)
Subjects with new T2 lesions, n (%)	205 (42.4%)	175 (42.1%)	30 (44.8%)
Subjects with new gad lesions, n (%)	59 (12.2%)	52 (12.5%)	7 (10.4%)
% BVL, MIR	-0.7 (-1.3, -0.4) (-8.6-2)	-0.7 (-1.3, -0.4) (-8.6-2)	-0.6 (-1.5, -0.4) (-6.2-0.8)
Combined MRI and Clinical metric			
NEDA, n (%)	84 (17.4%)	73 (17.5%)	11 (16.4%)
Laboratory values			
Average 25-OH vitamin D level (ng/mL), mean ± sd	24.2 ± 8.2	24.1 ± 8.1	24.6 ± 8.8

Subjects include those 10-year follow up, deceased subjects and subjects who worsened after baseline but did not complete the year 10 visit. 43 subjects missed the year 2 visit and EDSS scores were not obtained on 2 other subjects. For the 4 subjects who died of non-MS causes by year 2, EDSS scores were not available. Percentages are of the total number of subjects completing long-term follow-up. %BVL is the percent brain volume loss between baseline and year 2.
MIR = median (IQR) (range).

Supplementary Table 12: 25-OH vitamin D Level (ng/mL) from baseline to year 2 and new gad lesions

Predictor	All subjects			RMS subjects			PMS subjects		
	OR	95% CI	*p*-value	OR	95% CI	*p*-value	OR	95% CI	*p*-value
Baseline to year 2 mean vitamin D level (deseasonalized)	0.95	[0.91, 0.98]	0.008	0.94	[0.9, 0.98]	0.006	0.99	[0.9, 1.08]	0.803
Baseline to year 2 mean vitamin D level lowest quartile	2.74	[1.55, 4.81]	4.58e-04	2.77	[1.51, 5.05]	9e-04	2.54	[0.45, 13.01]	0.26

Supplementary Table 13: Analysis on clinical outcomes from year 2 to last visit – RMS subjects with long-term follow up excluding clinically and radiographically stable CIS (n = 391)

Response and Predictors	wo/PS adjustment			w/PS adjustment					
	OR	95% CI	*p* -value	OR	95% CI	*p*-value	$\beta 0$	$\beta 1$	$\beta 2$
EDSS Worsening									
Baseline to year 2 increase in EDSS	0.34	[0.21, 0.55]	9.6e-06	0.34	[0.21, 0.54]	9.04e-06	-0.25	-1.09	0.33
Baseline to year 2 new T2 lesions	0.98	[0.91, 1.03]	0.533	0.98	[0.91, 1.03]	0.468	-0.62	-0.02	0.37
Baseline to year 2 NEDA	1.51	[0.87, 2.62]	0.14	1.54	[0.89, 2.68]	0.124	-0.76	0.43	0.38
Baseline to year 3 therapeutic escalation	1.06	[0.59, 1.89]	0.843	0.13	[0.62, 2.03]	0.687	-0.67	0.12	0.37
Baseline to year 2 mean vitamin D level	0.99	[0.96, 1.02]	0.471	0.99	[0.96, 1.02]	0.456	-0.4	-0.01	0.38
PASAT Worsening									
Baseline to year 2 % BVL	1.38	[1.01, 1.87]	0.034	1.36	[0.99, 1.84]	0.047	-3.08	0.3	0.57
Baseline to year 2 new T2 lesions	0.84	[0.6, 1.02]	0.203	0.82	[0.58, 1.01]	0.174	-2.88	-0.19	1
Baseline to year 2 NEDA	0.75	[0.21, 2.04]	0.604	0.77	[0.22, 2.12]	0.641	-2.86	-0.26	0.76
Baseline to year 3 therapeutic escalation	1.93	[0.72, 4.64]	0.16	2.13	[0.78, 5.26]	0.114	-3.04	0.76	0.78
Baseline to year 2 mean vitamin D level	1.04	[0.99, 1.09]	0.085	1.04	[0.99, 1.09]	0.1	-3.76	0.04	0.57
T25W Worsening									
Baseline to year 2 new T2 lesions	0.99	[0.9, 1.04]	0.663	0.98	[0.89, 1.04]	0.613	-1.28	-0.02	0.3
Baseline to year 2 NEDA	0.79	[0.39, 1.52]	0.497	0.79	[0.39, 1.52]	0.504	-1.24	-0.23	0.29
Baseline to year 3 therapeutic escalation	1.98	[1.04, 3.7]	0.034	2.05	[1.06, 3.87]	0.029	-1.47	0.72	0.37
Baseline to year 2 mean vitamin D level	1.01	[0.98, 1.05]	0.417	1.01	[0.98, 1.05]	0.447	-1.56	0.01	0.24
9HPT Worsening									
Baseline to year 2 delta grey matter volume (dL)	0.12	[0.02, 0.69]	0.018	0.12	[0.02, 0.7]	0.019	-2.18	-2.12	0.27
Baseline to year 2 new T2 lesions	1	[0.92, 1.06]	0.878	1	[0.91, 1.06]	0.927	-2.05	0	0.27
Baseline to year 2 NEDA	0.4	[0.12, 1.05]	0.094	0.41	[0.12, 1.09]	0.107	-1.91	-0.88	0.24
Baseline to year 3 therapeutic escalation	2.2	[1.02, 4.54]	0.037	2.37	[1.08, 4.97]	0.026	-2.29	0.86	0.37
Baseline to year 2 mean vitamin D level	1	[0.96, 1.05]	0.872	1	[0.96, 1.05]	0.865	-2.09	0	0.21

Regression analysis with and without propensity score adjustment on 4 clinical outcomes for RMS subjects: EDSS worsening from year 2 to year 10, PASAT worsening from year 2 to year 10, T25W worsening from year 2 to year 10, and 9HPT worsening from year 2 to year 10 (responses are underlined in the first column). A propensity score for treatment tier at baseline was included as a covariate in the adjusted model: $ln(p_{worse}/1\text{-}p_{worse}) = \beta_0 + \beta_{1(\text{predictor})}X_1 + \beta_{2(\text{PS})}X_2$. The baseline to year 2 predictors that were tested for each outcome include ARR, 1 or more new T2 lesions, % BVL, grey matter volume change, white matter volume change, cortical grey matter volume change, cerebrospinal fluid volume change, change in T2 volume, EDSS worsening, EDSS worsening confirmed at year 3, worsening PASAT, worsening T25W, worsening 9HPT, 1 or more new gad+ lesions, NEDA, average vitamin D level, and baseline to year 3 increase in tier of therapy. % BVL is the percent brain volume loss between baseline and year 2 multiplied by -1. Associations with *p*-values < 0.05 and clinically relevant variables are shown. Mean 25-OH vitamin D levels are deseasonalized.

Supplementary Table 14: Analysis on clinical outcomes from year 2 to last visit – RMS subjects with long-term follow up including those who worsened after baseline but did not complete a year 10 visit and excluding clinically and radiographically stable CIS (n = 399)

Response and Predictors	wo/PS adjustment			w/PS adjustment					
	OR	95% CI	*p*-value	OR	95% CI	*p*-value	$\beta 0$	$\beta 1$	$\beta 2$
EDSS Worsening									
Baseline to year 2 increase in EDSS	0.35	[0.21, 0.55]	9.6e-06	0.34	[0.21, 0.54]	9.06e-06	-0.19	-1.08	0.24
Baseline to year 2 new T2 lesions	0.98	[0.92, 1.03]	0.571	0.98	[0.91, 1.03]	0.516	-0.57	-0.02	0.27
Baseline to year 2 NEDA	1.52	[0.87, 2.63]	0.138	1.54	[0.88, 2.68]	0.125	-0.71	0.43	0.31
Baseline to year 3 therapeutic escalation	1.06	[0.59, 1.89]	0.836	0.12	[0.61, 2.01]	0.71	-0.62	0.11	0.29
Baseline to year 2 mean vitamin D level	0.99	[0.96, 1.02]	0.429	0.99	[0.96, 1.02]	0.421	-0.33	-0.01	0.31
PASAT Worsening									
Baseline to year 2 % BVL	1.36	[1, 1.84]	0.04	1.34	[0.98, 1.81]	0.053	-3.11	0.3	0.62
Baseline to year 2 new T2 lesions	0.83	[0.6, 1.02]	0.196	0.82	[0.58, 1.01]	0.169	-2.88	-0.2	0.99
Baseline to year 2 NEDA	0.76	[0.21, 2.07]	0.619	0.78	[0.22, 2.15]	0.663	-2.91	-0.25	0.8
Baseline to year 3 therapeutic escalation	1.95	[0.73, 4.7]	0.153	2.17	[0.8, 5.37]	0.105	-3.08	0.78	0.82
Baseline to year 2 mean vitamin D level	1.04	[0.99, 1.09]	0.078	1.04	[0.99, 1.09]	0.096	-3.78	0.04	0.58
T25W Worsening									
Baseline to year 2 new T2 lesions	0.99	[0.91, 1.04]	0.718	0.99	[0.9, 1.04]	0.678	-1.22	-0.01	0.21
Baseline to year 2 NEDA	0.79	[0.39, 1.51]	0.488	0.79	[0.39, 1.51]	0.488	-1.19	-0.24	0.22
Baseline to year 3 therapeutic escalation	1.97	[1.03, 3.68]	0.035	2.01	[1.05, 3.8]	0.032	-1.41	0.7	0.3
Baseline to year 2 mean vitamin D level	1.01	[0.98, 1.04]	0.515	1.01	[0.98, 1.04]	0.544	-1.44	0.01	0.17
9HPT Worsening									
Baseline to year 2 new T2 lesions	1	[0.91, 1.06]	0.905	1	[0.91, 1.06]	0.947	-2	0	0.22
Baseline to year 2 NEDA	0.39	[0.12, 1.03]	0.088	0.41	[0.12, 1.09]	0.099	-1.86	-0.9	0.17
Baseline to year 3 therapeutic escalation	2.16	[1, 4.43]	0.042	2.29	[1.05, 4.79]	0.031	-2.22	0.83	0.3
Baseline to year 2 mean vitamin D level	1	[0.96, 1.05]	0.958	1	[0.96, 1.04]	0.966	-1.93	0	0.15

Regression analysis with and without propensity score adjustment on 4 clinical outcomes for RMS subjects: EDSS worsening from year 2 to year 10, PASAT worsening from year 2 to year 10, T25W worsening from year 2 to year 10, and 9HPT worsening from year 2 to year 10 (responses are underlined in the first column). A propensity score for treatment tier at baseline was included as a covariate in the adjusted model: $ln(p_{worse}/1\text{-}p_{worse}) = \beta_0 + \beta_{1(\text{predictor})}X_1 + \beta_{2(PS)}X_2$. The baseline to year 2 predictors that were tested for each outcome include ARR, 1 or more new T2 lesions, % BVL, grey matter volume change, white matter volume change, cortical grey matter volume change, cerebrospinal fluid volume change, change in T2 volume, EDSS worsening, EDSS worsening confirmed at year 3, worsening PASAT, worsening T25W, worsening 9HPT, 1 or more new gad+ lesions, NEDA, average vitamin D level, and baseline to year 3 increase in tier of therapy. % BVL is the percent brain volume loss between baseline and year 2 multiplied by -1. Associations with *p*-values < 0.05 and clinically relevant variables are shown. Mean 25-OH vitamin D levels are deseasonalized.