ORIGINAL ARTICLE

Evolution from a first clinical demyelinating event to multiple sclerosis in the REFLEX trial: Regional susceptibility in the conversion to multiple sclerosis at disease onset and its amenability to subcutaneous interferon beta-1a

Marco Battaglini[1] | Hugo Vrenken[2] | Riccardo Tappa Brocci[1] | Giordano Gentile[1] |
Ludovico Luchetti[1] | Adriaan Versteeg[2] | Mark S. Freedman[3] | Bernard M. J. Uitdehaag[4] |
Ludwig Kappos[5] | Giancarlo Comi[6] | Andrea Seitzinger[7] | Dominic Jack[8] |
Maria Pia Sormani[9] | Frederik Barkhof[2,10] | Nicola De Stefano[1]

[1]Department of Medicine, Surgery and Neurosciences, University of Siena, Siena, Italy

[2]Department of Radiology and Nuclear Medicine, Amsterdam UMC, Amsterdam, The Netherlands

[3]Department of Medicine, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada

[4]Department of Neurology, Amsterdam UMC, Location VUmc, Amsterdam, The Netherlands

[5]Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB) and Neurology, Departments of Head, Spine and Neuromedicine, Biomedical Engineering and Clinical Research, University Hospital, University of Basel, Basel, Switzerland

[6]Casa di Cura Privata del Policlinico, Università Vita-Salute San Raffaele, Milan, Italy

[7]Global Biostatistics, Merck Healthcare KGaA, Darmstadt, Germany

[8]Global Medical Affairs, Merck Serono Ltd (an affiliate of Merck KGaA), Feltham, UK

[9]Department of Health Sciences, University of Genoa, Genoa, Italy

[10]UCL Institutes of Neurology and Healthcare Engineering, London, UK

Correspondence
Marco Battaglini, Department of Medicine, Surgery and Neurosciences, University of Siena,Viale Bracci 2, 53100 Siena, Italy.
Email: battaglini@sienaimaging.it

Present address
Adriaan Versteeg, Department of Radiology and Nuclear Medicine, Biomedical Imaging Group Rotterdam, Erasmus MC-University Medical Center Rotterdam, Rotterdam, The Netherlands

Abstract

Background and purpose: In the REFLEX trial (ClinicalTrials.gov identifier: NCT00404352), patients with a first clinical demyelinating event (FCDE) displayed significantly delayed onset of multiple sclerosis (MS; McDonald criteria) when treated with subcutaneous interferon beta-1a (sc IFN β-1a) versus placebo. This *post hoc* analysis evaluated the effect of sc IFN β-1a on spatio-temporal evolution of disease activity, assessed by changes in T2 lesion distribution, in specific brain regions of such patients and its relationship with conversion to MS.

Methods: *Post hoc* analysis of baseline and 24-month magnetic resonance imaging data from FCDE patients who received sc IFN β-1a 44 μg once or three times weekly, or placebo in the REFLEX trial. Patients were grouped according to McDonald MS status

---

(converter/non-converter) or treatment (sc IFN β-1a/placebo). For each patient group, a baseline lesion probability map (LPM) and longitudinal new/enlarging and shrinking/disappearing LPMs were created. Lesion location/frequency of lesion occurrence were assessed in the white matter.

Results: At Month 24, lesion frequency was significantly higher in the anterior thalamic radiation (ATR) and corticospinal tract (CST) of converters versus non-converters ($p < 0.05$). Additionally, the overall distribution of new/enlarging lesions across the brain at Month 24 was similar in placebo- and sc IFN β-1a-treated patients (ratio: 0.95). Patients treated with sc IFN β-1a versus placebo showed significantly lower new lesion frequency in specific brain regions (cluster corrected): ATR ($p = 0.025$), superior longitudinal fasciculus ($p = 0.042$), CST ($p = 0.048$), and inferior longitudinal fasciculus ($p = 0.048$).

Conclusions: T2 lesion distribution in specific brain locations predict conversion to McDonald MS and show significantly reduced new lesion occurrence after treatment with sc IFN β-1a in an FCDE population.

INTRODUCTION

In the REFLEX trial (ClinicalTrials.gov identifier: NCT00404352), patients with a first clinical demyelinating event (FCDE) displayed significantly delayed onset of McDonald multiple sclerosis (MS) when treated with subcutaneous interferon beta-1a (sc IFN β-1a) versus placebo [1]. This was accompanied by lower occurrence and volume of new T2, T1 hypo-intense, and T1 gadolinium-enhancing (Gd+) lesions [2].

New advances in magnetic resonance imaging (MRI) analysis techniques facilitate further analysis of topographical lesion distribution and frequency. Proton density (PD) magnetic resonance subtraction imaging (SubI) is a relatively new, effective, imaging technique that allows monitoring of patient lesion evolution between time-points [3]. Lesion probability mapping (LPM) enables between-group comparisons of local lesion frequency throughout the brain using voxel-wise statistical analysis [4].

Multiple studies indicate that an increased white matter (WM) lesion load positively correlates with conversion to clinically definite MS (CDMS) [5-8]. Lesion location in specific brain regions has emerged as an important predictor of conversion to CDMS in patients with FCDE [9], as well as being a major factor determining the degree of disability and disease prognosis [9,10]. However, the relationship between the spatio-temporal dynamics of MS lesions and conversion to CDMS needs further elucidation.

This *post hoc* study of data from the REFLEX trial [1,2] aimed to assess, using LPM of SubI, whether inflammatory activity during the disease course is widespread across the brain or confined to specific areas. In doing so, this may allow for identification of specific brain areas that, if affected by an inflammatory activity, may predict the conversion to MS, and whether sc IFN β-1a affects the development of new/enlarging lesions or promotes the shrinking/disappearing of pre-existing lesions.

METHODS

This study used data from the phase III, placebo-controlled REFLEX clinical trial, which evaluated treatment with sc IFN β-1a 44 μg once (qw) or three times weekly (tiw) in patients with a FCDE. Details of this trial are described in detail elsewhere [1,2].

Patients

A total of 457 patients with a FCDE provided data for analysis. Patients were grouped according to their conversion to McDonald MS status (converted to McDonald MS vs. non-converted patients), which was defined by the 2005 McDonald criteria [11]. A secondary analysis grouping patients as converters or non-converters to CDMS using the Poser criteria was also performed (see Appendix S1) [12]. Patients were further analyzed according to their treatment regimen (sc IFN β-1a 44 μg qw or tiw vs. placebo). Those treated with sc IFN β-1a 44 μg qw or tiw were pooled into a single group and are referred to as sc IFN β-1a-treated patients throughout.

MRI data

MRI data from REFLEX [2] comprised a PD-, T2-, T1-, and Gd+T1-weighted image, as well as a T1-weighted brain mask image for each patient. In addition, for the purposes of this *post hoc* analysis, the Image Analysis Center at VU University Medical Center

(Amsterdam, The Netherlands) provided manually created, precise outlines of each PD-/T2-weighted lesion at screening (referred hereafter as baseline) and Month 24, and manually edited brain extraction masks originally obtained from a software library (FSL)-Brain Extraction Tool (BET).

The methodology for the creation and analysis of SubIs is described in full in Appendix S1. Briefly, prior to generating SubIs, slice-to-slice variation in signal intensity on the PD-weighted images was corrected. Baseline and Month 24 PD images were registered to a common halfway space using a similar procedure to that used in the FSL-SIENA software in which the required transformation matrix was determined from the corresponding T2-weighted images obtained from the same dual-echo acquisition. The screening scan images were subtracted from the corresponding Month 24 images to obtain the SubIs. To give a robust analysis, the SubIs were further normalized to account for the differences between the sites and MRI scanners from which the images were obtained.

Analysis of SubI lesion volume change was restricted to evaluation of manual lesion masks, and lesions were classified as follows. First, individual lesions were classified as *new*, *disappearing*, or *changing* based on comparison of the expert manual lesion outlines of both time-points. Specifically, if a lesion outline was present on the Month 24 image but not the baseline image, the lesion was classified as a *new* lesion; vice versa, a lesion outline present on the baseline image without a corresponding lesion outline on the Month 24 image was classified as a *disappearing* lesion. Finally, wholly or partially overlapping lesion outlines on both time-points' images were classified as *changing* lesions. Volume change was analyzed for individual lesions (see Appendix S1), and total lesion volume change was calculated from these individual changes. Within changing lesions, the net volume change thus obtained was used to determine whether to classify the lesion as a net *growing* or *shrinking* lesion. All voxels belonging to lesion outlines were assigned a label indicating the type of lesion for use in the LPM analyses.

Image post-processing

A study template representative of the study population was created as described in Appendix S1. This study-group-specific template was used for all the regional and voxel-wise analyses (baseline and longitudinal). For each patient group comparison (i.e., converters to McDonald MS versus non-converters, and sc IFN β-1a-treated versus placebo), a baseline new/enlarging and shrinking/disappearing lesion mask for each patient was registered to template space, followed by statistical analysis.

Global analysis

At baseline, between-group differences in age, sex, and lesion volume were assessed for patients converting to McDonald MS versus non-converters, and patients who were treated with sc IFN β-1a versus placebo. In the longitudinal analysis (baseline to Month 24), the number and volume of lesions were measured for each class of new/enlarging and shrinking/disappearing lesion masks.

Tract analysis

For each patient group, the numbers of new/enlarging and shrinking/disappearing lesions were determined for nine WM tracts from the Johns Hopkins University White-Matter Tractography Atlas provided with the FSL Library. The tracts analyzed were: anterior thalamic radiation (ATR), cingulum, cortical spinal tract (CST), forceps major, forceps minor, inferior fronto-occipital fasciculus (IFOF), inferior longitudinal fasciculus (ILF), superior longitudinal fasciculus (SLF), and uncinate fasciculus. For each mask, the number of connected clusters within each WM tract was counted; only clusters greater than three voxels were retained.

Voxel-wise analysis

For each patient group, a baseline LPM was generated by merging and averaging all standard space lesion masks, according to the procedure described by Di Perri and colleagues [4]. For the longitudinal analysis, LPMs of new/enlarging and shrinking/disappearing lesions were also generated. On the LPMs, the signal intensity at each voxel represents the probability of a voxel being a lesion.

Statistical analysis

For the global baseline analysis, a Mann–Whitney U test determined any statistically significant differences in age and lesions between groups. An Fisher exact test was used for between-group comparisons for sex (significance defined as $p < 0.05$). For the longitudinal global analysis, in terms of changes from baseline to Month 24, a between-group comparison for the number and volume change of each type of lesion was performed using a non-parametric permutation test (adjusted for age, sex, center, and conversion or treatment; significance defined as $p < 0.05$).

The longitudinal tract analysis used a non-parametric permutation test within the General Linear Model framework to assess the differences in number of new/enlarging and shrinking/disappearing lesions, between those converting to McDonald MS and non-converting patients, and between sc IFN β-1a- and placebo-treated patients, for each WM region (adjusted for age, sex, center, and conversion or treatment; significance defined as $p < 0.01$). To test whether differences in treatment effect (sc IFN β-1a vs. placebo) between tracts were likely real or related to statistical power, these age- and sex-adjusted comparisons were also performed using a mixed model adjusted for treatment-by-tract interactions (significance defined as $p < 0.01$).

Voxel-wise between-group differences in lesion frequency at baseline, adjusted for age and sex, and new lesion frequency at Month 24 (longitudinal analysis), adjusted for age, sex, center, and conversion or treatment, were analyzed using an unpaired t-test within the General Linear Model framework using the Randomise tool from the FMRIB software library (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki; significance defined as $p < 0.05$ and corrected for multiple comparisons). In the voxel-wise baseline and longitudinal analyses, the anatomical location of the local maxima within significant clusters were established using the probabilistic tractography WM atlas included with FSL.

RESULTS

Baseline analysis of lesion distribution and frequency

In total, 457 baseline MRI scans obtained from REFLEX trial participants were available for *post hoc* analysis; 307 were treated with sc IFN β-1a (i.e., sc IFN β-1a 44 μg qw or tiw) versus 150 who received placebo. Of these, 342 patients converted to McDonald MS and 115 were non-converters. LPMs were created for 442 patients: 333 converting to McDonald MS versus 109 non-converters, and 299 treated with sc IFN β-1a (qw $n = 142$; tiw: $n = 157$) versus 143 placebo recipients. Fifteen scans were discarded at baseline due to errors of nonlinear registration on the specific study-group template.

The baseline demographics and MRI characteristics of the patients included in the LPM analyses stratified by conversion to McDonald MS status and treatment group are presented in Table 1.

Patients converting to McDonald MS versus non-converters

Of the 442 patients with MRI scans available at baseline, those who subsequently converted to McDonald MS were significantly younger than those who did not convert (mean age 30.5 vs. 32.6 years; $p < 0.03$) and had a significantly greater overall mean lesion volume (3.92 vs. 1.90 cm^3; $p < 0.001$). No significant sex-related differences were apparent for those converting to McDonald MS versus non-converting groups (Table 1).

Using the voxel-wise approach, baseline LPMs showed that there was little difference between the maximum local probability of lesions for patients who converted to McDonald MS (15.9%) and those who did not convert (13.8%; Figure 1a,b). However, the overall distribution of lesions across the brain differed between the groups, and the number of cerebral voxels occupied by lesions was 2.8-fold larger in those converting to McDonald MS compared with the non-converting group. In addition, a higher frequency of lesions ($p < 0.05$, cluster-corrected) was present in the projection, association, and commissural WM tracts of patients who converted to McDonald MS compared with non-converters

(Figure 1c). In particular, lesion frequency was higher in the ATR (left and right), CST (left and right), forceps major (right), cingulum (right), and IFOF (left) tracts of patients who converted to McDonald MS ($p < 0.001$; Table S1).

sc IFN β-1a- versus placebo-treated patients

Among the 442 patients with MRI scans available at baseline, there were no significant differences in mean age (31.1 vs. 30.8 years), proportion of females (63.2% vs. 66.4%) or mean lesion volume (3.58 vs. 3.08 cm^3) between the sc IFN β-1a- and placebo-treated groups (Table 1).

Voxel-wise analysis of baseline LPMs showed that the maximum local probability of lesions at baseline was comparable between sc IFN β-1a-treated (15.4%) and placebo-treated (15.3%) patients, and in both groups reached the left side of the CST. The overall lesion distribution across the brain was similar in both groups, with no observed differences in the frequency of lesions in specific WM tracts (Figure 1d,e).

Longitudinal analysis of lesion distribution and frequency

Longitudinal analysis of lesion occurrence utilized 407 SubI scans, including 305 patients converting to McDonald MS versus 102 non-converters, and 273 patients treated with sc IFN β-1a versus 134 placebo recipients. Longitudinal MRI data from 35 patients were not included because 17 subtraction images failed during creation, two had artefacts, two had errors in registration, two were mutually inverted in orientation, and 12 images from Month 24 were unavailable.

The baseline patient demographics and MRI characteristics for this population are summarized in Table S2.

Patients converting to McDonald MS versus non-converters

Patients who converted to McDonald MS had a higher mean (±standard deviation [SD]) total number of (6.48 ± 9.3) and volume increase (0.77 ± 1.3 cm^3) due to new/enlarging lesions compared with patients who did not convert (1.15 ± 2.1 and 0.10 ± 0.2 cm^3, respectively; both $p < 0.001$), when adjusted for age, sex, center, and treatment. Patients who converted to McDonald MS also had a higher mean (±SD) number of (5.41 ± 6.9) and a larger mean volume decrease (0.85 ± 1.8 cm^3) due to shrinking/disappearing lesions compared with patients who did not convert (1.84 ± 3.2 and 0.39 ± 1.2 cm^3, respectively; both $p < 0.001$), when adjusted for age, sex, center, and treatment.

Compared with patients who converted to McDonald MS, the number of new/enlarging and shrinking/disappearing lesions

TABLE 1 Patient demographics and magnetic resonance imaging characteristics in the baseline lesion probability map analysis

Parameter	Converters to CDMS (n = 333)	Non-converters to CDMS (n = 109)	Placebo (n = 143)	sc IFN β-1a (n = 299)
Baseline demographics				
Age, years				
Mean ± SD	30.5 ± 8.1	32.6 ± 8.6	30.8 ± 8.0	31.1 ± 8.4
Median (Q1, Q3)	29.0 (24.0, 36.0)	32.0 (25.0, 38.0)	30 (24.0, 37.0)	30.0 (24.0, 37.0)
Female, n (%)	207 (62.2)	77 (70.6)	95 (66.4)	189 (63.2)
Classification of FCDE as monofocal[a], n (%)	176 (52.9)	78 (71.6)	81 (56.6)	173 (57.9)
Steroid use at FCDE, n (%)	233 (70.0)	76 (69.7)	97 (67.8)	212 (70.9)
EDSS score				
Mean ± SD	1.56 ± 0.74	1.36 ± 0.82	1.47 ± 0.75	1.53 ± 0.77
Median (Q1, Q3)	1.50 (1.00, 2.00)	1.50 (1.00, 2.00)	1.50 (1.00, 2.00)	1.50 (1.00, 2.00)
MRI characteristics				
T1 Gd+ lesions, n	524	45	147	422
Mean ± SD	1.57 ± 3.1	0.41 ± 1.0	1.03 ± 1.9	1.41 ± 3.1
Median (Q1, Q3)	0.0 (0.0, 2.0)	0.0 (0.0, 0.0)	0.0 (0.0, 1.0)	0.0 (0.0, 1.0)
Presence of at least 1 T1 Gd+ lesion, n (%)	161 (48.3)	30 (27.5)	57 (39.9)	134 (44.8)
T1 Gd+ lesion volume, mm^3				
Mean ± SD	198.83 ± 554.8	94.13 ± 355.9	145.98 ± 441.9	185.94 ± 546.2
Median (Q1, Q3)	0.00 (0.00, 128.8)	0.00 (0.00, 0.00)	0.00 (0.00, 108.2)	0.00 (0.00, 91.6)
T1 hypo-intense lesions, n	2162	364	728	1798
Mean ± SD	6.49 ± 7.7	3.34 ± 4.9	5.09 ± 7.4	6.01 ± 7.2
Median (Q1, Q3)	4.0 (1.0, 8.3)	1.0 (0.0, 4.3)	3.0 (0.3, 7.8)	4.0 (1.0, 8.0)
T1 hypo-intense lesion volume, mm^3				
Mean ± SD	772.8 ± 1125.9	466.8 ± 1013.1	623.7 ± 1008.0	732.6 ± 1149.8
Median (Q1, Q3)	343.3 (83.0, 989.9)	80.1 (0.0, 383.2)	177.4 (4.3, 778.2)	263.2 (52.2, 935.5)
T2 lesions, n	8502	1289	2829	6962
Mean ± SD	25.53 ± 21.3	11.83 ± 10.8	19.78 ± 19.6	23.28 ± 20.3
Median (Q1, Q3)	19 (11.0, 33.0)	8 (4.0, 17.0)	14 (6.3, 25.8)	18 (8.0, 32.0)
≥9 T2 lesions, n (%)	277 (83.2)	45 (41.3)	99 (69.2)	223 (74.6)
T2 lesion volume, mm^3				
Mean ± SD	3915.1 ± 4230.5	1899.4 ± 2753.5	3077.3 ± 3792.1	3581.0 ± 4107.6
Median (Q1, Q3)	2429.1 (1035.0, 5035.4)	755.3 (252.5, 2150.8)	1436.3 (640.2, 3963.9)	2131.5 (783.3, 4609.0)
Normalized brain volume, cm^3				
Mean ± SD	1535.9 ± 69.5	1540.6 ± 66.2	1543.3 ± 64.1	1534.1 ± 70.7
Median (Q1, Q3)	1544.7 (1490.5, 1582.7)	1533.4 (1505.9, 1585.7)	1546.8 (1497.8, 1588.7)	1536.9 (1493.2, 1580.1)

Abbreviations: CDMS, clinically definite multiple sclerosis; EDSS, Expanded Disability Status Scale; FCDE, first clinical demyelinating event; Gd+, gadolinium-enhancing; IFN, interferon; LPM, lesion probability map; MRI, magnetic resonance imaging; Q, quartile; sc, subcutaneous; SD, standard deviation.

[a]According to the adjudication committee.

(adjusted for age, sex, center, and treatment) was lower for those who did not convert in all WM tracts analyzed ($p < 0.01$ and $p < 0.05$, respectively; Table 2).

In the voxel-wise analysis, there was no difference between the maximum local probability of new/enlarging lesions for patients who converted to McDonald MS (3.28%) compared with those who did not convert (2.94%; Figure 2a,b). However, the distribution of new lesions across the brain was very different in those patients who converted to McDonald MS versus patients who did not. The number of cerebral voxels occupied by new/enlarging lesions in the converting to McDonald MS group was 14.6-fold greater than in the non-converting group. Patients who

FIGURE 1 Baseline T2-weighted (T2-W) lesion probability maps (LPMs). Baseline T2-W LPMs in standard space for patients who converted to McDonald multiple sclerosis (MS) (n = 333; (a)) and those who did not convert (n = 109; (b)), and for placebo-treated (n = 144; (d)) and subcutaneous interferon beta-1a (sc IFN β-1a)-treated patients (n = 298; bottom panel, (e)). The colour overlay created on the top of the Montreal Neurological Institute (MNI) standard brain shows the probability of each voxel containing a lesion in each patient group; the colour bar denotes the probability range. Results of the frequency comparison between patients who converted to McDonald MS and those who did not convert (c). The colour overlay created on the top of the MNI standard brain registered on the study-specific template shows voxels with a probability of being lesional that was significantly higher (p < 0.05, cluster corrected) in the converting than non-converting group; the colour bar denotes the p value range [Colour figure can be viewed at wileyonlinelibrary.com]
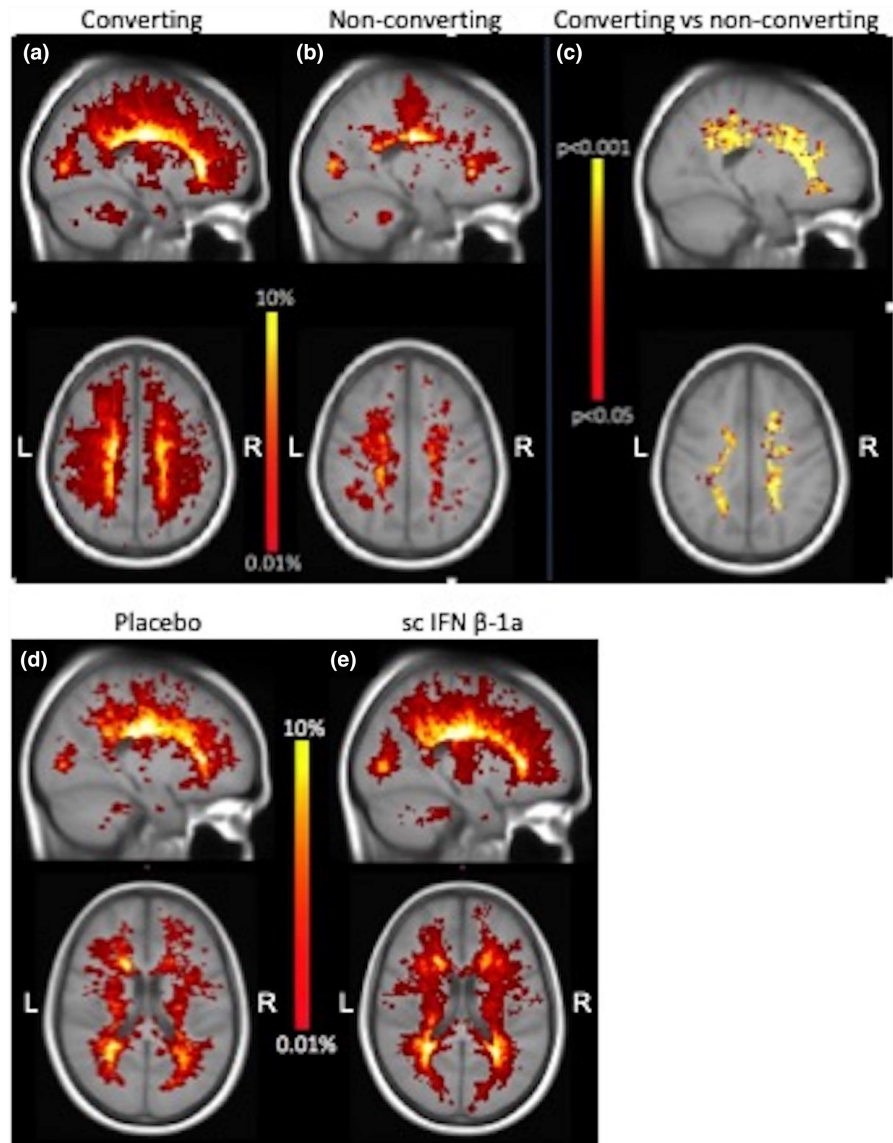
converted to McDonald MS had a higher new/enlarging lesion frequency than those who did not convert in the projection, association, and commissural WM tracts (p < 0.05, cluster-corrected; Figure 2c). New lesion frequency was higher in the ATR (left) and CST (right) tracts of patients who converted to McDonald MS compared with those who did not convert (p < 0.05; Table S3). No differences in lesion distribution were observed for shrinking/disappearing lesions in patients who converted to McDonald MS compared with those who did not convert.

Additional voxel-wise analyses were performed comparing patients who converted to CDMS (n = 110) with those who did not convert (n = 296; see Appendix S1). Although patient distribution between converting and non-converting groups differed using the two definitions of conversion, new brain lesions were more frequently found in converting than non-converting patients in the same regions using both definitions. The affected regions were less extensive when conversion was defined as CDMS compared with McDonald MS.

sc IFN β-1a- versus placebo-treated patients

Patients treated with sc IFN β-1a had a lower mean (±SD) number of new/enlarging lesions (n = 273; 6.9 ± 10) compared with placebo-treated patients (n = 134; 10.9 ± 16, p < 0.01), while the lesion volume increase of new/enlarging lesions showed a trend towards reduction in treated patients (0.47 ± 0.99 vs. 0.88 ± 1.48 cm^3, p = 0.067), after adjusting for age, sex, center, and conversion. There was no difference in the mean (±SD) number and volume change of shrinking/disappearing lesions for sc IFN β-1a-treated (9.7 ± 12 and −0.71 ± 1.53 cm^3, respectively) and placebo-treated patients (8.6 ± 11.0 and −0.80 ± 1.87 cm^3, respectively).

The number of new/enlarging lesions was significantly reduced in sc IFN β-1a-treated patients compared with placebo-treated patients in the ATR, cingulum, CST, ILF, and SLF (n = 407; Table 3). When analyses were adjusted for treatment-by-tract interactions, the differences between sc IFN β-1a- and

TABLE 2 Percentage reduction in the number of new/enlarging and shrinking/disappearing lesions in specific white matter regions of patients who converted to McDonald multiple sclerosis versus patients who did not convert

Tract	New/enlarging lesions, mean ± SD			Shrinking/disappearing lesions, mean ± SD		
	Patients converting to McDonald MS	Non-converting patients	Percentage reduction, %	Patients converting to McDonald MS	Non-converting patients	Percentage reduction, %
ATR	1.12 ± 1.9	0.19 ± 0.6***	83	0.89 ± 1.4	0.33 ± 1.0***	63
Cingulum	0.35 ± 1.0	0.02 ± 0.1***	94	0.26 ± 0.7	0.05 ± 0.2***	81
CST	0.84 ± 1.3	0.16 ± 0.4***	81	0.76 ± 1.2	0.23 ± 0.5***	70
Forceps major	0.50 ± 1.0	0.13 ± 0.6***	74	0.36 ± 0.8	0.18 ± 0.5**	50
Forceps minor	0.30 ± 0.8	0.01 ± 0.1***	97	0.22 ± 0.6	0.06 ± 0.3***	73
IFOF	1.06 ± 1.8	0.22 ± 0.7***	79	0.66 ± 1.2	0.25 ± 0.8***	62
ILF	0.79 ± 1.5	0.14 ± 0.4***	82	0.46 ± 1.0	0.20 ± 0.8**	56
SLF	1.34 ± 2.2	0.21 ± 0.6***	84	0.95 ± 1.7	0.26 ± 0.6***	73
UF	0.12 ± 0.4	0.01 ± 0.1**	92	0.05 ± 0.2	0.01 ± 0.1*	80

Note: Regions considered were based on the Johns Hopkins University White-Matter Tractography Atlas.

Abbreviations: ATR, anterior thalamic radiation; CST, cortical spinal tract; IFOF, inferior fronto-occipital fasciculus; ILF, inferior longitudinal fasciculus; MS, multiple sclerosis; SD, standard deviation; SLF, superior longitudinal fasciculus; UF, uncinate fasciculus.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$ versus converting group.

placebo-treated patients in the number of such lesions were no longer significant ($p = 0.9$). This suggests that differences were widespread rather than tract-related, not reaching statistical significance in those tracts that, irrespective of treatment, showed a low frequency of new/enlarging lesions. Within the different tracts analyzed, no significant differences between sc IFN β-1a- and placebo-treated patients were observed for the number of shrinking/disappearing lesions.

In voxel-wise analyses, the maximum local probability of new/enlarging lesions was almost 2-fold higher for placebo recipients (5.2%) compared with sc IFN β-1a-treated patients (3.2%). The overall volume of new/enlarging lesions was similar in sc IFN β-1a- and placebo-treated patients (153.3 vs. 146.1 cm³; ratio for placebo- to sc IFN β-1a-treated patients: 0.95; Figure 2d,e). Patients treated with sc IFN β-1a showed lower new lesion frequency than placebo-treated patients in specific brain regions (cluster-corrected) including the ATR ($p = 0.025$), SLF ($p = 0.042$), CST ($p = 0.048$), and ILF ($p = 0.048$; Figure 3). In specific areas of the CST (left and right), ATR (left), ILF (right), forceps major (left), and SCR (right), new lesion frequency was lower in sc IFN β-1a-treated versus placebo-treated patients (Table S4). In specific areas of the forceps major (right) and SCR (left), new lesion frequency was greater in sc IFN β-1a-treated compared with placebo-treated patients (Table S5).

No significant differences were observed in the voxel-wise analysis for sc IFN β-1a- and placebo-treated patients for shrinking/disappearing lesions.

Additional voxel-wise analyses, performed with adjustment for conversion to CDMS instead of McDonald MS, showed that the same regions of importance were found with respect to those obtained by previous analysis but with smaller (half) spatial extension (see Appendix S1).

DISCUSSION

The definition of conversion to McDonald criteria MS is based on radiological disease progression. Thus, it is not surprising in the present study that new/enlarging lesion accumulation over 2 years was globally more prominent in patients who converted to McDonald MS than in those who did not convert. However, we identified specific WM regions where inflammatory activity related to MS was more pronounced, using LPM methodology to map the distribution of new/enlarging lesions from SubIs, highlighting the clinically relevant role of specific tracts such as the CST and ATR. Interestingly, the absence of treatment-by-tract interactions strongly suggests that suppression of this activity does not follow specific neural pathways, thus supporting the view that although microstructural changes in MS are widespread across selective brain areas, there is no drug-specific tract activity.

To test whether the clinical transition from FCDE to MS was driven by a higher presence of new/enlarging lesions in selective brain areas, we repeated our analyses both at global and voxel-wise levels using CDMS Poser criteria [12], which is a more clinical definition of conversion and is not related to radiological features. Although using this definition of conversion changed the composition of the analysis groups, results were in general agreement with those using McDonald criteria. Pathological inflammatory (lesional) activity was again more pronounced in patients who converted than in those who did not convert, and, most importantly, the areas involved in driving the conversions were approximately the same as those obtained when using the McDonald criteria, although with a smaller spatial extension. Our results are also consistent with those of previous studies in patients with a FCDE that compared topographical lesion frequency in those who converted to CDMS versus non-converting patients [9,13–15]. Furthermore, in patients with

FIGURE 2 Longitudinal new/enlarging T2-weighted (T2-W) lesion probability maps (LPMs). New/enlarging T2-W LPMs in standard space for patients who converted to McDonald multiple sclerosis (MS) (*n* = 305; (a)) and those who did not convert (*n* = 102; (b)), and for placebo-treated (*n* = 134; (d)) and subcutaneous interferon beta-1a (sc IFN β-1a)-treated patients (*n* = 273; (e)). The colour overlay created on the top of the Montreal Neurological Institute (MNI) standard brain shows the probability of each voxel containing a lesion in each patient group; the colour bar denotes the probability range. The blue mask overlay created on top of the MNI standard brain represents the cortical spinal tract region (d/e). Results of the frequency comparison between the two groups (c). The colour overlay created on the top of the MNI standard brain registered on the study specific template shows voxels with a probability of being lesional significantly higher (*p* < 0.05, cluster corrected) in the converting than non-converting group; the colour bar denotes the *p* value range [Colour figure can be viewed at wileyonlinelibrary.com]
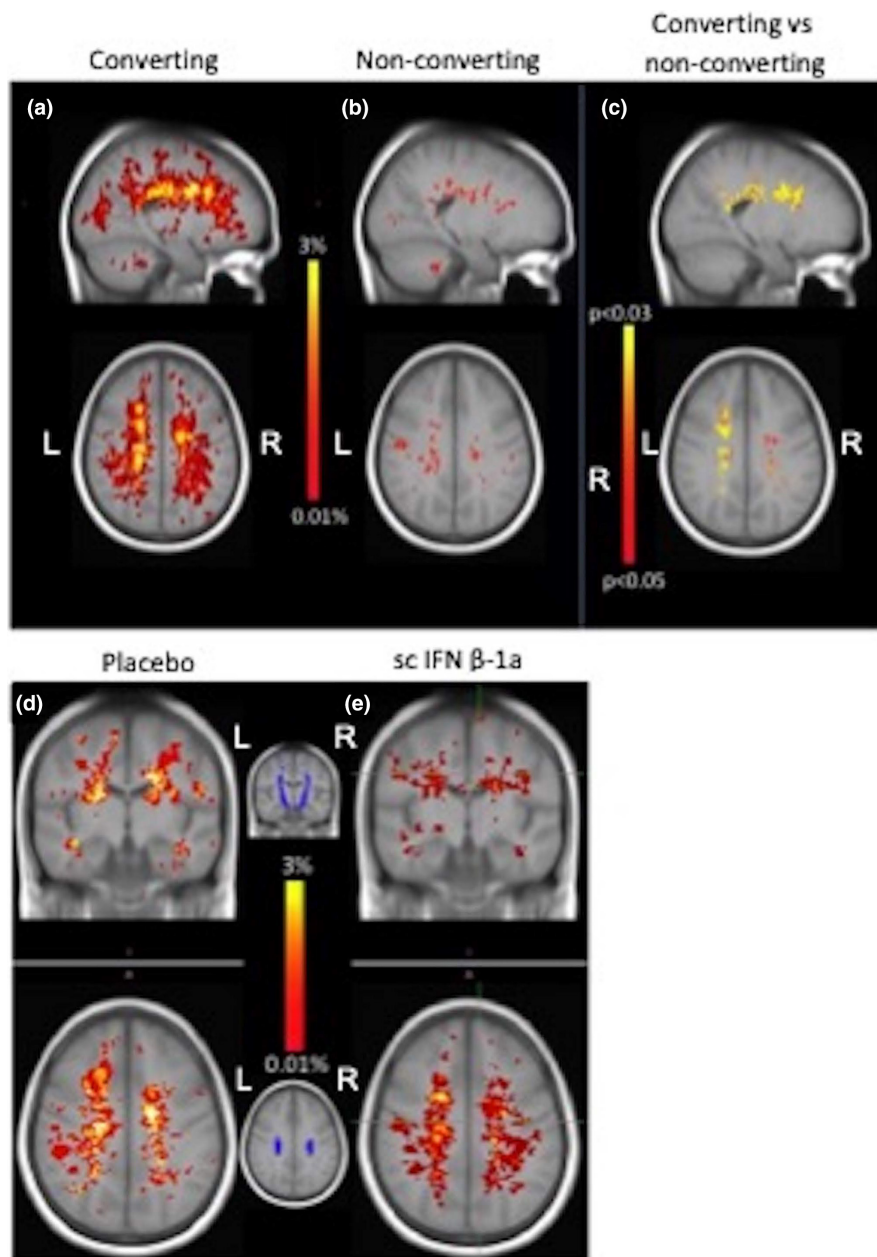
early relapsing-remitting MS, sustained disease progression, when compared with stable disease, was initially associated with significantly greater T2 lesion volume and whole brain, subcortical deep grey matter, WM and cortical volume, and significantly larger decreases in whole brain, cortex, grey matter, and thalamus volume over a 5-year period [16].

At follow-up, treatment with sc IFN β-1a had no effect on shrinking/disappearing lesions but significantly reduced the number and volume of new/enlarging lesions compared with placebo. These data align with the primary findings from the REFLEX clinical trial, which used conventional MRI techniques to determine the number and volume of new T2 lesions in patients treated with sc IFN β-1a qw or tiw versus placebo [2]. The lack of treatment effect on shrinking/disappearing lesions would be difficult to detect given that the voxel-wise analysis was performed on areas of

low lesion frequency in patients with a relatively low lesion load at baseline. However, since it is thought that different pathological mechanisms are responsible for new/enlarging and shrinking/disappearing lesions [17], it is possible that sc IFN β-1a mostly reduces inflammation rather than directly promoting repair of tissue damage, although we are unable to confirm this based on our findings. Our study identified only a few specific WM regions where inflammatory activity was more pronounced in the placebo-treated than in the sc IFN β-1a-treated group, but the absence of treatment-by-tract interactions suggests that sc IFN β-1a targets brain areas with a high presence of inflammatory activity irrespective of tract. These results complement those of Vrenken et al. [18] who showed that treatment with sc IFN β-1a 44 µg tiw, but not qw, reduced evolution of the absolute number of new lesions into black holes in patients with a FCDE.

Tract	Lesions in sc IFN β-1a-treated patients, mean ± SD	Lesions in placebo-treated patients, mean ± SD	Percentage reduction, %
ATR	0.72±1.5	1.22±2.0*	41
Cingulum	0.17±0.6	0.46±1.2**	63
CST	0.53±1.1	0.96±1.4**	45
Forceps Major	0.35±0.8	0.51±1.0	31
Forceps Minor	0.18±0.6	0.34±0.9	47
IFOF	0.71±1.5	1.13±1.9	37
ILF	0.49±1.1	0.91±1.6*	46
SLF	0.83±1.6	1.53±2.5**	46
UF	0.08±0.4	0.12±0.3	33

TABLE 3 Number and percentage reduction in T2 lesions in specific white matter regions after treatment with subcutaneous interferon β-1a versus placebo

Note: Regions considered were based on the Johns Hopkins University White-Matter Tractography Atlas.

Abbreviations: ATR, anterior thalamic radiation; CST, cortical spinal tract; IFOF, inferior fronto-occipital fasciculus; IFN, interferon; ILF, inferior longitudinal fasciculus; sc, subcutaneous; SD, standard deviation; SLF, superior longitudinal fasciculus; UF, uncinate fasciculus.

*$p < 0.05$, **$p < 0.01$ versus sc IFN β-1a-treated patients.
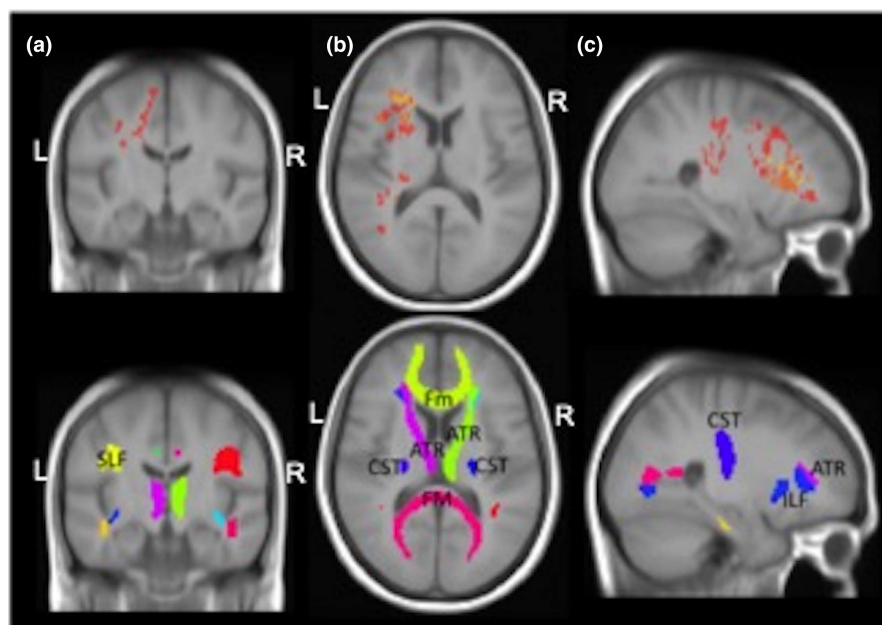
FIGURE 3 Voxel cluster analysis in placebo-treated versus subcutaneous interferon beta-1a (sc IFN β-1a)-treated patients. Clusters of voxels where new brain lesions were more frequent in placebo- versus sc IFN β-1a-treated patients (controlled for age, sex, and center; top row). Colour overlap created on top of the Montreal Neurological Institute standard brain registered on the study-specific template showing the tracts considered (bottom row). ATR, anterior thalamic radiation; CST, cortical spinal tract; FM, forceps major; Fm, forceps minor; ILF, inferior longitudinal fasciculus; SLF, superior longitudinal fasciculus [Colour figure can be viewed at wileyonlinelibrary.com]

Strengths and limitations

The present analysis used MRI-based LPMs generated from SubI to assess treatment effect in a clinical trial. SubI has multiple advantages for monitoring lesion evolution in WM regions of the brain, as it is a more sensitive technique compared with conventional, single time-point imaging [3]. This is particularly important in a clinical setting for deciphering treatment effects on disease activity. SubI also allows in-depth characterization of lesion activity, by enabling the differentiation between new and enlarging, shrinking, and disappearing lesions. This is advantageous over conventional MRI techniques that only measure net volume change [3]. In addition, use of SubI is less time consuming compared with other methods, since measurements are obtained from a single image rather than analysis of individual images at different time-points. Since SubI enhances study power, fewer subjects are required than with conventional techniques [19]. Nevertheless, the findings of this study were based on a very complete set of analyses obtained in parallel from a well-defined set of more than 100 patients with early MS.

Despite these strengths, some limitations must be acknowledged. A very precise overlap of lesions was not required in the tract analysis, and lesions were only counted provided they resided in a specific tract. A clear mechanistic explanation for the observed anatomical distribution of lesion changes is also lacking. In addition, these analyses were *post hoc* and the calculated *p* values are therefore descriptive in nature with no adjustment for multiple testing. Finally, a plausible explanation of our results cannot exclude the greater statistical power of the LPM analysis in those regions where the incidence of injury is greatest. Figure 4 shows the LPM of the 407 patients employed in the longitudinal lesion analysis both at
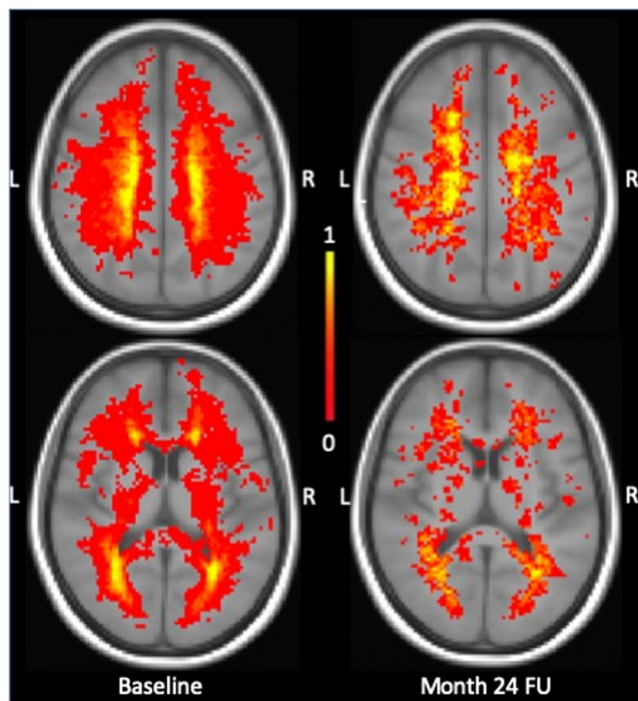
FIGURE 4 Comparison of lesion probability map (LPM) distributions at baseline (left), using T2-weighted (T2-W) lesion masks, and at Month 24 (right), using new/enlarging T2-W lesion masks. Given the differences in magnitude of the two lesion maps, each of them was scaled to the highest probability value. Thus the maximum peaks of both the LPM is 1, and the yellow colour indicates areas where the probability of being a lesion (left) or new lesion (right) is similar to that voxel with the higher probability of being a lesion within the group [Colour figure can be viewed at wileyonlinelibrary.com]

baseline, using the baseline T2-W lesion masks, and at Month 24, using the new enlarging T2-W lesion masks. It can be noted that although not strictly concordant at the voxel level, there is an accumulation of lesions in those areas affected at baseline. Further investigations are also warranted to assess the group-level sensitivity of the LPM analysis to detect significant differences in regional lesion distribution, such as use of synthetic data to control the level of probability in lesional voxel distributions and the spatial extent of lesion size.

CONCLUSIONS AND FUTURE DIRECTIONS

We found that microstructural damage in specific brain areas, as reflected by the accumulation of new/enlarging lesions, may play an important role in the transition to MS, both clinically and radiologically, in patients at disease onset. Our results also showed that treatment with sc IFN β-1a reduced the development of new/enlarging lesions over time, and that this effect became detectable at the voxel level only in those regions with a high presence of inflammatory activity. By contrast, sc IFN β-1a may not play a role in the mechanisms of repair, as shown by the absence of differences in shrinking/

disappearing lesions between the treated and non-treated groups both at global and voxel-wise levels.

In parallel work on the same dataset [18], new black hole lesions were found to be less likely to develop in sc IFN β-1a-treated patients, although this seemed to be driven by an overall reduction of development of new lesions. This is in line with the findings reported here, which suggest that sc IFN β-1a does not promote lesion repair. Further analyses could be performed to investigate the spatio-temporal relationship between new lesion appearance and atrophy in driving the conversion from FCDE to MS, and how long-term treatment with sc IFN β-1a can delay this conversion. In this regard, studies with a longer follow-up are needed to evaluate the predictive value of the early changes observed in these analyses, and to assess how they are related to conversion to McDonald MS and CDMS. The MRI-based LPM generated from the SubI approach used in this manuscript could be used in further studies given its sensitivity and convenience.

CONFLICT OF INTEREST

HV has received research support from Merck, Novartis, Pfizer, and Teva, consulting fees from Merck, and speaker honoraria from Novartis; all funds were paid to his institution. AV has received research support from Merck. MSF has received honoraria or consultation fees from Alexion, Atara Biotherapeutics, Bayer, BeiGene, BMS (Celgene), EMD Inc., Canada (an affiliate of Merck KGaA), Hoffmann La-Roche, Janssen (J&J), Merck, Novartis, Pendopharm, and Sanofi-Genzyme; has been a member of a company advisory board, board of directors, or other similar group for Alexion, Atara Biotherapeutics, Bayer, BeiGene, BMS (Celgene), Clene Nanomedicine, Hoffmann La-Roche, Janssen (J&J), McKesson, Merck, Novartis, and Sanofi-Genzyme; and has participated in a company sponsored speaker's bureau for EMD Serono Inc., USA (an affiliate of Merck KGaA) and Sanofi-Genzyme. BMJU has received consultation fees from Biogen, Genzyme, Merck, Novartis, Roche, and Teva. LK's institution (University Hospital, University of Basel) has received the following exclusively for research support: steering committee, advisory board, and consultancy fees (Actelion [Janssen/J&J], Bayer, Biogen, BMS, Genzyme, Janssen, Merck, Novartis, Roche, Sanofi, Santhera, and TG Therapeutics); speaker fees (Bayer, Biogen, Merck, Novartis, Roche, and Sanofi); support of educational activities (Allergan, Bayer, Biogen, CSL Behring, Desitin, Genzyme, Merck, Novartis, Roche, Pfizer, Sanofi, Shire, and Teva); license fees for Neurostatus products;

AUTHOR CONTRIBUTIONS

Marco Battaglini: Conceptualization (equal); Investigation (equal); Methodology (equal); Software (equal); Supervision (equal); Writing – original draft (equal). **Hugo Vrenken:** Conceptualization (equal); Investigation (equal); Methodology (equal); Software (equal); Supervision (equal); Writing – original draft (equal). **Riccardo Tappa Brocci:** Validation (equal); Writing – review & editing (equal). **Giordano Gentile:** Software (equal); Validation (equal); Writing – review & editing (equal). **Ludovico Luchetti:** Software (equal); Validation (equal); Writing – review & editing (equal). **Adriaan Versteeg:** Writing – review & editing (equal). **Mark S. Freedman:** Writing – review & editing (equal). **Bernard M. J. Uitdehaag:** Writing – review & editing (equal). **Ludwig Kappos:** Writing – review & editing (equal). **Giancarlo Comi:** Writing – review & editing (equal). **Andrea Seitzinger:** Writing – review & editing (equal). **Dominic Jack:** Writing – review & editing (equal). **Maria Pia Sormani:** Writing – review & editing (equal). **Frederik Barkhof:** Conceptualization (equal); Investigation (equal); Supervision (equal); Writing – review & editing (equal). **Nicola De Stefano:** Conceptualization (equal); Investigation (equal); Supervision (equal); Writing – review & editing (equal).

ETHICAL APPROVAL

This *post hoc* study used data from the REFLEX trial, which was undertaken in compliance with the Declaration of Helsinki and standards of Good Clinical Practice according to the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. For each center, the relevant institutional review board or independent ethics committee reviewed and approved the trial protocol, patient information leaflet, informed consent forms, and investigator brochure. Written informed consent was obtained for all patients at the screening visit.

DATA AVAILABILITY STATEMENT

Any requests for data by qualified scientific and medical researchers for legitimate research purposes will be subject to Merck's Data Sharing Policy. All requests should be submitted in writing to Merck's data sharing portal https://www.merckgroup.com/en/research/our-approach-to-research-and-development/healthcare/clinical-trials/commitment-responsible-data-sharing.html. When Merck has a co-research, co-development, or co-marketing or co-promotion agreement, or when the product has been out-licensed, the responsibility for disclosure might be dependent on the agreement between parties. Under these circumstances, Merck will endeavour to gain agreement to share data in response to requests.

ORCID

Marco Battaglini https://orcid.org/0000-0002-9188-4408
Giordano Gentile https://orcid.org/0000-0002-8481-2801
Giancarlo Comi https://orcid.org/0000-0002-6989-1054
Andrea Seitzinger https://orcid.org/0000-0001-8088-9695
Dominic Jack https://orcid.org/0000-0001-8629-553X
Frederik Barkhof https://orcid.org/0000-0003-3543-3706

REFERENCES

1. Comi G, De Stefano N, Freedman MS, et al. Comparison of two dosing frequencies of subcutaneous interferon beta-1a in patients with a first clinical demyelinating event suggestive of multiple sclerosis (REFLEX): a phase 3 randomised controlled trial. *Lancet Neurol.* 2012;11:33-41.
2. De Stefano N, Comi G, Kappos L, et al. Efficacy of subcutaneous interferon beta-1a on MRI outcomes in a randomised controlled trial of patients with clinically isolated syndromes. *J Neurol Neurosurg Psychiatry.* 2014;85:647-653.
3. Duan Y, Hildenbrand PG, Sampat MP, et al. Segmentation of subtraction images for the measurement of lesion change in multiple sclerosis. *Am J Neuroradiol.* 2008;29:340-346.
4. Di Perri C, Battaglini M, Stromillo ML, et al. Voxel-based assessment of differences in damage and distribution of white matter lesions between patients with primary progressive and relapsing-remitting multiple sclerosis. *Arch Neurol.* 2008;65:236-243.
5. Tintore M, Rovira A, Rio J, et al. Baseline MRI predicts future attacks and disability in clinically isolated syndromes. *Neurology.* 2006;67:968-972.
6. Ruet A, Deloire MS, Ouallet JC, Molinier S, Brochet B. Predictive factors for multiple sclerosis in patients with clinically isolated spinal cord syndrome. *Mult Scler.* 2011;17:312-318.
7. The Optic Neuritis Study Group. Multiple sclerosis risk after optic neuritis: final optic neuritis treatment trial follow-up. *Arch Neurol.* 2008;65:727-732.
8. Fisniku LK, Brex PA, Altmann DR, et al. Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain.* 2008;131:808-817.
9. Giorgio A, Battaglini M, Rocca MA, et al. Location of brain lesions predicts conversion of clinically isolated syndromes to multiple sclerosis. *Neurology.* 2013;80:234-241.

10. Vellinga MM, Geurts JJ, Rostrup E, et al. Clinical correlations of brain lesion distribution in multiple sclerosis. *J Magn Reson Imaging*. 2009;29:768-773.

11. Polman CH, Reingold SC, Edan G, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the "McDonald Criteria". *Ann Neurol*. 2005;58:840-846.

12. Poser CM, Paty DW, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Ann Neurol*. 1983;13:227-231.

13. Bendfeldt K, Taschler B, Gaetano L, et al. MRI-based prediction of conversion from clinically isolated syndrome to clinically definite multiple sclerosis using SVM and lesion geometry. *Brain Imaging Behav*. 2019;13:1361-1374.

14. Dalton CM, Bodini B, Samson RS, et al. Brain lesion location and clinical status 20 years after a diagnosis of clinically isolated syndrome suggestive of multiple sclerosis. *Mult Scler*. 2012;18:322-328.

15. Gaetano L, Magnusson B, Kindalova P, et al. White matter lesion location correlates with disability in relapsing multiple sclerosis. *Mult Scler J Exp Transl Clin*. 2020;6:2055217320906844.

16. Zivadinov R, Bergsland N, Dolezal O, et al. Evolution of cortical and thalamus atrophy and disability progression in early relapsing-remitting MS during 5 years. *Am J Neuroradiol*. 2013;34:1931-1939.

17. Pongratz V, Schmidt P, Bussas M, et al. Prognostic value of white matter lesion shrinking in early multiple sclerosis: an intuitive or naïve notion? *Brain Behavior*. 2019;9:e01417.

18. Vrenken H, de Vos ML, Battaglini M, et al. Evolution of new lesions and its temporal patterns in patients with clinically isolated syndrome treated with subcutaneous interferon beta-1a [P1025]. *Mult Scler*. 2017;23(S3):529-530.

19. Moraal B, Meier DS, Poppe PA, et al. Subtraction MR images in a multiple sclerosis multicenter clinical trial setting. *Radiology*. 2009;250:506-514.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.