

# Disease progression in the first 5 years of treatment in multiple sclerosis: Predictive value of early brain and lesion volume changes

Rozemarijn M Mattiesing , Eline Kramer , Eva MM Strijbis , Iman Brouwer , Ronald A van Schijndel , Giordano Gentile , Marco Battaglini, Nicola De Stefano , Bernard MJ Uitdehaag, Frederik Barkhof , Hugo Vrenken\*  and Menno M Schoonheim\* 

## Abstract

**Background:** Whether the degree of inflammation (and its resolution) and neurodegeneration after treatment initiation predicts disease progression in multiple sclerosis (MS) remains unclear.

**Objectives:** To assess the predictive value of magnetic resonance imaging (MRI)-derived brain and lesion volume (LV) changes in years 1 and 2 of treatment for disease progression.

**Methods:** Patients receiving early interferon beta-1a treatment in REFLEX/REFLEXION ( $N=262$ ) were included. Predictive regression models included new/enlarging LV (positive activity), disappearing/shrinking LV (negative activity), and global/central atrophy during years 1 and 2.

**Results:** Faster global atrophy and/or pseudo-atrophy and positive lesion activity in years 1 and 2 related to an increased probability and faster conversion to clinically definite multiple sclerosis (CDMS). Negative lesion activity in year 1 and slower central atrophy in year 2 were predictive of confirmed disability progression (9-Hole Peg Test). Positive lesion activity in year 2 was predictive of faster global atrophy, while positive lesion activity in years 1 and 2 was predictive of faster central atrophy.

**Conclusions:** A higher degree of global atrophy and/or pseudo-atrophy in year 1 was predictive of CDMS. Positive lesion activity in any year was related to CDMS and neurodegeneration. Disability was related to negative lesion activity in year 1 and slower central atrophy in year 2.

**Keywords:** Brain atrophy, disease progression, early multiple sclerosis, prognostic factors, pseudo-atrophy, white matter lesions

Date received: 28 June 2023; revised: 18 October 2023; accepted: 20 October 2023.

## Introduction

Treatment decisions in multiple sclerosis (MS) are mostly based on the clinical disease course and radiological assessment of new/enlarging T2 lesion activity.<sup>1</sup> However, MS is also characterized by a neurodegenerative component present from the earliest phases of the disease.<sup>2</sup>

Although previous studies predicting clinical progression mostly focused on the predictive value of lesions,<sup>3</sup> recent studies pointed out the clinical importance of other pathological mechanisms such as neurodegeneration.<sup>4</sup> The severity of atrophy can predict disability in MS (independent of lesion metrics)<sup>5</sup> and

is a main candidate for future clinical implementation.<sup>6</sup> In addition to cross-sectional atrophy, methodological advances now enable longitudinal radiological quantifications of disease progression, such as dynamic changes in lesions and longitudinal global atrophy rates. Regional information, such as ventricular enlargement,<sup>7,8</sup> can also predict disability progression.

The quantification of “true” neurodegeneration is, however, difficult in settings where treatment is initiated due to pseudo-atrophy. This refers to the rapid resolution of global edema inducing brain volume loss in the first 6 to 12 months after the initiation of anti-inflammatory



Correspondence to:

**RM Mattiesing**  
MS Center Amsterdam,  
Department of Radiology  
and Nuclear Medicine, Vrije  
Universiteit Amsterdam,  
Amsterdam Neuroscience,  
Amsterdam UMC Location  
VUmc, P.O. Box 7057,  
1007 MB Amsterdam, The  
Netherlands.

[r.mattiesing@  
amsterdamumc.nl](mailto:r.mattiesing@amsterdamumc.nl)

**Rozemarijn M Mattiesing**  
**Eline Kramer**  
**Iman Brouwer**  
**Ronald A van Schijndel**  
**Hugo Vrenken**  
MS Center Amsterdam,  
Department of Radiology  
and Nuclear Medicine, Vrije  
Universiteit Amsterdam,  
Amsterdam Neuroscience,  
Amsterdam UMC Location  
VUmc, Amsterdam, The  
Netherlands

**Eva MM Strijbis**  
**Bernard MJ Uitdehaag**  
MS Center Amsterdam,  
Department of Neurology,  
Vrije Universiteit  
Amsterdam, Amsterdam  
Neuroscience, Amsterdam  
UMC Location VUmc,  
Amsterdam, The Netherlands

**Giordano Gentile**  
**Marco Battaglini**  
Department of Medicine,  
Surgery and Neuroscience,  
University of Siena, Siena,  
Italy/SIENA Imaging SRL,  
Siena, Italy

**Nicola De Stefano**  
Department of Medicine,  
Surgery and Neuroscience,  
University of Siena, Siena,  
Italy

**Frederik Barkhof**  
MS Center Amsterdam,  
Department of Radiology  
and Nuclear Medicine, Vrije  
Universiteit Amsterdam,  
Amsterdam Neuroscience,  
Amsterdam UMC Location

VUmc, Amsterdam, The Netherlands/Queen Square Institute of Neurology and Centre for Medical Image Computing, University College London, London, UK

**Menno M Schoonheim**  
MS Center Amsterdam, Department of Anatomy and Neurosciences, Vrije Universiteit Amsterdam, Amsterdam Neuroscience, Amsterdam UMC Location VUmc, Amsterdam, The Netherlands

\*These authors contributed equally to this manuscript.

treatment.<sup>9</sup> Treatment efficacy is usually assessed by defining breakthrough disease using “positive” lesion activity, that is, new/enlarging T2-hyperintense lesions, T1 gadolinium-enhancing lesions, and spinal cord lesions, all of which have predictive value for disability worsening.<sup>10</sup> Negative lesion activity (disappearing/shrinking lesions) can reflect the resolution of edema and thus pseudo-atrophy early after the start of treatment. The prognostic potential of both positive and negative lesion activity has not been investigated together. In addition, it remains unclear how to best predict disease progression using a combination of factors during the initial pseudo-atrophy period after treatment initiation and/or the stable treatment period after resolution of edema where neurodegeneration can be quantified more accurately.

In this study, we aimed to investigate which magnetic resonance imaging (MRI) volumetric changes in the first year and second year after treatment initiation are predictive of different measures of clinical progression and neurodegeneration. We used data from the REFLEX/REFLEXION studies, including patients that started early subcutaneous interferon beta-1a treatment, as we have shown the presence of pseudo-atrophy in the first year in this data set.<sup>11</sup> We now hypothesized that brain and lesion volume (LV) changes in the pseudo-atrophy period have different predictive value for disease progression compared with the period after the resolution of edema.

## Materials and methods

### Data set

REFLEXION (REbif FLEXible dosing in early MS extensION; NCT00813709) was a planned 3-year continuation of the 2-year REFLEX (NCT00404352) study, aimed at assessing the impact of subcutaneous interferon beta-1a treatment on patients with early MS throughout a prolonged period of observation. The REFLEX/REFLEXION studies lasted from 16 November 2006 until 30 August 2013.<sup>12</sup> Only those patients who started treatment immediately at the beginning of REFLEX ( $N=262$ ),<sup>11</sup> the early treatment group, were included in this study. This was chosen due to the study design, as patients in the delayed treatment group started treatment immediately after conversion to clinically definite multiple sclerosis (CDMS), resulting in lowered probabilities of progression in this subgroup over time.

### Measurements

The Expanded Disability Status Scale (EDSS), timed 25-foot walk (T25FW), 9-Hole Peg Test (9HPT), and

CDMS were assessed every 3 months during REFLEX and every 6 months during REFLEXION. MRI scans consisting of  $1 \times 1 \times 3 \text{ mm}^3$  two-dimensional (2D) dual-echo proton density (PD)-/T2-, T1-weighted, and postgadolinium T1-weighted images were acquired yearly within the follow-up period of 5 years.

### Ethical approval

The REFLEX and REFLEXION studies were undertaken in compliance with the Declaration of Helsinki and standards of Good Clinical Practice according to the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. Before initiation of the studies at each center, the relevant institutional review board or independent ethics committee reviewed and approved the study protocols, patient information leaflets, informed consent forms, and investigator brochures. All patients provided written informed consent at the screening visit of REFLEX and before enrollment to REFLEXION.

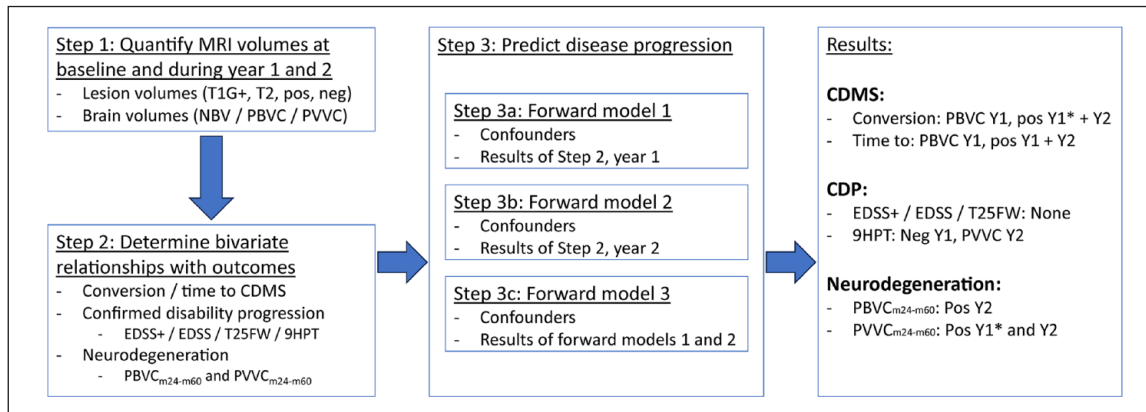
### Clinical outcome measures

**CDMS.** In REFLEX/REFLEXION,<sup>13,14</sup> conversion to CDMS was defined as a relapse or sustained progression in EDSS score of  $\geq 1.5$  points compared with baseline and confirmed at the next visit. Time to CDMS conversion was also recorded.

**Confirmed disability progression.** Confirmed disability progression (CDP) was defined by an increase of 1.5, 1, or 0.5 points on the EDSS (for baseline EDSS 0, 1–5, or  $\geq 5.5$ , respectively), or  $\geq 20\%$  worsening on T25FW or 9HPT (in seconds) at any time during the study compared with baseline, all confirmed at the next visit.<sup>15</sup> EDSS-Plus<sup>16</sup> progression was also considered, that is CDP on EDSS, T25FW, and/or 9HPT.

### Radiological predictors and outcome measures

Baseline measurements included normalized brain volume (NBV, in L, from SIENAX)<sup>17</sup>, T2-hyperintense lesion volume (T2LV, in mL), and T1 gadolinium-enhancing lesion volume (T1GdLV, in mL). Manual lesion outlines were provided by the Image Analysis Center of Amsterdam UMC (Location VUmc). A semi-automated method<sup>18</sup> quantified yearly LV changes (in  $\mu\text{L}$ ). Briefly, PD images of two visits were registered to a halfway space, after slice-to-slice intensity variation and bias-field correction. Then, after intensity matching, images were subtracted and voxel values were converted to Z-scores, based on the



**Figure 1.** Flowchart depicting the different statistical steps and main results. For more details, see “Materials and methods” section of the main manuscript. 9HPT: 9-Hole Peg Test; CDMS: clinically definite multiple sclerosis; CDP: confirmed disability progression; EDSS: Expanded Disability Status Scale; m24–m60: month 24 to month 60; PBVC: percentage brain volume change; PVVC: percentage ventricular volume change; pos: new and/or enlarging lesions; NBV: normalized brain volume; neg: disappearing and/or shrinking lesions; T1G+: T1 gadolinium-enhancing lesion volume; T2: T2-hyperintense lesion volume; T25FW: timed 25-foot walk; Y1: year 1; Y2: year 2; \*did not remain significant in Step 3c.

mean and standard deviation (SD) within a static “healthy” brain tissue mask. By applying a Z-score threshold inside manual T2 lesion masks (used as the region of interest), the volumes of changing voxels were quantified in four categories: new, enlarging, disappearing, or shrinking; these are reported below as “New LV in year 1,” and so on.

As described in the work by Mattiesing *et al.*<sup>11</sup> the yearly percentage brain volume change (PBVC) as a measure of global atrophy and percentage ventricular volume change (PVVC) as a measure of central atrophy were calculated using SIENA<sup>17</sup> and VIENA,<sup>19</sup> respectively. A more negative value for PBVC and more positive value for PVVC are indicative of faster atrophy. Neurodegeneration as an outcome measure was based on PBVC and PVVC in the last 3 years of the study, that is, from month 24 to month 60 (PBVC<sub>m24-m60</sub> and PVVC<sub>m24-m60</sub>, respectively), calculated with formulas provided in the supplemental material.

### Statistical analyses

Statistical analyses were performed with IBM SPSS-28. Stepwise logistic and linear regression were used (manual forward selection, with  $p < 0.05$ ) to identify early imaging features predictive of clinical progression (conversion to CDMS, CDP based on EDSS-Plus, EDSS, T25FW, and 9HPT) and neurodegeneration (PBVC<sub>m24-m60</sub> and PVVC<sub>m24-m60</sub>), respectively; Cox regression was used for time to CDMS as clinical outcome measure.

See Figure 1 for a flowchart depicting subsequent analyses. LV changes and PBVC/PVVC in year 1 and year 2 were first assessed in separate year 1 and year 2 MRI change models as potential individual predictors. Significant predictors from these models were combined in a final model using the “Enter” method. Variables of noninterest were included as confounders for all three models; that is, age, sex, baseline value of the specific clinical outcome measure (EDSS, 9HPT, or T25FW in seconds), and baseline MRI measures (NBV, T2LV, and T1GdLV). PBVC and PVVC year 1 and year 2 were not used as predictors for PVVC<sub>m24-m60</sub> and PBVC<sub>m24-m60</sub>, respectively.

The following models with potential predictors for all clinical and radiological outcome measures were analyzed:

Model 1. Year 1 MRI change: New LV in year 1, enlarging LV in year 1, disappearing LV in year 1, shrinking LV in year 1, PBVC year 1, and PVVC year 1.

Model 2. Year 2 MRI change: New LV in year 2, enlarging LV in year 2, disappearing LV in year 2, shrinking LV in year 2, PBVC year 2, and PVVC year 2.

Model 3. Combined MRI change: A combination of all significant predictors of models 1 and 2.

All models always included the following variables as confounders: Age, sex, clinical outcome-specific baseline value, baseline NBV, baseline T2LV, and baseline T1GdLV.

**Table 1.** Demographic, clinical, and radiological variables.

Variables	Baseline		
Age (years), mean (SD)	31.676 (8.430)		
Sex (female), <i>N</i> (%)	162 (61.832)		
EDSS, median (min, max)	1.5 (0.0, 4.0)		
T25FW (seconds), mean (SD)	6.484 (5.474)		
9HPT (seconds), mean (SD)	20.616 (11.000)		
T1GdLV (mL), median (min, max)	0.000 (0.000, 2.953) <sup>a</sup>		
T2LV (mL), median (min, max)	2.139 (0.020, 25.318)		
NBV (L), mean (SD)	1.531 (0.070)		
Variables	Year 1	Year 2	
PBVC (%), mean (SD)	-0.541 (0.722)	-0.353 (0.603)	
PVVC (%), mean (SD)	6.868 (7.095)	2.373 (4.582)	
NewLV (μL), median (min, max)	19.067 (0.000, 6259.377)	0.000 (0.000, 5430.227)	
EnLV (μL), median (min, max)	22.889 (0.000, 4091.259)	21.458 (0.000, 6784.006)	
DisLV (μL), median (min, max)	0.000 (0.000, 703.812)	0.000 (0.000, 151.634)	
ShrinkLV (μL), median (min, max)	165.939 (0.000, 10,703.097)	32.902 (0.000, 3298.756)	
<p>9HPT: 9-Hole Peg Test; DisLV: disappearing lesion volume; EDSS: Expanded Disability Status Scale; EnLV: enlarging lesion volume; NBV: normalized brain volume; NewLV: new lesion volume; PBVC: percentage brain volume change; PVVC: percentage ventricular volume change; SD: standard deviation; ShrinkLV: shrinking lesion volume; T1GdLV: T1 gadolinium-enhancing lesion volume; T25FW: timed 25-foot walk; T2LV: T2-hyperintense lesion volume.</p> <p><sup>a</sup>At baseline, 58.397% had no T1 gadolinium-enhancing lesions.</p>			

### Data availability statement

Any requests for data by qualified scientific and medical researchers for legitimate research purposes will be subject to the Data Sharing Policy of the health care business of Merck KGaA, Darmstadt, Germany. All requests should be submitted in writing to the data sharing portal for the health care business of Merck KGaA, Darmstadt, Germany <https://www.emdgroup.com/en/research/our-approach-to-research-and-development/healthcare/clinical-trials/commitment-responsible-data-sharing.html>. When the health care business of Merck KGaA 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, the health care business of Merck KGaA will endeavor to gain agreement to share data in response to requests.

### Results

Table 1 describes the demographics of the cohort (e.g. age (*mean ± SD*) = 31.7 ± 8.4 years, % female = 61.8, and median EDSS = 1.5). The supplemental material shows the individual predictive value of baseline MRI confounders and the bivariate relationships for all outcome measures on which the order of the stepwise forward procedures was based.

### Clinical outcome: CDMS

A total of 105 (40.076%) of 262 patients converted to CDMS during the study period. Results of the stepwise forward logistic regressions for the models containing the year 1 MRI change (model 1) and year 2 MRI change (model 2) are shown in Table 2. The final model (model 3) with the combined significant MRI change predictors in year 1 (PBVC and new LV) and year 2 (enlarging LV) showed that PBVC year 1 (odds ratio (OR) = 0.568, 95% confidence interval (CI) = [0.368, 0.876], *p* = 0.010) and enlarging LV in year 2 (OR = 1.001, 95% CI = [1.000, 1.002], *p* = 0.007) remained significantly associated with conversion to CDMS, while new LV in year 1 (OR = 1.0008, 95% CI = [0.9998, 1.0018], *p* = 0.116) did not.

In patients who converted to CDMS during the study period, the mean number of days until conversion was 767.371 ± 486.517. Results for models 1 and 2 are shown in Table 3. The final model showed that new LV in year 1 (hazard ratio (HR) = 1.0007, 95% CI = [1.0003, 1.0011], *p* < 0.001), PBVC year 1 (HR = 0.629, 95% CI = [0.466, 0.849], *p* = 0.002), and new LV in year 2 (HR = 1.0004, 95% CI = [1.0001, 1.0007], *p* = 0.003) remained significantly associated with the time to CDMS, while enlarging LV in year 2 (HR = 1.0001, 95% CI = [0.9999, 1.0003], *p* = 0.276) did not.

**Table 2.** Stepwise forward logistic (clinical outcomes) and linear (neurodegenerative outcomes) regression analyses for the year 1 and year 2 MRI change models

Conversion to CDMS						
Model	Nagelkerke $R^2$	$\chi^2$	$p$	Model	Nagelkerke $R^2$	$\chi^2$
Year 1 MRI change	0.104	18.423	<b>0.010*</b>	Year 2 MRI change	0.099	18.701
Predictors year 1 MRI change	OR [95% CI]	Wald	$p$	Predictors year 2 MRI change	OR [95% CI]	Wald
Age	0.968 [0.933, 1.005]	2.903	0.088	Age	0.974 [0.941, 1.009]	2.094
Sex (female)	0.740 [0.416, 1.315]	1.055	0.304	Sex (female)	0.743 [0.432, 1.278]	1.153
T1GdLV-BL (mL)	0.788 [0.347, 1.789]	0.324	0.570	T1GdLV-BL (mL)	0.825 [0.379, 1.795]	0.235
T2LV-BL (mL)	1.008 [0.930, 1.093]	0.040	0.841	T2LV-BL (mL)	0.990 [0.913, 1.074]	0.056
NBV-BL (L)	0.308 [0.003, 27.288]	0.265	0.607	NBV-BL (L)	0.156 [0.002, 11.494]	0.717
PBVC-Y1 (%)	0.554 [0.365, 0.841]	7.689	<b>0.006*</b>	EnLV-Y2 ( $\mu$ L)	1.001 [1.000, 1.002]	8.177
NewLV-Y1 ( $\mu$ L)	1.001 [1.000, 1.002]	5.023	<b>0.025*</b>			
CDP: 9HPT						
Model	Nagelkerke $R^2$	$\chi^2$	$p$	Model	Nagelkerke $R^2$	$\chi^2$
Year 1 MRI change	0.118	14.734	<b>0.040*</b>	Year 2 MRI change	0.118	14.703
Predictors year 1 MRI change	OR [95% CI]	Wald	$p$	Predictors year 2 MRI change	OR [95% CI]	Wald
Age	0.983 [0.930, 1.039]	0.388	0.533	Age	0.989 [0.938, 1.044]	0.152
Sex (female)	0.998 [0.413, 2.413]	0.00002	0.996	Sex (female)	0.946 [0.392, 2.284]	0.015
9HPT-BL (seconds)	0.913 [0.797, 1.047]	1.694	0.193	9HPT-BL (seconds)	0.838 [0.730, 0.962]	6.288
T1GdLV-BL (mL)	0.366 [0.067, 1.989]	1.354	0.245	T1GdLV-BL (mL)	0.966 [0.312, 2.987]	0.004
T2LV-BL (mL)	0.999 [0.884, 1.128]	0.0004	0.983	T2LV-BL (mL)	1.047 [0.924, 1.187]	0.528
NBV-BL (L)	0.0611 [0.0001, 43.4595]	0.696	0.404	NBV-BL (L)	0.566 [0.001, 376.567]	0.030
DisLV-Y1 ( $\mu$ L)	1.007 [1.002, 1.012]	8.255	<b>0.004*</b>	PVVC-Y2 (%)	0.871 [0.781, 0.971]	6.216

(Continued)

**Table 2.** (Continued)

PBVC <sub>m24-m60</sub>									
Model	<i>R</i> <sup>2</sup>	<i>F</i>	<i>p</i>	Model	<i>R</i> <sup>2</sup>	<i>F</i>	<i>p</i>		
Year 1 MRI change	0.123	4.534	< <b>0.001</b> *	Year 2 MRI change	0.300	9.624	< <b>0.001</b> *		
Predictors year 1 MRI change									
	<i>B</i> [95% CI]	<i>t</i>	<i>p</i>	Predictors year 2 MRI change	<i>B</i> [95% CI]	<i>t</i>	<i>p</i>		
Age	-0.009 [-0.028, 0.009]	-1.014	0.312	Age	-0.014 [-0.030, 0.003]	-1.611	0.109		
Sex (female)	-0.298 [-0.588, -0.008]	-2.028	<b>0.044</b> *	Sex (female)	-0.280 [-0.546, -0.014]	-2.083	<b>0.039</b> *		
T1GdLV-BL (mL)	-0.040 [-0.418, 0.339]	-0.207	0.836	T1GdLV-BL (mL)	0.168 [-0.198, 0.533]	0.906	0.366		
T2LV-BL (mL)	-0.052 [-0.091, -0.012]	-2.576	<b>0.011</b> *	T2LV-BL (mL)	-0.042 [-0.079, -0.005]	-2.233	<b>0.027</b> *		
NBV-BL (L)	1.873 [-0.419, 4.164]	1.614	0.109	NBV-BL (L)	1.495 [-0.586, 3.577]	1.419	0.158		
				NewLV-Y2 (μL)	-0.0004 [-0.0008, -0.0001]	-2.341	<b>0.020</b> *		
				EnLV-Y2 (μL)	-0.0005 [-0.0007, -0.0003]	-4.650	< <b>0.001</b> *		
PVVC <sub>m24-m60</sub>									
Model	<i>R</i> <sup>2</sup>	<i>F</i>	<i>p</i>	Model	<i>R</i> <sup>2</sup>	<i>F</i>	<i>p</i>		
Year 1 MRI change	0.136	4.118	< <b>0.001</b> *	Year 2 MRI change	0.399	14.875	< <b>0.001</b> *		
Predictors year 1 MRI change									
	<i>B</i> [95% CI]	<i>t</i>	<i>p</i>	Predictors year 2 MRI change	<i>B</i> [95% CI]	<i>t</i>	<i>p</i>		
Age	0.161 [-0.016, 0.337]	1.800	0.074	Age	0.210 [0.064, 0.355]	2.842	<b>0.005</b> *		
Sex (female)	1.953 [-0.810, 4.717]	1.396	0.165	Sex (female)	1.374 [-0.939, 3.688]	1.174	0.242		
T1GdLV-BL (mL)	0.349 [-3.291, 3.989]	0.189	0.850	T1GdLV-BL (mL)	-1.502 [-4.680, 1.676]	-0.934	0.352		
T2LV-BL (mL)	0.474 [0.098, 0.851]	2.488	<b>0.014</b> *	T2LV-BL (mL)	0.429 [0.106, 0.752]	2.624	<b>0.010</b> *		
NBV-BL (L)	-5.584 [-27.298, 16.130]	-0.508	0.612	NBV-BL (L)	-0.070 [-18.188, 18.048]	-0.008	0.994		
EnLV-Y1 (μL)	0.0027 [0.0001, 0.0053]	2.034	<b>0.044</b> *	NewLV-Y2 (μL)	0.006 [0.002, 0.009]	3.470	< <b>0.001</b> *		
				EnLV-Y2 (μL)	0.006 [0.004, 0.007]	6.187	< <b>0.001</b> *		

9HPT: 9-Hole Peg Test; BL: baseline; CDMS: clinically definite multiple sclerosis; CDP: confirmed disability progression; CI: confidence interval; DisLV: disappearing lesion volume; EnLV: enlarging lesion volume; m24-m60: month 24 to month 60; MRI: magnetic resonance imaging; NBV: normalized brain volume; NewLV: new lesion volume; OR: odds ratio; PBVC: percentage brain volume change; PVVC: percentage ventricular volume change; T1GdLV: T1 gadolinium-enhancing lesion volume; T2LV: T2-hyperintense lesion volume; Y1: year 1; Y2: year 2. ORs and regression coefficients are reported per unit increase for the continuous predictor. \**p* < 0.05. For significant effects the *p*-value is printed in bold.

**Table 3.** Stepwise forward Cox regression analyses with time to conversion to CDMS as outcome for the year 1 and year 2 MRI change models.

Time to CDMS							
Model		$\chi^2$	<i>p</i>	Model		$\chi^2$	<i>p</i>
Year 1 MRI change		31.899	< <b>0.001*</b>	Year 2 MRI change		20.477	<b>0.005*</b>
Predictors year 1 MRI change	HR [95% CI]	Wald	<i>p</i>	Predictors year 2 MRI change	HR [95% CI]	Wald	<i>p</i>
Age	0.976 [0.949, 1.003]	2.970	0.085	Age	0.984 [0.959, 1.010]	1.455	0.228
Sex (female)	0.764 [0.499, 1.171]	1.528	0.216	Sex (female)	0.757 [0.500, 1.145]	1.738	0.187
T1GdLV-BL (mL)	0.824 [0.460, 1.475]	0.425	0.515	T1GdLV-BL (mL)	0.872 [0.490, 1.551]	0.218	0.640
T2LV-BL (mL)	1.013 [0.952, 1.077]	0.157	0.692	T2LV-BL (mL)	1.005 [0.944, 1.070]	0.029	0.865
NBV-BL (L)	0.395 [0.012, 13.530]	0.266	0.606	NBV-BL (L)	0.270 [0.009, 8.328]	0.560	0.454
PBVC-Y1 (%)	0.618 [0.461, 0.828]	10.367	<b>0.001*</b>	NewLV-Y2 ( $\mu$ L)	1.0003 [1.0000, 1.0006]	4.298	<b>0.038*</b>
NewLV-Y1 ( $\mu$ L)	1.0007 [1.0003, 1.0010]	14.133	< <b>0.001*</b>	EnLV-Y2 ( $\mu$ L)	1.0002 [1.0000, 1.0005]	5.002	<b>0.025*</b>

BL: baseline; CDMS: clinically definite multiple sclerosis; CI: confidence interval; EnLV: enlarging lesion volume; HR: hazard ratio; MRI: magnetic resonance imaging; NBV: normalized brain volume; NewLV: new lesion volume; PBVC: percentage brain volume change; T1GdLV: T1 gadolinium-enhancing lesion volume; T2LV: T2-hyperintense lesion volume; Y1: year 1; Y2: year 2.  
HRs are reported per unit increase for the continuous predictor.  
\**p* < 0.05. For significant effects the *p*-value is printed in bold.

### Clinical outcome: CDP

**EDSS-Plus, EDSS, and T25FW.** A total of 115 (43.893%) of 262, 48 (18.321%) of 262, and 72 (27.481%) of 262 patients showed CDP based on the EDSS-Plus, EDSS, and T25FW, respectively. For all three clinical outcome measures, models 1 and 2 were not significant (hence not included in Table 2).

**9HPT.** A total of 33 (12.595%) of 262 patients showed CDP based on the 9HPT, where models 1 and 2 were significant as shown in Table 2. The final model showed that all predictors in models 1 and 2 remained significant: disappearing LV in year 1 (OR=1.010, 95% CI=[1.003, 1.016], *p*=0.003), PVVC year 2 (OR=0.846, 95% CI=[0.753, 0.950], *p*=0.005), and baseline 9HPT (OR=0.830, 95% CI=[0.712, 0.969], *p*=0.018) were significantly associated with 9HPT-based CDP.

### Neurodegenerative outcomes: global atrophy and central atrophy

Data availability and quality<sup>11</sup> allowed the quantification of atrophy in 169 patients, that is global atrophy (PBVC<sub>m24-m60</sub>:  $-1.009 \pm 0.967\%$ ) and central atrophy (PVVC<sub>m24-m60</sub>:  $7.511 \pm 9.117\%$ ) from month 24 to month 60. Results of the stepwise forward linear regressions for models 1 and 2 are shown in Table 2. For global atrophy, only the cross-sectional model 1

was significant (see Table 2), while longitudinal MRI changes in year 1 did not provide additional information (hence, no model 3 was run). For central atrophy, longitudinal measures did provide additional information and the final model showed that new LV in year 2 (*B*=0.006, 95% CI=[0.002, 0.009], *p*<0.001), enlarging LV in year 2 (*B*=0.006, 95% CI=[0.004, 0.008], *p*<0.001), age (*B*=0.184, 95% CI=[0.035, 0.333], *p*=0.016), and baseline T2LV (*B*=0.432, 95% CI=[0.109, 0.754], *p*=0.009) were significantly associated with central atrophy, while enlarging LV in year 1 (*B*=-0.00003, 95% CI=[-0.00239, 0.00233], *p*=0.982) did not remain significant.

### Discussion

This study investigated the predictive value of early brain and LV changes on clinical disability and neurodegeneration. Faster global atrophy and/or pseudo-atrophy and higher new LV in year 1 (the pseudo-atrophy period), and higher enlarging/new LV in year 2 (after resolution of edema) were related to a more likely, and faster, conversion to CDMS. Global, lower limb, and composite (EDSS-Plus) disability were not predicted by volumetric changes in year 1 or year 2. Worsening hand function was predicted by higher disappearing LV in year 1 and slower ventricular widening in year 2. Neurodegeneration was predicted by earlier positive lesion activity, with some differences between global atrophy and central atrophy.

The predictive value of positive lesion activity for CDMS is evident as observed previously, for instance in another trial on interferon beta.<sup>20</sup> The observation of the 5-year predictive value of faster global atrophy and/or pseudo-atrophy in year 1 in our study is a novel finding. In early MS patients treated with interferon beta-1a, this could indicate that patients with a stronger immune response have more inflammation and edema. In fact, the severity of pseudo-atrophy directly after treatment initiation in these patients could represent a marker of more aggressive disease and for a higher risk of experiencing a relapse, as the majority of the converted patients in this data set reached CDMS based on the relapse criterion.<sup>14</sup> For this reason, in this study conversion to CDMS and time to CDMS conversion can be considered a reasonable proxy for the occurrence of a relapse.

For EDSS-Plus, EDSS, and T25FW, atrophy and LV changes in year 1 and year 2 were not predictive in any model. Although imaging predictors were not significant, we found that a worse clinical score at baseline was generally related to a lower probability of disease progression. Although this may sound counterintuitive, this might be explained by regression to the mean.<sup>21</sup> Alternatively, other factors could be at play, for instance that those who were physically better at baseline might have more to lose than those who started off worse and might have already experienced a deterioration, perhaps due to spinal cord pathology. The probability of progression is then higher for individuals with better baseline scores.

Progression based on hand function was predicted by higher disappearing LV in year 1 and slower central atrophy in year 2. The fact that predictors of 9HPT-based CDP were not relevant for prediction of the other clinical outcome measures might indicate that it is better to look at the separate tests, that is that ambulation and hand dexterity provide complementary information. In addition, this indicates the value of individual measures, instead of a composite (like the EDSS-Plus), since these predictors were not significant for EDSS-Plus-based CDP. Clinically, the importance of individual measures is also clear, as the overlap between 9HPT progressors and T25FW progressors was only 5.3%. The different measures of physical functioning (EDSS, T25FW, and 9HPT) might touch upon different domains and mechanisms in the brain (especially the 9HPT, as the EDSS highly relies on ambulation). Scores on the EDSS and T25FW are associated with spinal cord damage,<sup>22</sup> which was not assessed in this study, while the 9HPT also touches upon areas that are related to motor planning, information processing, and cognitive pathways

in the brain.<sup>23</sup> Disappearing lesions in year 1 might be interpreted as beneficial, but our results could indicate that a higher degree of resolution of edema could, as discussed above, be a marker of a more aggressive immune response. A recent study also found that shrinking LV was related to worse clinical outcome<sup>24</sup> but their measurement concerned atrophied LV which is distinct from the negative lesion activity as assessed in this study. The applied semi-automated method is unable to quantify the replacement of lesion tissue by cerebrospinal fluid over time. Another study that also included early MS patients showed that white matter lesion shrinking was not associated with clinical outcome based on the EDSS,<sup>25</sup> which is in accordance with our findings as we only found a predictive value of disappearing LV for 9HPT-based CDP. A systematic review suggested that T2 lesion counts might have a stronger predictive value than volumes for concurrent EDSS progression.<sup>26</sup> Effects might become blurred due to the wider variety in volumes, potentially explaining the lack of predictive value for several clinical outcomes in our study. The prognostic value of slower ventricular widening in year 2 for worsening hand function seems counterintuitive, but could potentially indicate that the brain is more slowly recovering from the inflammatory processes of the first year related to global edema in progressors. Alternatively, the amount of brain volume loss in year 1 could have been driven especially by neurodegeneration leading to a floor effect in year 2 in progressors. However, we cannot exclude that this should be considered as a coincidental finding.

Higher positive lesion activity in year 1 was predictive of faster central atrophy and in year 2 of faster global and central atrophy. This is consistent with earlier work showing the predictive value of lesions for subsequent atrophy;<sup>27,28</sup> however, the novelty of this study lies in predicting atrophy rates while most studies<sup>28</sup> predicted brain volume at a specific timepoint (cross-sectional atrophy). Previous studies predicting neurodegeneration also focused on gray matter volume or cortical thickness as outcome measures,<sup>29</sup> which was not feasible in the current data set because of the limited contrast between the white matter–gray matter border in these 2D images. When combining the significant year 1 and year 2 MRI change predictors in the same (final) model, only positive lesion activity in year 2 remained predictive of central atrophy. As such, neurodegeneration seems to have been mainly driven by positive lesion activity in year 2, indicating that resolution of edema (negative lesion activity and pseudo-atrophy) might not be related to eventual neurodegenerative outcomes. This could indicate that the severity of pseudo-atrophy might be



related to a worse clinical prognosis (as discussed above), while pseudo-atrophy is not necessarily related to worse neurodegeneration.

### Limitations

Spinal cord lesions were not quantified in this study, although these have been shown to be predictive of disease progression and functioning, which could have driven our lack of predictive power for the EDSS and T25FW.<sup>22,30</sup> Gray matter atrophy could not be determined on these images, although this has been shown to be more predictive of disease progression than white matter atrophy,<sup>31</sup> and would also have allowed investigations of the compartmentalization of pseudo-atrophy in these tissue types.<sup>32</sup> The co-occurrence of pseudo- and actual atrophy in year 1 could have obscured a predictive effect of pseudo-atrophy (except for CDMS) and shrinking lesions. Individual responses to treatment (and treatment adherence) could have introduced more heterogeneity in this data set, as high efficacy therapy was not applied, where the pseudo-atrophy phenomenon might be more pronounced.<sup>33</sup> In addition, such heterogeneity might also be different at different stages and phenotypes of the disease. It would have been interesting to precisely study the relationship with relapses but this was not feasible with the available data. However, given that in REFLEX only 1 out of 130 patients converted based on EDSS progression,<sup>14</sup> conversion to CDMS and time to CDMS conversion can be considered a reasonable proxy for the occurrence of a relapse. In future studies, the predictive values of year 1 and year 2 MRI volumetric changes should be investigated with more advanced sequences such as isotropic 3D-FLAIR (three-dimensional fluid-attenuated inversion recovery).

### Conclusion

In the first year of MS, an increased degree of atrophy and/or pseudo-atrophy and disappearing lesions, reflecting resolving inflammation, seem to, respectively, have predictive value for a higher risk of/earlier time to conversion to MS and worsening hand function, while slower ventricular widening in year 2 was predictive of worsening hand function. In both the first year and second year of MS, higher enlarging and new lesion activity were predictive of a higher risk of/earlier time to conversion to MS and faster neurodegeneration across the last 3 years of the study.

### Acknowledgements

F.B. acknowledges support by the NIHR Biomedical Research Center at UCLH. This research has been executed within the MS Center Amsterdam,

Amsterdam UMC, Siena Imaging SRL, and the University of Siena.

### Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: R.M.M. has received research support from Merck. I.B. has received research support from Merck, Novartis, Teva, and the Dutch MS Research Foundation. N.D.S. is a consultant for Biogen, Merck, Novartis, Roche, Sanofi-Genzyme, and Teva; has grants or grants pending from FISM and Novartis; is on the speakers' bureaus of Biogen, Merck, Novartis, Roche, Sanofi-Genzyme, and Teva; and has received travel funds from Merck, Novartis, Roche, Sanofi-Genzyme, and Teva. B.M.J.U. reports research support and/or consultancy fees from Biogen, Genzyme, Merck, Novartis, Roche, Teva, and Immunic Therapeutics. F.B. is a steering committee or Data Safety Monitoring Board member for Biogen, Merck, ATRI/ACTC, and Prothena; is a consultant for Roche, Celltrion, Rewind Therapeutics, Merck, IXICO, Jansen, and Combinostics; is a co-founder and shareholder of Queen Square Analytics Ltd.; and has research agreements with Merck, Biogen, GE Healthcare, and Roche. H.V. 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. M.M.S. serves on the editorial board of *Neurology* and *Frontiers in Neurology*; receives research support from the Dutch MS Research Foundation, Eurostars-EUREKA, ARSEP, Amsterdam Neuroscience, MAGNIMS, and ZonMW (Vidi grant, project number 09150172010056); and has served as a consultant for or received research support from Atara Biotherapeutics, Biogen, Celgene/Bristol Meyers Squibb, EIP, Sanofi, MedDay, and Merck. E.K., E.M.M.S., R.A.v.S., G.G., and M.B. report no disclosures.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The REFLEXION study was supported by Merck (CrossRef Funder ID: 10.13039/100009945).

### ORCID iDs

Rozemarijn M Mattiesing  <https://orcid.org/0000-0002-5876-4260>

Eline Kramer  <https://orcid.org/0009-0002-4213-3138>

Eva MM Strijbis  <https://orcid.org/0000-0001-6705-5864>

Iman Brouwer  <https://orcid.org/0000-0002-0408-0461>  
 Ronald A van Schijndel  <https://orcid.org/0000-0002-2294-1050>  
 Giordano Gentile  <https://orcid.org/0000-0002-8481-2801>  
 Nicola De Stefano  <https://orcid.org/0000-0003-4930-7639>  
 Frederik Barkhof  <https://orcid.org/0000-0003-3543-3706>  
 Hugo Vrenken  <https://orcid.org/0000-0001-9106-1172>  
 Menno M Schoonheim  <https://orcid.org/0000-0002-2504-6959>

### Supplemental Material

Supplemental material for this article is available online.

### References

- Filippi M, Bar-Or A, Piehl F, et al. Multiple sclerosis. *Nat Rev Dis Primers* 2018; 4: 43.
- Charil A and Filippi M. Inflammatory demyelination and neurodegeneration in early multiple sclerosis. *J Neurol Sci* 2007; 259: 7–15.
- 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(Pt. 3): 808–817.
- Eshaghi A, Prados F, Brownlee WJ, et al. Deep gray matter volume loss drives disability worsening in multiple sclerosis. *Ann Neurol* 2018; 83(2): 210–222.
- Popescu V, Agosta F, Hulst HE, et al. Brain atrophy and lesion load predict long term disability in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2013; 84(10): 1082–1091.
- Sastre-Garriga J, Pareto D, Battaglini M, et al. MAGNIMS consensus recommendations on the use of brain and spinal cord atrophy measures in clinical practice. *Nat Rev Neurol* 2020; 16(3): 171–182.
- Dalton CM, Brex PA, Jenkins R, et al. Progressive ventricular enlargement in patients with clinically isolated syndromes is associated with the early development of multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2002; 73(2): 141–147.
- Lukas C, Minneboo A, de Groot V, et al. Early central atrophy rate predicts 5 year clinical outcome in multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2010; 81(12): 1351–1356.
- De Stefano N, Airas L, Grigoriadis N, et al. Clinical relevance of brain volume measures in multiple sclerosis. *CNS Drugs* 2014; 28(2): 147–156.
- Filippi M, Preziosa P, Langdon D, et al. Identifying progression in multiple sclerosis: New perspectives. *Ann Neurol* 2020; 88(3): 438–452.
- Mattiesing RM, Gentile G, Brouwer I, et al. The spatio-temporal relationship between white matter lesion volume changes and brain atrophy in clinically isolated syndrome and early multiple sclerosis. *Neuroimage Clin* 2022; 36: 103220.
- Comi G, De Stefano N, Freedman MS, et al. Subcutaneous interferon beta-1a in the treatment of clinically isolated syndromes: 3-year and 5-year results of the phase III dosing frequency-blind multicentre REFLEXION study. *J Neurol Neurosurg Psychiatry* 2017; 88: 285–294.
- Comi G, De Stefano N, Freedman MS, et al. Subcutaneous interferon  $\beta$ -1a in the treatment of clinically isolated syndromes: 3-year and 5-year results of the phase III dosing frequency-blind multicentre REFLEXION study. *J Neurol Neurosurg Psychiatry* 2017; 88(4): 285–294.
- 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(1): 33–41.
- Uitdehaag BMJ. Disability outcome measures in phase III clinical trials in multiple sclerosis. *CNS Drugs* 2018; 32(6): 543–558.
- Cadavid D, Cohen JA, Freedman MS, et al. The EDSS-Plus, an improved endpoint for disability progression in secondary progressive multiple sclerosis. *Mult Scler* 2017; 23(1): 94–105.
- Smith SM, Zhang Y, Jenkinson M, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. *Neuroimage* 2002; 17(1): 479–489.
- Battaglini M, Vrenken H, Tappa Brocci R, et al. 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. *Eur J Neurol* 2022; 29(7): 2024–2035.
- Vrenken H, Vos EK, van der Flier WM, et al. Validation of the automated method VIENA: An accurate, precise, and robust measure of ventricular enlargement. *Hum Brain Mapp* 2014; 35(4): 1101–1110.

20. Nielsen JM, Pohl C, Polman CH, et al. MRI characteristics are predictive for CDMS in monofocal, but not in multifocal patients with a clinically isolated syndrome. *BMC Neurol* 2009; 9: 19.
21. Vickers AJ and Altman DG. Statistics notes: Analysing controlled trials with baseline and follow up measurements. *BMJ* 2001; 323: 1123–1124.
22. Fritz NE, Keller J, Calabresi PA, et al. Quantitative measures of walking and strength provide insight into brain corticospinal tract pathology in multiple sclerosis. *Neuroimage Clin* 2017; 14: 490–498.
23. Strik M, Eijlers AJ, Dekker I, et al. Sensorimotor network dynamics predict decline in upper and lower limb function in people with multiple sclerosis. *Mult Scler* 2023; 29(1): 81–91.
24. Dwyer MG, Bergsland N, Ramasamy DP, et al. Atrophied brain lesion volume: A new imaging biomarker in multiple sclerosis. *J Neuroimaging* 2018; 28(5): 490–495.
25. 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 Behav* 2019; 9(12): e01417.
26. Fahrbach K, Huelin R, Martin AL, et al. Relating relapse and T2 lesion changes to disability progression in multiple sclerosis: A systematic literature review and regression analysis. *BMC Neurol* 2013; 13: 180.
27. Chard DT, Brex PA, Ciccarelli O, et al. The longitudinal relation between brain lesion load and atrophy in multiple sclerosis: A 14 year follow up study. *J Neurol Neurosurg Psychiatry* 2003; 74(11): 1551–1554.
28. Zivadinov R and Zorzon M. Is gadolinium enhancement predictive of the development of brain atrophy in multiple sclerosis? A review of the literature. *J Neuroimaging* 2002; 12(4): 302–309.
29. Lie IA, Weeda MM, Mattiesing RM, et al. Relationship between white matter lesions and gray matter atrophy in multiple sclerosis: A systematic review. *Neurology* 2022; 98: e1562–e1573.
30. Brownlee WJ, Altmann DR, Prados F, et al. Early imaging predictors of long-term outcomes in relapse-onset multiple sclerosis. *Brain* 2019; 142: 2276–2287.
31. Fisniku LK, Chard DT, Jackson JS, et al. Gray matter atrophy is related to long-term disability in multiple sclerosis. *Ann Neurol* 2008; 64(3): 247–254.
32. De Stefano N, Giorgio A, Gentile G, et al. Dynamics of pseudo-atrophy in RRMS reveals predominant gray matter compartmentalization. *Ann Clin Transl Neurol* 2021; 8(3): 623–630.
33. De Stefano N, Giorgio A, Battaglini M, et al. Reduced brain atrophy rates are associated with lower risk of disability progression in patients with relapsing multiple sclerosis treated with cladribine tablets. *Mult Scler* 2018; 24(2): 222–226.

Visit SAGE journals online  
[journals.sagepub.com/  
home/msj](https://journals.sagepub.com/home/msj)

 Sage journals