NeuroImage: Clinical 32 (2021) 102802

Contents lists available at ScienceDirect

NeuroImage: Clinical

journal homepage: www.elsevier.com/locate/ynicl

Brain atrophy and lesion burden are associated with disability progression in a multiple sclerosis real-world dataset using only T2-FLAIR: The NeuroSTREAM MSBase study

Michael Barnett [a,b,*,1], Niels Bergsland [c,d,1], Bianca Weinstock-Guttman [e], Helmut Butzkueven [f], Tomas Kalincik [g,h], Patricia Desmond [i], Frank Gaillard [i], Vincent van Pesch [j], Serkan Ozakbas [k], Juan Ignacio Rojas [l], Cavit Boz [m], Ayse Altintas [n], Chenyu Wang [a,b], Michael G. Dwyer [c,o,p], Suzie Yang [a], Dejan Jakimovski [c], Kain Kyle [a,b], Deepa P. Ramasamy [c], Robert Zivadinov [c,o,p]

[a] Sydney Neuroimaging Analysis Centre, Camperdown, Sydney, Australia
[b] Brain and Mind Centre, University of Sydney, Sydney, Australia
[c] Buffalo Neuroimaging Analysis Center, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, NY, USA
[d] IRCCS, Fondazione Don Carlo Gnocchi ONLUS, Italy
[e] Jacobs Comprehensive MS Treatment and Research Center, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, NY, USA
[f] Box Hill Hospital, Melbourne, Australia
[g] CORe, Department of Medicine, The University of Melbourne, Melbourne, Australia
[h] MS Centre, The Royal Melbourne Hospital, University of Melbourne, Melbourne, Australia
[i] Department of Radiology, Royal Melbourne Hospital, The University of Melbourne, Melbourne, Australia
[j] Cliniques Universitaires Saint-Luc, Brussels, UCLouvain, Belgium
[k] Dokuz Eylül University, Izmir, Turkey
[l] Hospital Italiano de Buenos Aires, Buenos Aires, Argentina
[m] KTU Medical Faculty Farabi Hospital, Trabzon, Turkey
[n] Koç University School of Medicine, Koç University Research Center for Translational Medicine (KUTTAM), İstanbul, Turkey
[o] Center for Biomedical Imaging, Clinical Translational Science Institute, USA
[p] University at Buffalo, NY, USA

ARTICLE INFO

Keywords:
Multiple sclerosis
Disability progression
Brain atrophy
Lateral ventricle volume
Lesion burden
Salient central lesion volume

ABSTRACT

Background: Methodological challenges limit the use of brain atrophy and lesion burden measures in the follow-up of multiple sclerosis (MS) patients on clinical routine datasets.
Objective: To determine the feasibility of T2-FLAIR-only measures of lateral ventricular volume (LVV) and salient central lesion volume (SCLV), as markers of disability progression (DP) in MS.
Methods: A total of 3,228 MS patients from 9 MSBase centers in 5 countries were enrolled. Of those, 2,875 (218 with clinically isolated syndrome, 2,231 with relapsing-remitting and 426 with progressive disease subtype) fulfilled inclusion and exclusion criteria. Patients were scanned on either 1.5 T or 3 T MRI scanners, and 5,750 brain scans were collected at index and on average after 42.3 months at post-index. Demographic and clinical data were collected from the MSBase registry. LVV and SCLV were measured on clinical routine T2-FLAIR images.
Results: Longitudinal LVV and SCLV analyses were successful in 96% of the scans. 57% of patients had scanner-related changes over the follow-up. After correcting for age, sex, disease duration, disability, disease-modifying therapy and LVV at index, and follow-up time, MS patients with DP (n = 671) had significantly greater absolute LVV change compared to stable (n = 1,501) or disability improved (DI, n = 248) MS patients (2.0 mL vs. 1.4 mL vs. 1.1 mL, respectively, ANCOVA $p < 0.001$, post-hoc pair-wise DP vs. Stable $p = 0.003$; and DP vs. DI, $p = 0.002$). Similar ANCOVA model was also significant for SCLV ($p = 0.03$).

* Corresponding author.
E-mail address: michael@sydneyneurology.com.au (M. Barnett).
[1] Equal contribution

https://doi.org/10.1016/j.nicl.2021.102802
Received 12 April 2021; Received in revised form 28 July 2021; Accepted 18 August 2021
Available online 24 August 2021
2213-1582/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Conclusions: LVV-based atrophy and SCLV-based lesion outcomes are feasible on clinically acquired T2-FLAIR scans in a multicenter fashion and are associated with DP over mid-term.

1. Introduction:

While MRI is used on a daily basis for diagnostic and monitoring purposes in clinical routine care, quantitative MRI measures of lesion burden and brain atrophy are not yet widely incorporated into real-world clinical routine monitoring. (De Stefano et al., 2014; Rocca et al., 2017; Wattjes et al., 2015; Zivadinov et al., 2016) In the last decade, lesion burden and brain atrophy have been assessed in all major phase 3 clinical MS trials. They are viewed as important endpoints for determining the effectiveness of disease-modifying treatment (DMT). (Carlos et al., 2015; De Stefano et al., 2014; Tsivgoulis et al., 2015a; Tsivgoulis et al., 2015b; Zivadinov et al., 2008) Therefore, there is an increasing need to translate these quantitative MRI measures into the clinical routine follow-up of MS patients to allow more informed clinical and treatment decisions. Additionally, legacy clinical routine datasets outside of specific studies or trials are often not leveraged to better understand real-world observations.

Most recently, the use of fully-automated cloud-based and proprietary software systems that receive MRI data and compute lesion burden and/or regional or global brain atrophy metrics, returning the information to the user almost in real-time, has been proposed. For example, NeuroQuant® and IcoMetrix® are Food and Drug Administration (FDA) approved and CE (European Conformity) marked fully-automated implementations that compute cross-sectional and longitudinal measures of lesion and brain volume MRI outcomes using three dimensional (3D) images of sufficiently high resolution. (Beadnall et al., 2019; Jain et al., 2015; Jain et al., 2016; Wang et al., 2016) However, these (Beadnall et al., 2019; Jain et al., 2015; Jain et al., 2016; Wang et al., 2016) and other similar research approaches (Anderson et al., 2006; Bermel and Bakshi, 2006; Miller et al., 2002; Zivadinov and Bakshi, 2004) often require pre-standardization of MRI protocols that are not common practice for many clinical centers, are not applicable across scanner-related changes, and are not easily clinically translated or applied to legacy datasets, especially in mid- to long-term retrospective studies. (Zivadinov et al., 2018a)

There is an increasing need to develop simple, accurate, reproducible, and easily obtainable lesion and brain volume measures in clinical routine. The Neurological Software Tool for REliable Atrophy Measurement (NeuroSTREAM) is a research-based, fully-automated software that computes cross-sectional and longitudinal lateral ventricular volume (LVV) and salient central lesion volume (SCLV) on low- and high-resolution 2D- and 3D-T2-fluid attenuated inversion recovery (FLAIR) images in MS patients. (Dwyer et al., 2017; Dwyer et al., 2018; Dwyer et al., 2019) While the NeuroSTREAM T2-FLAIR approach was tested on multiple scanners, field strengths, and imaging protocols in a single-center (Ghione et al., 2018; Ghione et al., 2019; Ghione et al., 2020; Jakimovski et al., 2021) and multi-center (Ghione et al., 2018; Weinstock-Guttman et al., 2018; Zivadinov et al., 2018a; Zivadinov et al., 2018b; Zivadinov et al., 2019) fashion, more evidence is needed regarding the robustness of this approach in the clinical routine datasets to predict disability progression (DP).

Therefore, the objective of this study was to investigate NeuroSTREAM-derived LVV and SCLV metrics in the multicenter MSBase study to assess the utility of NeuroSTREAM as a biomarker of DP in an independent, heterogeneous cohort of patients imaged during routine clinical practice over mid-term follow-up.

2. Methods

2.1. Design of the study

This was a multicenter, observational, retrospective, longitudinal cohort study of brain and lesion volume changes in MS patients that used data from participants in the MSBase registry (www.MSBase.org). This study was required to adhere to the Health Insurance Portability and Accountability Act (HIPAA), General Data Protection Regulation (GDPR) and Institutional Review Board (IRB)/ethics directives regarding participant privacy. It required central (University at Buffalo) and (where relevant) local IRB/ethics approvals. All subject data were anonymized via MSBase identifier codes for subjects and scans.

2.2. Population of the study:

The subject population that was included in the study fulfilled the following criteria: a) patients diagnosed with clinically isolated syndrome (CIS) or MS according to McDonald 2010 criteria, (Polman et al., 2011) b) access to MRI images that had an index and post-index T2-FLAIR sequence obtained on a 1.5 T or 3 T scanner, c) age 18–85 at index, and d) have a minimum amount of demographic and clinical data available at index (age, age at onset, sex, disease duration, and disease subtype), and e) no steroid use or relapses occurring in the 30 days prior to MRI scan. Index MRI scan was defined as a first MRI scan that was followed by a follow-up MRI scan of the same individual MS patient. Post-index MRI scan was defined as any MRI scan of an individual MS patient acquired 12 to 120 months after the index scan. If multiple post-index scans were available for the analysis, the last to follow-up scan, fulfilling inclusion criteria, was used.

2.3. Clinical outcomes

All demographic and clinical data, physical examinations, and neurologic assessments were pre-captured within the MSBase database from which data were directly exported to the investigators using standard MSBase substudy procedures. Specifically, age, age at onset, sex, race, education, disease duration, disease subtype, Expanded Disability Status Scale (EDSS), and type of DMT were collected.

DP was defined as an increase from index EDSS of at least 1.0 point, or 0.5 if the index EDSS score was >5.5. (Ghione et al., 2020) Disability improvement (DI) was defined as a reduction from the index EDSS score of at least 1.0 point if the index score was 2.0–5.5, or 0.5 if the index score was >5.5. (Ghione et al., 2020) Stable disability status was defined as non-occurrence of DP or DI. (Ghione et al., 2020) Because of a relatively small number of patients with primary-progressive (PP) MS (n = 64) in the study, the PP and secondary-progressive (SP) MS (n = 362) patients were combined, as a single progressive MS (PMS) group (n = 426).

2.4. MRI acquisition

MRI scans were retrospectively acquired at a minimum of two time points (at index and 12 to 120 months post-index) throughout clinical follow-up of MS patients, and no MRI protocol standardization or *a posteriori* selection based on MRI protocol changes was performed. The same individual patient did not need to have the scan performed on the same scanner type/field strength at the two time points. Brain MRI images were digitally transferred from participating centers via an online transfer portal and automatically de-identified by the Sydney

Neuroimaging Analysis Centre (SNAC), Sydney, Australia. 2D or 3D T2-FLAIR images were collected.

2.5. MRI analyses

The NeuroSTREAM LVV and SCLV analyses were performed by the Buffalo Neuroimaging Analysis Center (BNAC), Buffalo, NY, USA. Additionally, although LVV and SCLV were produced automatically, an imaging expert (N.B.) manually reviewed all final LVV and SCLV segmentations and rated them on a pass/fail quality control basis.

Briefly, a T2-FLAIR proxy measure for global brain atrophy, the LVV, was acquired using the previously described (NeuroSTREAM) method. (Dwyer et al., 2017; Dwyer et al., 2018) This tool performs automated processing, including basic pre-processing, multi-atlas template-based segmentation, level-set refinement, and partial volume estimation of the LVV. Pre-processing steps include reorientation, robust field of view selection, inhomogeneity correction, winsorization, sinc upsampling, anisotropic diffusion, deformation to low-, mid-, and high atrophy atlases. Then, joint-label fusion is carried out, followed by voxelwise logistic regression/masking, level set evolution, and partial volume estimation.

The T2-FLAIR proxy measure for T2-LV, the SCLV, was acquired as previously described. (Dwyer et al., 2019) This measure is comprised of a subset of lesion voxels within a specific distance (20 mm) of the lateral ventricles (centrality) and with intensity substantially brighter than normal-appearing tissue (salience). The images were bias-field corrected, and a fully-automated random-forest based lesion classifier was run to produce a lesion probability map. Then, centrality relative to the lateral ventricles was determined according to a whole-brain voxel-wise Euclidian distance map, relative to the LVV map derived using the NeuroSTREAM method described above. Lesional voxels within 20 mm were selected and then retained if they were at least one standard deviation brighter than the central normal-appearing brain tissue (cNABT) or brighter than cNABT by at least 50% of the difference between cNABT and CSF.

2.6. Statistical analyses

Demographic, clinical and MRI databases were compiled and harmonized using R version 3.6. All statistical analyses were performed with R version 3.6 and SPSS version 26.0 (IBM, Armonk, NY, USA).

The distribution of the data and their residuals was determined by visual inspection of the Q-Q plots. Normally distributed variables are shown as mean and standard deviation (SD), whereas data without normal distribution are shown with medians and interquartile range (IQR).

Comparisons for the MRI variables between the disease subtypes was performed by age, sex, disease duration at index, and time of follow-up, Bonferroni-corrected analysis of covariance (ANCOVA).

Binary logistic regression was used to ascertain the presence of DP by demographic/clinical (age, sex, disease duration, EDSS, MRI center, time of follow-up) and MRI measures. MRI measures of LVV and SCLV were investigated at both index and in terms of relative and absolute change over the follow-up. Regression metrics of beta, standard error (SE), Wald statistics are reported and significant variables are considered as DP predictors.

The differences between the disability status groups were calculated using Bonferroni-adjusted ANCOVA for age, sex, disease duration, DMT category and MRI measures at index, and time of follow-up, if longitudinal measure, in patients with all available information. The MRI outcomes in the tables are shown as estimated means after the adjustment of all covariates.

Exploratory analysis examining the effect of MRI scanner model, software or protocol changes on MRI measures between DP and non-DP MS patients was performed using Student's *t*-test and Bonferroni-adjusted ANCOVA for age, sex, disease duration, DMT category and MRI measures at index, and time of follow-up, if longitudinal measure, in patients with all available information. Additional post-hoc analyses determined the effect of scanner field strength change (four longitudinal scanner combinations) and determined the differences in demographic, clinical and MRI outcomes between all 9 MRI centers. Lastly, an additional analysis utilized time-restricted subset of patients (follow-up time of 48 months ± 3 months) and determined the difference in MRI outcomes between the disability status outcomes (disability progression, disability improvement and stable disease).

P-values lower than 0.05 were considered statistically significant.

3. Results

3.1. Demographic and clinical characteristics of the study cohort

Of the 3,228 patients recruited in the study, 2,875 were enrolled according to the inclusion and exclusion criteria (Fig. 1). Of those, 218 had CIS, 2,231 relapsing-remitting (RR) MS and 426 PMS disease subtype.

Table 1 shows demographic and clinical characteristics of the study population, according to disease subtype. The average time of follow-up was 42.3 (SD 25.3) months. As expected, PMS and RRMS patients had higher age, disease duration and EDSS at index, and greater follow-up time in months compared to CIS patients. PMS patients had a lower number of relapses compared to RRMS and CIS patients 12 and 24 months prior to index. Over the follow-up, the highest EDSS absolute score change was observed in PMS patients, followed by the RRMS and CIS patients. As expected, more PMS patients (41%) showed DP over the follow-up, compared to RRMS (26%) and CIS (16%) patients. About half (51%) of CIS patients converted to a RRMS disease subtype over the follow-up, and 3.8% of CIS/RRMS patients converted to PMS disease subtype. 323 (13%) of patients were not treated at index and the figure was lower at the follow-up (n = 216, 7.5%). More than 50% of the patients remained on the same DMT over the follow-up, while 727 (25.3%) switched DMT, 149 (5.8%) started DMT and 130 (4.5%) stopped DMT.

3.2. Participating center scanner and T2-FLAIR sequence characteristics of the study cohort

A total of 9 centers from 5 countries participated in the study. Supplement Table 1 displays the scanner and T2-FLAIR sequence characteristics of the study cohort, according to the study center. There were differences between the centers in scanner field strength combinations, scanner types at index and post-index, scanner manufacturer at index and post-index, T2-FLAIR type at index and post-index, slice thickness at index and post-index, and T2-FLAIR type and slice thickness changes over the follow-up.

Patient demographics, clinical characteristics and MRI outcomes for each separate MRI center are shown in Supplement Table 2. Due to significant differences in age, time of follow-up, disease duration and particularly in the distribution of clinical phenotypes (CIS/RRMS/PMS), no further MRI comparisons between individual MRI centers were performed.

3.3. Scanner model, software and protocol changes

At the post-index, 1,640 (57%) of patients were acquired on scanners that used a different model, software or protocol compared to index. Most patients were scanned at index on General Electric (n = 1,937, 67.4%) scanners, followed by Siemens (n = 723, 25.1%) and Philips (n = 215, 7.5%) scanners (Supplement Table). 996 (34.6%) of patients had scanner field strength change over the follow-up. Most of the patients were scanned at index and post-index on 3 T-3 T (n = 1,135, 39.5%), followed by 1.5 T-3 T (n = 755, 26.3%), 1.5 T-1.5 T (n = 744, 25.9%) and 3 T-1.5 T (n = 241, 8.4%) combination pairs.

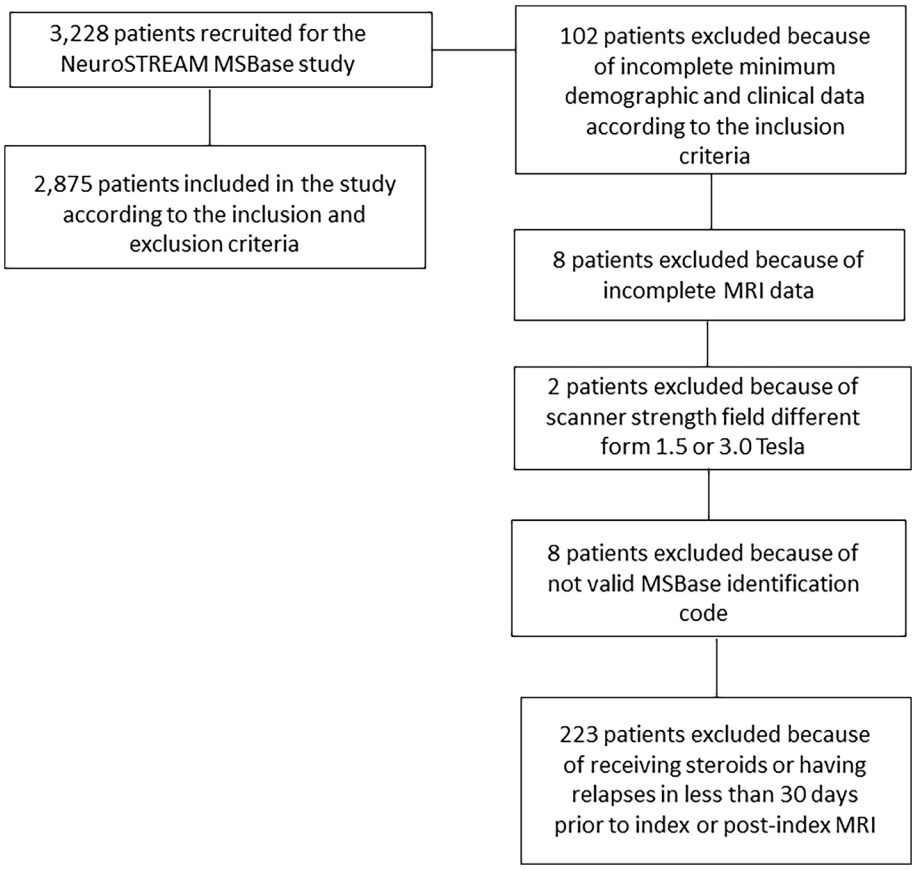


Fig. 1. Flow chart of the study population, according to the total recruitment, inclusion and exclusion criteria.

3.4. Feasibility of the MRI outcomes

In total, LVV analysis passed quality control, as previously reported, (Dwyer et al., 2017; Dwyer et al., 2019) in 2,843 (98.9%) of the index and 2,825 (98.3%) of the post-index scans, while the absolute and % LVV changes were available in 2,759 (96%) of the scan pairs.

In total, SCLV analysis passed quality control, as previously reported, (Dwyer et al., 2017; Dwyer et al., 2019) in 2,724 (94.7%) of the index and 2,728 (94.9%) of the post-index scans, while the absolute SCLV change was available in 2,749 (95.6%) of the scan pairs.

3.5. MRI outcomes by disease subtype

Table 2 shows LVV and SCLV outcomes at index and post-index, and over the follow-up, according to disease subtype. Compared to RRMS, CIS patients had significantly lower LVV at post-index ($p = 0.02$), SCLV at index ($p = 0.028$) and SCLV at post-index ($p = 0.001$), even when accounted for differences in age, sex and disease duration. Compared to RRMS, PMS patients had significantly higher LVV at index and post-index ($p < 0.001$), SCLV at index and at post-index ($p < 0.001$) and absolute SCLV change over the follow-up ($p < 0.001$).

3.6. Binary logistic regression models correlating with disability progression status

In the LVV binary logistic model (Table 3), presence of DP was significantly correlated with age at index (Beta = 0.021, Wald = 15.329, $p < 0.001$), time of follow-up in months (Beta = 0.014, Wald = 55.708, $p < 0.001$) and absolute change in LVV (Beta = 0.036, Wald = 7.7, p-value = 0.006). MRI center was not significantly associated with the presence of DP. Therefore, for every increase in age year, month of follow-up and mL of LVV change, there was a higher 2.1%, 1.4% and 3.6% chance of having DP status.

In the equivalent SCLV binary logistic model (Table 3), the presence of DP was significantly correlated with age at index (Beta = 0.025, Wald = 20.67, $p < 0.001$), time of follow-up in months (Beta = 0.015, Wald = 59.86, $p < 0.001$) and SCLV at index (B = 0.011, Wald = 7.752, $p = 0.005$). Concordant with the LVV model, MRI center was not associated with DP status. In addition to the age and time of follow-up effects, for every mL of SCLV at baseline, there was 1.1% greater chance of being diagnosed with DP status over the follow-up.

3.7. MRI outcomes according to disability status at post-index

Table 4 shows LVV and SCLV outcomes at index and post-index, and over the follow-up, according to the disability status at post-index. After adjustment for age, sex, disease duration, EDSS score, DMT category and LVV at index, and differences in follow-up time, MS patients with DP had greater absolute LVV change when compared to MS patients with stable or DI status (2.0 mL, 95% CI 1.7–2.3 vs. 1.4 mL, 95% CI 1.2–2.6 vs. 1.1 mL, 95% CI 0.6–1.6, respectively, ANCOVA $p < 0.001$, post-hoc pairwise DP vs. Stable $p = 0.003$; and DP vs. DI, $p = 0.002$). Similarly, after adjustment for all variables at index, and differences in follow-up time, MS patients with DP had greater absolute SCLV change when compared to MS patients with stable or DI status (2.2 mL, 95% CI 1.7–2.7 vs. 1.5 mL, 95% CI 1.2–1.9 vs. 0.9 mL, 95% CI 0.1–1.8, respectively, ANCOVA $p = 0.03$).

Additional analysis (Supplement Table 3) compared the MRI-based outcomes in a subset of patients that were followed for 48 months (±3 months). MS patients with DP status showed significant LVV expansion when compared to the MS patients with DI (9.75% 95 CI 4.37–15.12 vs. −4.08% 95% CI −13.72–5.56, ANCOVA $p = 0.046$).

Table 1
Demographic and clinical characteristics of the study population, according to the disease subtype, when patients who received steroids or had relapse in the 30 days prior to MRI examination were excluded.

Characteristic	CIS n = 218[1]	RR n = 2,231[1]	PMS n = 426[1]
Sex			
Female	170 (78%)	1,685 (76%)	301 (71%)
Male	48 (22%)	546 (24%)	125 (29%)
Race			
White	165 (91%)	1,634 (92%)	363 (94%)
Hispanic or Latino	1 (0.6%)	8 (0.4%)	2 (0.5%)
Asian	2 (1.1%)	13 (0.7%)	2 (0.5%)
Black or African American	12 (6.6%)	128 (7.2%)	20 (5.2%)
American Indian or Alaska Native	0 (0%)	1 (<0.1%)	1 (0.3%)
Native Hawaiian or Other Pacific Islander	1 (0.6%)	1 (<0.1%)	0 (0%)
Unknown	37	446	38
Education			
Completed college	31 (34%)	482 (39%)	79 (27%)
Completed high school	28 (31%)	354 (29%)	79 (27%)
Completed elementary school	10 (11%)	61 (5.0%)	21 (7.3%)
Completed graduate school	3 (3.3%)	76 (6.2%)	27 (9.3%)
Some college	17 (19%)	213 (17%)	72 (25%)
Some graduate school	2 (2.2%)	36 (2.9%)	11 (3.8%)
Unknown	127	1,009	137
Age at index (yrs)	37.9 (10.6)	42.5 (10.6)	53.1 (9.7)
Age at onset of the first clinical event (yrs)	35.8 (10.5)	32.3 (9.9)	34.3 (10.5)
Time of follow-up (months)	37.1 (23.4)	43.5 (25.9)	45.1 (25.7)
Disease duration at index (yrs)	2.4 (4.0)	10.3 (8.3)	18.7 (10.1)
EDSS at index	1.5 (1.0 2.0)	2.0 (1.5 3.5)	6.0 (4.5 6.5)
Unknown	32	219	42
Number of relapses in 12 months before index	0.3 (0.5)	0.3 (0.6)	0.1 (0.4)
Number of relapses in 24 months before index	0.4 (0.5)	0.4 (0.7)	0.2 (0.5)
DMT at index			
Interferon beta	47 (25%)	778 (38%)	134 (39%)
Non-therapy	81 (43%)	229 (11%)	63 (18%)
Glatiramer Acetate	31 (16%)	289 (14%)	70 (20%)
Oral therapy	15 (7.9%)	424 (20%)	36 (10%)
Natalizumab	7 (3.7%)	292 (14%)	22 (6.3%)
Alemtuzumab	1 (0.5%)	17 (0.8%)	0 (0%)
B-cell therapy	0 (0%)	3 (0.1%)	2 (0.6%)
Immunosuppressive therapy	6 (3.2%)	19 (0.9%)	18 (5.2%)
Stem cell transplantation	0 (0%)	1 (<0.1%)	0 (0%)
Other	2 (1.1%)	9 (0.4%)	0 (0%)
Combination therapy	0 (0%)	10 (0.5%)	3 (0.9%)
Unknown	28	160	78
EDSS at post-index	1.5 (1.0 2.0)	2.5 (1.5 3.5)	6.5 (6.0 7.0)
Unknown	22	175	36
EDSS absolute change over follow-up	0.0 (0.9)	0.3 (1.2)	0.4 (1.1)
Unknown	40	346	69
Number of relapses over follow-up	0.2 (0.5)	0.4 (0.9)	0.2 (0.6)
Unknown	8	46	4
Disability status at post-index			
Improved	19 (11%)	194 (10%)	35 (9.8%)
Stable	130 (73%)	1,194 (63%)	177 (50%)
Progressed	29 (16%)	497 (26%)	145 (41%)
Unknown	40	346	69
Disease subtype status at post-index			
Converted to PMS	1 (0.5%)	84 (3.8%)	0 (0%)
Converted to RRMS	111 (51%)	0 (0%)	0 (0%)
Did not convert	106 (49%)	2,147 (96%)	426 (100%)
DMT at post-index			
Interferon beta	52 (26%)	533 (25%)	89 (24%)
Non-therapy	61 (30%)	233 (11%)	75 (20%)
Glatiramer Acetate	33 (16%)	271 (13%)	86 (23%)
Oral therapy	29 (14%)	557 (26%)	48 (13%)
Natalizumab	14 (7.0%)	381 (18%)	26 (7.1%)
Alemtuzumab	3 (1.5%)	39 (1.8%)	1 (0.3%)
B-cell therapy	1 (0.5%)	20 (0.9%)	8 (2.2%)
Immunosuppressive therapy	2 (1.0%)	35 (1.6%)	14 (3.8%)
Stem cell transplantation	1 (0.5%)	4 (0.2%)	1 (0.3%)
Other	3 (1.5%)	34 (1.6%)	14 (3.8%)
Combination therapy	1 (0.5%)	15 (0.7%)	5 (1.4%)
Unknown	18	109	59
DMT status at post-index			
Remained on same DMT	75 (40%)	1,116 (55%)	161 (47%)
Started DMT	28 (15%)	102 (5.0%)	19 (5.5%)
Non Therapy	51 (27%)	122 (6.0%)	43 (13%)
Switched DMT	26 (14%)	603 (29%)	98 (29%)
Stopped DMT	6 (3.2%)	102 (5.0%)	22 (6.4%)
Unknown	32	186	83

Legend: multiple sclerosis; CIS-clinically isolated syndrome; yrs-years; RR-relapsing-remitting; PMS-primary-progressive; n-number; EDSS-Expanded Disability Status Scale; DMT-disease modifying treatment.
[1] Statistics presented: n (%); mean (SD); median (25% 75%).

Table 2
MRI outcomes, according to the disease subtype.

MRI-based measures	CIS N = 218	RRMS N = 2,231	PMS N = 426	ANCOVA p-value	CIS vs. RRMS p-value	RRMS vs PMS p-value
LVV at index	15.8 (8.4)	20.7 (11.9)	30.1 (17.4)	**<0.001**[a]	0.069 [a]	**<0.001** [a]
Failed analysis	4	27	1			
LVV at post-index	16.5 (8.6)	22.3 (13.3)	32.5 (19.2)	**<0.001** [a]	**0.02** [a]	**<0.001** [a]
Failed analysis	8	37	5			
Absolute LVV change	0.7 (3.1)	1.6 (3.6)	2.3 (4.5)	**0.008**[b]	0.05[b]	0.174[b]
Failed analysis	8	53	5			
Percent LVV change	5.7 (19.9)	7.8 (18.4)	7.9 (17.6)	0.219[b]	0.245[b]	1.0[b]
Failed analysis	8	53	5			
SCLV at index	3.7 (5.9)	9.3 (12.1)	19.9 (19.4)	**<0.001** [a]	**0.028** [a]	**<0.001** [a]
Unknown	14	99	38			
SCLV at post-index	4.0 (5.3)	10.9 (13.8)	22.7 (21.3)	**<0.001** [a]	**0.001** [a]	**<0.001** [a]
Failed analysis	11	98	38			
Absolute SCLV change	0.2 (4.3)	1.5 (5.9)	3.2 (8.7)	**<0.001**[b]	0.092[b]	**<0.001**[b]
Failed analysis	12	93	21			

Legend: LVV – lateral ventricular volume, SCLV – salient central lesion volume, CIS-clinically isolated syndrome; RR-relapsing-remitting multiple sclerosis; PMS-progressive MS.
[a]Bonferroni-adjusted analysis of Covariance (ANCOVA) corrected for age, sex, and disease duration, [b] Bonferroni-adjusted ANCOVA corrected for age, sex, disease duration at index and time of follow-up.
p-values lower or equal than 0.05 were considered statistically significant and shown in bold. All measures are shown as mean (standard deviation). Both LVV and SCLV are shown in milliliters (mL).

Table 3
Logistic regression determining the ability of lateral ventricular volume and salient central lesion volume in ascertaining disability progression in the study population.

Variables in the Equation	B	S.E.	Wald	Sig.	Exp(B)
LVV analysis					
Age at index	0.021	0.005	15.329	**<0.001**	1.021
Sex	−0.202	0.115	3.054	0.081	0.817
Disease duration at index	0.011	0.006	3.101	0.078	1.011
EDSS at index	0.003	0.027	0.014	0.906	1.003
Time of follow-up	0.014	0.002	55.708	**<0.001**	1.014
MRI Center	0.018	0.02	0.858	0.354	1.018
LVV at index	0.004	0.004	1.315	0.251	1.005
Absolute LVV change	0.036	0.013	7.7	**0.006**	1.037
SCLV analysis					
Age at index	0.025	0.006	20.67	**<0.001**	1.026
Sex	−0.153	0.118	1.685	0.194	0.858
Disease duration at index	0.011	0.006	2.851	0.091	1.011
EDSS at index	−0.015	0.029	0.274	0.601	0.985
Time of follow-up	0.015	0.002	59.86	**<0.001**	1.015
MRI Center	0.008	0.02	0.159	0.69	1.008
SCLV at index	0.011	0.004	7.752	**0.005**	1.012
Absolute SCLV change	0.013	0.008	2.683	0.101	1.013

Legend: EDSS – Expanded Disability Status Scale, LVV – lateral ventricular volume; SCLV – salient central lesion volume.
Binary logistic regression was used to determine the correlation of DP with demographic (age, sex, disease duration, EDSS, time of follow-up, MRI center) and MRI measures (baseline LVV and absolute LVV change or baseline SCLV and absolute SCLV change, respectively).
Regression metrics of beta, standard error (SE), Wald statistics are reported.
p-values lower or equal than 0.05 were considered statistically significant and shown in bold.

3.8. MRI differences between non-DP and DP MS patients, according to the MRI scanner model, software or protocol changes

Compared to non-DP MS patients, DP patients were older (46.5 vs. 42.9 years), had longer disease duration (12.7 vs. 10.3 years), were followed for a longer period of time (51.8 vs. 40.5 months) and had higher EDSS (2.5 vs. 2.0) at index ($p < 0.001$ for all). The differences between DP and non-DP MS patients in those variables were similar, independent of whether patients were acquired on scanners that used a different model, software or protocol at follow-up ($n = 1,371$) or had no scanner-related changes ($n = 1,049$).

Table 5 shows the differences in LVV and SCLV at index and over the follow-up in DP and non-DP MS patients, according to the MRI scanner model, software or protocol changes. Segmentation examples are shown in Fig. 2. Despite changes in MRI scanner model, software or protocol, DP MS patients had significantly greater absolute LVV change (2.4 mL vs. 1.2 mL, Student's *t*-test, $p < 0.001$; ANCOVA, $p = 0.001$), % LVV change (9.9% vs. 5.6%, Student's *t*-test, $p < 0.001$; ANCOVA, $p = 0.014$) and absolute SCLV change (2.8 mL vs. 1.5 mL, Student's *t*-test, $p = 0.006$; ANCOVA, $p = 0.108$), compared to non-DP patients. DP MS patients who had no scanner-related changes had greater absolute LVV change (2.5 mL vs. 1.5 mL, Student's *t*-test, $p < 0.001$; ANCOVA, $p = 0.08$), % LVV change (10.9% vs. 8.3%, Student's *t*-test, $p = 0.033$; ANCOVA, $p = 0.999$) and absolute SCLV change (2.5 mL vs. 1.1 mL, Student's *t*-test, $p = 0.002$; ANCOVA, $p = 0.046$) compared to non-DP patients.

The differences in LVV and SCLV at index and over the follow-up in DP and non-DP MS patients based on the four MRI scanner field strength combinations are shown in the Supplement Table 4. In the group of MS patients within the 1.5 T-1.5 T scanner combination, there were significantly greater absolute and % LVV change differences between the stable DP and DI MS patients ($p = 0.005$ for both). Within the 1.5 T-3.0 T and 3.0 T-1.5 T scanner combination, the disability status groups also had significantly different absolute LVV change over the follow-up ($p = 0.018$ and $p = 0.04$, respectively). Lastly, the patient outcome subgroups within the 3.0 T-3.0 T scanner combination were significantly different in index LVV and absolute change in LVV ($p = 0.009$ and $p = 0.003$, respectively).

4. Discussion

This study investigated lesion burden accumulation and brain atrophy progression in a large cohort of CIS and MS patients enrolled in MSBase and followed in clinical routine, with respect to their DP status at follow-up. The study utilized brain volume and lesion measures from retrospective MRI exams, collected over approximately 10 years from more than 3,000 individuals who were followed for an average of 3.5 years. We showed that LVV-based atrophy and SCLV-based lesion outcomes are feasible on T2-FLAIR scans in a multicenter fashion and are associated with DP over mid-term. The association between LVV expansion and DP was further confirmed in smaller time-restricted subanalysis within patients that were followed for 48 months (±3 months). Changes in MRI scanner model, software or protocol did not interfere with the ability of LVV and SCLV to differentiate between DP and non-

Table 4
Lateral ventricle volume and salient central lesion volume outcomes in MS patients, according to disability status at post-index.

MRI measures	DI	Stable	DP	p-value	p-value	p-value
	Estimated Mean (95% CI)	Estimated Mean (95% CI)	Estimated Mean (95% CI)	(DP, DI, Stable)	(DP, DI)*	(DP, Stable)*
LVV at index	21.9 (20.4–23.5)	21.1 (20.5–21.8)	22.7 (20.4–23.5)	**0.03**	1.00	**0.027**
SCLV at index	11.4 (9.7–13.1)	9.5 (8.8–10.2)	11.7 (10.6–12.7)	**<0.001**	1.000	**0.002**
Absolute LVV change	1.1 (0.6–1.6)	1.4 (1.2–1.6)	2.0 (1.7–2.3)	**<0.001**	**0.002**	**0.003**
% LVV change	5.8 (3.4–8.2)	7.4 (6.4–8.3)	8.8 (7.4–10.3)	0.082	0.104	0.32
Absolute SCLV change	0.9 (0.1–1.8)	1.5 (1.2–1.9)	2.2 (1.7–2.7)	**0.03**	0.053	0.104

Legend: MS- multiple sclerosis; DP- disease progression; DI- disease improvement; Stable- stable disability status; LVV-lateral ventricle volume; PLVVC- percentage lateral ventricle volume change; SCLV – salient central lesion volume, CI- confidence interval; ANCOVA-analysis of covariance.
The data are presented as estimated mean (95% confidence intervals). The absolute values are expressed in milliliters.
The differences between the groups were calculated using Bonferroni-adjusted ANCOVA for: age at index, sex, disease duration, disease modifying treatment category and MRI related measure at index, and time of follow-up, if longitudinal measure, in patients with all available information
DMT was categorically classified as 0 – no treatment, 1 – moderate efficacy (interferon-beta, glatiramer acetate, teriflunomide, sphingosine-1-phosphate receptor modulators, dimethyl fumarate, intravenous immunoglobulins and azathioprine) and 2 – high efficacy (alemtuzumab, natalizumab, ocrelizumab, mitoxantrone, stem cell transplantation and B-cell therapies). Please note that the number of patients included in the ANCOVAs was reduced based on the availability of DMT.
* - Bonferroni adjusted pairwise comparison
p-values lower or equal than 0.05 were considered statistically significant and shown in bold.

Table 5
MRI differences between MS patients with and without disability progression, according to the MRI scanner model, software or protocol changes.

MRI measures	MRI scanner model, software or protocol changes Yes (n = 1,371)			MRI scanner model, software or protocol changes No (n = 1,049)		
LVV and SCLV measures	Non-DP MS(n = 987)	DP MS (n = 384)	p-value	Non-DP MS(n = 762)	DP MS (n = 287)	p-value
LVV at index	21.3 (13.5)	24.3 (14.2)	**<0.001**[a] **0.027**[b]	20.9 (12.9)	23.3 (13.0)	**0.009**[a] 0.218[b]
SCLV at index	9.7 (13.3)	13.4 (16.5)	**<0.001**[a] **0.028**[b]	9.6 (12.1)	12.9 (16.1)	**0.002**[a] **0.032**[b]
Absolute LVV change	1.2 (3.5)	2.4 (4.9)	**<0.001**[a] **0.001**[b]	1.5 (3.0)	2.5 (4.5)	**<0.001**[a] 0.08[b]
% LVV change	5.6 (18.7)	9.9 (20.8)	**<0.001**[a] **0.014**[b]	8.3 (18.6)	10.9 (16.9)	**0.033**[a] 0.999[b]
Absolute SCLV change	1.5 (6.9)	2.8 (7.6)	**0.006**[a] 0.108[b]	1.2 (4.9)	2.5 (6.3)	**0.002**[a] **0.046**[b]

Legend: MS- multiple sclerosis; DP- disability progression; LVV-lateral ventricle volume; PLVVC- percentage lateral ventricle volume change; SCLV – salient central lesion volume.
DMT was categorically classified as 0 – no treatment, 1 – moderate efficacy (interferon-b, glatiramer acetate, teriflunomide, sphingosine-1-phosphate receptor modulators, dimethyl fumarate, intravenous immunoglobulins and azathioprine) and 2 – high efficacy (alemtuzumab, natalizumab, ocrelizumab, mitoxantrone, stem cell transplantation and B-cell therapies). Please note that the number of patients included in the ANCOVAs was reduced based on the availability of DMT.
p-values lower or equal than 0.05 were considered statistically significant and shown in bold.
[a] The differences between the groups were calculated using Student's *t*-test.
[b] The differences between the groups were calculated using Bonferroni-adjusted ANCOVA for: age, sex, disease duration, disease modifying treatment (DMT) category and MRI related measure at index, and time of follow-up, if longitudinal measure, in patients with all available information.

DP MS patients.

When adjusting for demographic, clinical and MRI differences at index, and time of follow-up between the MS patients with or without DP, LVV remained significantly related to DP. SCLV was also associated with DP, and SCLV at index was predictive of the subsequent disability status change. We also demonstrated that LVV and SCLV cross-sectional and longitudinal outcomes differ between CIS, RRMS, and PMS patients. Therefore, this multicenter study confirmed results of previous studies showing that changes in brain atrophy and lesion burden are associated with development of DP over mid-term follow-up. (Fisher et al., 2008; Fisniku et al., 2008; Horakova et al., 2008; Jacobsen et al., 2014)

In this study, we chose to report only absolute SCLV changes. Given the heterogeneity of the included MS subtypes, % SCLV change would be somewhat misleading, as a small absolute SCLV change in CIS patients would produce a relatively large % SCLV change, and similarly, a large absolute SCLV change in SPMS patients would produce a relatively small % SCLV change. While a similar concern can be raised for the LVV, it is our experience that LVV is much less influenced by the biological degree of inflammatory activity in MS, and more by age. (Ghione et al., 2018; Ghione et al., 2019; Ghione et al., 2020) Nevertheless, although we observed similar findings between absolute and % LVV change with respect to disability status, absolute LVV measures showed somewhat stronger associations, supporting previous studies using the NeuroSTREAM absolute vs. % LVV change. (Dwyer et al., 2017; Dwyer et al., 2018; Ghione et al., 2018; Ghione et al., 2019; Ghione et al., 2020)

The scanner model, software, and protocol changes over the follow-up did not significantly affect the ability of LVV and SCLV to distinguish MS patients with and without DP, at least when analyzed at the group level, in line with a recent study. (Jakimovski et al., 2021) Given the large heterogeneity of scanner types and field strengths between the centers in the present study, and more than half of patients having scans with scanner-related changes over the follow-up, our findings make LVV and SCLV promising biomarkers for differentiating patients by disability status at the group level, even in the presence of scanner model, software and protocol changes.

Assessing brain atrophy in clinical routine is important for a number of reasons. Several reports have shown it to be one of the most reliable

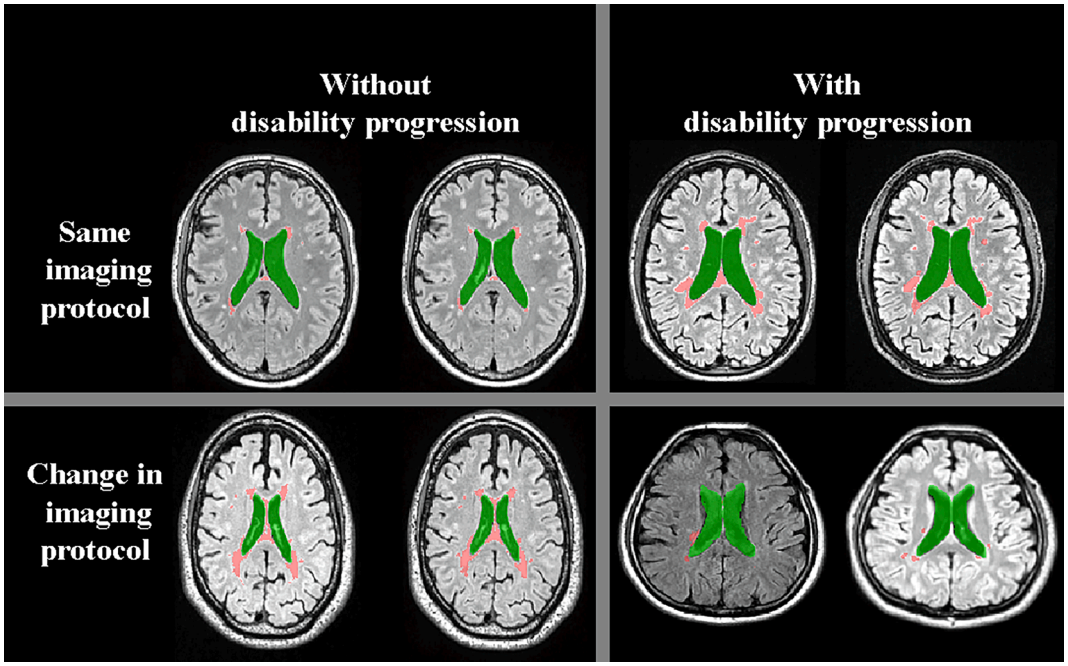


Fig. 2. The segmentation results in patients with/without disability progression and with/without scanner change over the follow-up. The lateral ventricles are shown in green while salient central lesions are shown in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

biomarkers of neurodegeneration that correlates with physical and cognitive impairment in MS patients. (Carlos et al., 2015; De Stefano et al., 2014; Horakova et al., 2008; Tsivgoulis et al., 2015a; Tsivgoulis et al., 2015b) For more than a decade, randomized-controlled trials have used brain atrophy measurement as a secondary or tertiary endpoint to determine effectiveness of treatment. As such, the application of these measures to clinical routine scans would enable the ability to drive research by providing quantitative metrics from a far broader pool of otherwise-latent scans with potentially wider demographic coverage and follow-up time. Unfortunately, though, assessing brain atrophy on clinical routine imaging can be challenging due to technical factors related to acquisition and measurement methods (Anderson et al., 2006; Bermel and Bakshi, 2006; De Stefano et al., 2014; Miller, 2002; Rocca et al., 2017; Wattjes, 2015; Zivadinov and Bakshi, 2004). In particular, these methods are affected by both acquisition stability over time and by acquisition protocol quality.

Regarding acquisition stability over time, a recent multicenter, retrospective, real-world study (MS-MRIUS) investigated the feasibility of brain atrophy measurement in a clinical routine without MRI protocol standardization, using academic and non-academic centers specializing in treatment and monitoring of MS. (Weinstock-Guttman et al., 2018; Zivadinov et al., 2018b; Zivadinov et al., 2018c; Zivadinov et al., 2019) The MS-MRIUS study showed that scanner/protocol changes occurred in more than 50% of the patients over the 16 month follow-up. (Zivadinov et al., 2018b) The results from the present NeuroSTREAM MSBase study are in line with this previous report, as over a period of 3.5 years, 57% of cases had scanner model, software, and protocol changes. Therefore, the present study confirms that frequent change of scanner field strength, model, software and protocol is inevitable in real-world clinical routine follow-up of MS patients.

Regarding acquisition protocol quality, image contrast and image resolution are important for reliable and optimal segmentation of lesion and brain volumes, and 3D pulse sequences are preferred as the gold standard, because of reduced partial voluming and more accurate co-registration, especially for serial imaging over time, compared to 2D. However, in the current NeuroSTREAM MSBase study, only 35.1% of the patients had 3D-FLAIR scan at index, compared to 64.9% with 2D-FLAIR. The figures were slightly better at post-index – 37.4% vs. 62.6% – but still less than half had 3D-FLAIR. No analysis was performed investigating performance of the NeuroSTREAM software on the 2D vs 3D T2-FLAIR in this real-word study, and that should be subject of future work.

The inclusion criteria for this study required that enrolled patients had a T2-FLAIR sequence available at index and post-index on a 1.5 T or 3 T scanner. The decision was based on the findings from a previous large feasibility multicenter study in the USA (MS-MRIUS), (Zivadinov et al., 2018a) where it was found that only T2-FLAIR sequence was a common denominator in the clinical routine scanning of patients with MS (presence of T2-FLAIR in 99.5% patients at index and 99.3% patients at post-index). Other sequences required for brain volumetry, like 2D- and 3D-weighted T1 were substantially less available (presence of 2D T1-WI in 79.7% patients at index and 75.6% patients at post-index, and of 3D T1-WI in 31.4% patients at index and 39.7% patients at post-index). While some of the participating sites made available all MRI sequences acquired at index and post-index, the other sites only provided a T2-FLAIR acquisition, as required by the original study design. Therefore, there was a limited availability to perform comparison of our T2-FLAIR based approach, to the other currently available brain volumetry methods requiring presence of 2D- and 3D-weighted T1 sequences. (Anderson et al., 2006; Ashburner and Friston, 2000; Bermel and Bakshi, 2006; Fischl and Dale, 2000; Nakamura et al., 2014; Patenaude et al., 2011; Smith et al., 2002)

While the imaging measures applied in this study do not replace MRI analyses measures obtained on research-quality MRI sequences, they may facilitate more broad use of legacy datasets for understanding predictors of clinical disability, especially in real-world treatment-related studies of MS patients. They also may contribute to increase research opportunities in clinical centers that did not obtain research quality MRI acquisitions as part of clinical routine imaging. Therefore, the current approach may be useful where only clinical T2-FLAIR is available, providing substantially more quantitative information about brain pathology and disability than is currently standard practice in MS. However, we need to be careful to avoid over-interpreting the translation of current findings to the follow-up of individual patients, as protocol stability and quality factors almost certainly have a much larger impact on individual analyses, compared to those obtained at the group-level.

In line with previous studies, (Dwyer et al., 2017; Dwyer et al., 2018) the NeuroSTREAM MSBase study showed that 4% of LVV outcomes did not pass quality control at index, post-index or over the follow-up, rendering it a potentially valuable measure to be utilized in the clinical routine from a feasibility point of view. The assessment of LVV presents some advantages for calculation of brain atrophy on clinical routine scans compared to whole brain volume measurements, as previously reported. (Dwyer et al., 2018; Zivadinov et al., 2018a). This is mainly because tissue borders of lateral ventricles have high contrast and simple topology with respect to the surrounding CSF, and the central position of the ventricles within the field of view renders them less likely to be affected by gradient distortions, inaccurate co-registration, segmentation errors, incomplete head coverage, and wrap-around artifacts. The NeuroSTREAM MSBase study further provides support for the notion that LVV measurement in T2-FLAIR sequence is a meaningful and reliable measure of brain atrophy assessment in real-word datasets, when scanning protocols cannot be standardized.

Similar figures were found for SCLV, with $<5\%$ of the SCLV analyses failing quality control at index, post-index or over the follow-up, confirming it as a robust measure to be utilized in the clinical routine. These results are in line with a recently published study. (Dwyer et al., 2019)

LVV and SCLV have been recently explored as predictors of cognitive outcome, based on the Symbol Digit Modalities Test (SDMT), in a smaller cohort of patients with MS (Fuchs et al., 2021). Moreover, additional T2-FLAIR-only based derived measures have also been developed, including medulla oblongata volume, as proxy of spinal cord atrophy, thalamic segmentation, and white matter network disruption (Fuchs et al., 2021). We plan to similarly explore the relationship of these metrics with SDMT and other focused clinical outcomes in larger, multicenter datasets in future work.

A major strength of this study is the large sample size, with data collected at 9 centers from 5 countries, comprising 16 different scanner types and 3 different scanner manufacturers. An important limitation of the study is its retrospective nature, and that the number of study participants enrolled by different centers was not balanced. Another limitation is that DP was not confirmed after 3- o 6-months of follow-up, however, we excluded patients with active relapses or recent steroid treatment. While we used information about the treatment status to adjust our analyses, it would be important to utilize this large multicenter dataset for more detailed analyses of DMT in relation to imaging outcomes, which will be subject of future work.

In conclusion, LVV-based atrophy and SCLV-based lesion outcomes are feasible on T2-FLAIR scans in a multicenter fashion and associated with DP over mid-term. Both LVV and SCLV are resistant to MRI scanner-related changes.

CRediT authorship contribution statement

Michael Barnett: Conceptualization, Writing – original draft, Writing - review & editing, Funding acquisition, Resources. **Niels Bergsland:** Software, Formal analysis, Writing - review & editing. **Bianca Weinstock-Guttman:** Resources, Writing - review & editing. **Helmut Butzkueven:** Resources, Writing - review & editing. **Tomas Kalincik:** Resources, Writing - review & editing. **Patricia Desmond:** Resources, Writing - review & editing. **Frank Gaillard:** Resources,

Writing - review & editing. **Vincent van Pesch:** Resources, Writing - review & editing. **Serkan Ozakbas:** Resources, Writing - review & editing. **Juan Ignacio Rojas:** Resources, Writing - review & editing. **Cavit Boz:** Resources, Writing - review & editing. **Ayse Altintas:** Resources, Writing - review & editing. **Chenyu Wang:** Formal analysis, Writing - review & editing. **Michael G. Dwyer:** Software, Formal analysis, Writing - review & editing. **Suzie Yang:** Data curation, Project administration, Writing - review & editing. **Dejan Jakimovski:** Resources, Writing - review & editing. **Kain Kyle:** Data curation, Writing - review & editing. **Deepa P. Ramasamy:** Data curation, Writing - review & editing. **Robert Zivadinov:** Conceptualization, Methodology, Writing - review & editing, Funding acquisition, Writing – original draft, Resources.

Study Disclosures

Research reported in this publication was funded by the Novartis.

Financial Relationships/Potential Conflicts of Interest

Michael Barnett reports research grants from Genzyme-Sanofi, Novartis, Biogen, and Merck outside the submitted work. He is a co-founder of RxMx and Research Director for the Sydney Neuroimaging Analysis Centre.

Niels Bergsland has nothing to disclose.

Bianca Weinstock-Guttman has participated in speaker's bureaus and/or served as a consultant for Biogen, Novartis, Genzyme & Sanofi, Genentech, Abbvie, Bayer AG, and Celgene/ BMS. Dr. Weinstock-Guttman also has received grant/research support from the agencies listed in the previous sentence as well as Mallinckrodt Pharmaceuticals, Inc. She serves in the editorial board for BMJ Neurology, Journal of International MS and CNS Drugs.

Helmut Butzkueven has nothing to disclose.

Tomas Kalincik served on scientific advisory boards for BMS/Celgene, Roche, Sanofi Genzyme, Novartis, Merck and Biogen, steering committee for Brain Atrophy Initiative by Sanofi Genzyme, received conference travel support and/or speaker honoraria from WebMD Global, Novartis, Biogen, Sanofi-Genzyme, Teva, BioCSL and Merck and received research or educational event support from Biogen, Novartis, Genzyme, Roche, BMS/Celgene and Merck.

Patricia Desmond has nothing to disclose.

Frank Gaillard has nothing to disclose.

Vincent van Pesch has received travel grants from Merck, Biogen, Sanofi, Celgene, Almirall and Roche. His institution has received research grants and consultancy fees from Roche, Biogen, Sanofi, Celgene, Merck, Almirall and Novartis Pharma.

Serkan Ozakbas has nothing to disclose.

Juan Ignacio Rojas has received honoraria in concept of lectures and traveling grants, research projects from Biogen, Genzyme, Roche, Novartis and Merck.

Cavit Boz has nothing to disclose.

Ayse Altintas has received travel funding or speaker honoraria from Merck, Generica and Deva, unrelated to the present study.

Chenyu Wang is an employee of the Sydney Neuroimaging Analysis Centre and has received research support from the Nerve Research Foundation and Multiple Sclerosis Research Australia. He has received institutional support from Novartis, Biogen and Roche for speaking.

Michael G. Dwyer received personal compensation from Keystone Heart for speaking and consultant fees. He received financial support for research activities from Bristol Myers Squibb, Novartis and Keystone Heart.

Suzie Yang is an employee of the Sydney Neuroimaging Analysis Centre.

Dejan Jakimovski has nothing to disclose.

Kain Kyle is an employee of the Sydney Neuroimaging Analysis Centre.

Deepa P. Ramasamy has nothing to disclose.

Robert Zivadinov received personal compensation from Bristol Myers Squibb, EMD Serono, Sanofi, Novartis and Keystone Heart for speaking and consultant fees. He received financial support for research activities from Bristol Myers Squibb, Sanofi, Novartis, Keystone Heart, V-WAVE Medical, Mapi Pharma and Protembis.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nicl.2021.102802.

References

Anderson, V.M., Fox, N.C., Miller, D.H., 2006. Magnetic resonance imaging measures of brain atrophy in multiple sclerosis. J. Magn. Reson. Imaging 23 (5), 605–618.

Ashburner, J., Friston, K.J., 2000. Voxel-Based Morphometry—The Methods. NeuroImage 11 (6), 805–821.

Beadnall, H.N., et al., 2019. Comparing longitudinal brain atrophy measurement techniques in a real-world multiple sclerosis clinical practice cohort: towards clinical integration? Ther Adv Neurol Disord. 12, 1756286418823462.

Bermel, R.A., Bakshi, R., 2006. The measurement and clinical relevance of brain atrophy in multiple sclerosis. The Lancet Neurology 5 (2), 158–170.

Carlos, R., 2015. Brain atrophy in Multiple Sclerosis. AJPN 3 (3), 40.

De Stefano, N., Airas, L., Grigoriadis, N., Mattle, H.P., O'Riordan, J., Oreja-Guevara, C., Sellebjerg, F., Stankoff, B., Walczak, A., Wiendl, H., Kieseier, B.C., 2014. Clinical Relevance of Brain Volume Measures in Multiple Sclerosis. CNS Drugs 28 (2), 147–156.

Dwyer, M.G., Silva, D., Bergsland, N., Horakova, D., Ramasamy, D., Durfee, J., Vaneckova, M., Havrdova, E., Zivadinov, R., 2017. Neurological software tool for reliable atrophy measurement (NeuroSTREAM) of the lateral ventricles on clinical-quality T2-FLAIR MRI scans in multiple sclerosis. NeuroImage: Clinical 15, 769–779.

Dwyer, M.G., Hagemeier, J., Bergsland, N., Horakova, D., Korn, J.R., Khan, N., Uher, T., Medin, J., Silva, D., Vaneckova, M., Havrdova, E.K., Zivadinov, R., 2018. Establishing pathological cut-offs for lateral ventricular volume expansion rates. NeuroImage: Clinical 18, 494–501.

Dwyer, M.G., Bergsland, N., Ramasamy, D.P., Weinstock-Guttman, B., Barnett, M.H., Wang, C., Tomic, D., Silva, D., Zivadinov, R., 2019. Salient Central Lesion Volume: A Standardized Novel Fully Automated Proxy for Brain FLAIR Lesion Volume in Multiple Sclerosis. Journal of Neuroimaging 29 (5), 615–623.

Fischl, B., Dale, A.M., 2000. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc. National Acad. Sci. 97 (20), 11050–11055.

Fisher, E., Lee, J.-C., Nakamura, K., Rudick, R.A., 2008. Gray matter atrophy in multiple sclerosis: A longitudinal study: Gray Matter Atrophy in MS. Ann Neurol. 64 (3), 255–265.

Fisniku, L.K., Chard, D.T., Jackson, J.S., Anderson, V.M., Altmann, D.R., Miszkiel, K.A., Thompson, A.J., Miller, D.H., 2008. Gray matter atrophy is related to long-term disability in multiple sclerosis: GM Atrophy and Disability in MS. Ann Neurol. 64 (3), 247–254.

Fuchs, T.A., Dwyer, M.G., Jakimovski, D., Bergsland, N., Ramasamy, D.P., Weinstock-Guttman, B., HB Benedict, R., Zivadinov, R., 2021. Quantifying disease pathology and predicting disease progression in multiple sclerosis with only clinical routine T2-FLAIR MRI. NeuroImage: Clinical 31, 102705. https://doi.org/10.1016/j.nicl.2021.102705.

Ghione, E., et al., 2018. Brain Atrophy Is Associated with Disability Progression in Patients with MS followed in a Clinical Routine. AJNR Am J Neuroradiol. 39, 2237-2242.

Ghione, E., et al., 2019. Aging and Brain Atrophy in Multiple Sclerosis. J Neuroimaging. 29, 527-535.

Ghione, E., et al., 2020. Disability Improvement Is Associated with Less Brain Atrophy Development in Multiple Sclerosis. AJNR Am J Neuroradiol.

Horakova, D., Cox, J.L., Havrdova, E., Hussein, S., Dolezal, O., Cookfair, D., Dwyer, M.G., Seidl, Z., Bergsland, N., Vaneckova, M., Zivadinov, R., 2008. Evolution of different MRI measures in patients with active relapsing-remitting multiple sclerosis over 2 and 5 years: a case-control study. J. Neurol., Neurosurg. Psychiatry 79 (4), 407–414.

Jacobsen, C., Hagemeier, J., Myhr, K.-M., Nyland, H., Lode, K., Bergsland, N., Ramasamy, D.P., Dalaker, T.O., Larsen, J.P., Farbu, E., Zivadinov, R., 2014. Brain atrophy and disability progression in multiple sclerosis patients: a 10-year follow-up study. J. Neurol. Neurosurg. Psychiatry 85 (10), 1109–1115.

Jain, S., Sima, D.M., Ribbens, A., Cambron, M., Maertens, A., Van Hecke, W., De Mey, J., Barkhof, F., Steenwijk, M.D., Daams, M., Maes, F., Van Huffel, S., Vrenken, H., Smeets, D., 2015. Automatic segmentation and volumetry of multiple sclerosis brain lesions from MR images. NeuroImage: Clinical 8, 367–375.

Jain, S., et al., 2016. Two Time Point MS Lesion Segmentation in Brain MRI: An Expectation-Maximization Framework. Front Neurosci. 10, 576.

Jakimovski, D., Zivadinov, R., Bergsland, N., Ramasamy, D.P., Hagemeier, J., Genovese, A.V., Hojnacki, D., Weinstock-Guttman, B., Dwyer, M.G., 2021. Clinical feasibility of longitudinal lateral ventricular volume measurements on T2-FLAIR across MRI scanner changes. NeuroImage: Clinical 29, 102554.

Miller, D.H., et al., 2002. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. Brain 125, 1676–1695.

Nakamura, K., Guizard, N., Fonov, V.S., Narayanan, S., Collins, D.L., Arnold, D.L., 2014. Jacobian integration method increases the statistical power to measure gray matter atrophy in multiple sclerosis. NeuroImage: Clinical 4, 10–17.
Patenaude, B., Smith, S.M., Kennedy, D.N., Jenkinson, M., 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. NeuroImage 56 (3), 907–922.
Polman, C.H., Reingold, S.C., Banwell, B., Clanet, M., Cohen, J.A., Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., Lublin, F.D., Montalban, X., O'Connor, P., Sandberg-Wollheim, M., Thompson, A.J., Waubant, E., Weinshenker, B., Wolinsky, J.S., 2011. Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. Ann Neurol. 69 (2), 292–302.
Rocca, M.A., Battaglini, M., Benedict, R.H.B., De Stefano, N., Geurts, J.J.G., Henry, R.G., Horsfield, M.A., Jenkinson, M., Pagani, E., Filippi, M., 2017. Brain MRI atrophy quantification in MS: From methods to clinical application. Neurology 88 (4), 403–413.
Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, A., De Stefano, N., 2002. Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis. NeuroImage 17 (1), 479–489.
Tsivgoulis, G., et al., 2015a. The effect of disease modifying therapies on brain atrophy in patients with relapsing-remitting multiple sclerosis: a systematic review and meta-analysis. PloS one. 10, e0116511.
Tsivgoulis, G., Katsanos, A.H., Grigoriadis, N., Hadjigeorgiou, G.M., Heliopoulos, I., Papathanasopoulos, P., Dardiotis, E., Kilidireas, C., Voumvourakis, K., 2015b. The effect of disease-modifying therapies on brain atrophy in patients with clinically isolated syndrome: a systematic review and meta-analysis. Ther. Adv. Neurol. Disord. 8 (5), 193–202.
Wang, C., et al., 2016. Automated brain volumetrics in multiple sclerosis: a step closer to clinical application. J. Neurol. Neurosurg. Psychiatry. 87, 754–757.
Wattjes, M.P., et al., 2015. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis–establishing disease prognosis and monitoring patients. Nat Rev Neurol. 11, 597–606.
Weinstock-Guttman, B., et al., 2018. Assessing 'No Evidence of Disease Activity' Status in Patients with Relapsing-Remitting Multiple Sclerosis Receiving Fingolimod in Routine Clinical Practice: A Retrospective Analysis of the Multiple Sclerosis Clinical and Magnetic Resonance Imaging Outcomes in the USA (MS-MRIUS) Study. CNS Drugs. 32, 75–84.
Zivadinov, R., Bakshi, R., 2004. Role of MRI in multiple sclerosis II: brain and spinal cord atrophy. Front Biosci. 9, 647–664.
Zivadinov, R., et al., 2008. Mechanisms of action of disease-modifying agents and brain volume changes in multiple sclerosis. Neurology. 71, 136–144.
Zivadinov, R., et al., 2016. Clinical relevance of brain atrophy assessment in multiple sclerosis. Implications for its use in a clinical routine. Expert Rev Neurother. 1–17.
Zivadinov, R., et al., 2018a. Feasibility of Brain Atrophy Measurement in Clinical Routine without Prior Standardization of the MRI Protocol: Results from MS-MRIUS, a Longitudinal Observational, Multicenter Real-World Outcome Study in Patients with Relapsing-Remitting MS. AJNR Am J Neuroradiol. 39, 289–295.
Zivadinov, R., et al., 2018b. No evidence of disease activity in patients receiving fingolimod at private or academic centers in clinical practice: a retrospective analysis of the multiple sclerosis, clinical, and magnetic resonance imaging outcomes in the USA (MS-MRIUS) study. Curr Med Res Opin. 34, 1431–1440.
Zivadinov, R., et al., 2018c. Fingolimod's Impact on MRI Brain Volume Measures in Multiple Sclerosis: Results from MS-MRIUS. J Neuroimaging. 28, 399–405.
Zivadinov, R., et al., 2019. Impact of fingolimod on clinical and magnetic resonance imaging outcomes in routine clinical practice: A retrospective analysis of the multiple sclerosis, clinical and MRI outcomes in the USA (MS-MRIUS) study. Mult. Scler. Relat. Disord. 27, 65–73.