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

R. Zivadinov, N. Bergsland, J.R. Korn, M.G. Dwyer, N. Khan, J. Medin, J.C. Price, B. Weinstock-Guttman, and D. Silva; on behalf of the MS-MRIUS Study Group

ABSTRACT

BACKGROUND AND PURPOSE: Feasibility of brain atrophy measurement in patients with MS in clinical routine, without prior standardization of the MRI protocol, is unknown. Our aim was to investigate the feasibility of brain atrophy measurement in patients with MS in clinical routine.

MATERIALS AND METHODS: Multiple Sclerosis and Clinical Outcome and MR Imaging in the United States (MS-MRIUS) is a multicenter (33 sites) retrospective study that included patients with relapsing-remitting MS who began treatment with fingolimod. Brain MR imaging examinations previously acquired at the baseline and follow-up periods on 1.5T or 3T scanners with no prior standardization were used, to resemble a real-world situation. Brain atrophy outcomes included the percentage brain volume change measured by structural image evaluation with normalization of atrophy on 2D-T1-weighted imaging and 3D-T1WI and the percentage lateral ventricle volume change, measured by VIENA on 2D-T1WI and 3D-T1WI and NeuroSTREAM on T2-fluid-attenuated inversion recovery examinations.

RESULTS: A total of 590 patients, followed for 16 months, were included. There were 585 (99.2%) T2-FLAIR, 425 (72%) 2D-T1WI, and 166 (28.2%) 3D-T1WI longitudinal pairs of examinations available. Excluding MR imaging examinations with scanner changes, the analyses were available on 388 (65.8%) patients on T2-FLAIR for the percentage lateral ventricle volume change, 259 and 257 (43.9% and 43.6%, respectively) on 2D-T1WI for the percentage brain volume change and the percentage lateral ventricle volume change, and 110 (18.6%) on 3D-T1WI for the percentage brain volume change and percentage lateral ventricle volume change. The median annualized percentage brain volume change was −0.31% on 2D-T1WI and −0.38% on 3D-T1WI. The median annualized percentage lateral ventricle volume change was 0.95% on 2D-T1WI, 1.47% on 3D-T1WI, and 0.90% on T2-FLAIR.

CONCLUSIONS: Brain atrophy was more readily assessed by estimating the percentage lateral ventricle volume change on T2-FLAIR compared with the percentage brain volume change or percentage lateral ventricle volume change using 2D- or 3D-T1WI in this observational retrospective study. Although measurement of the percentage brain volume change on 3D-T1WI remains the criterion standard and should be encouraged in future prospective studies, T2-FLAIR–derived percentage lateral ventricle volume change may be a more feasible surrogate when historical or other practical constraints limit the availability of percentage brain volume change on 3D-T1WI.

ABBREVIATIONS: MS-MRIUS = Multiple Sclerosis and Clinical Outcome and MR Imaging in the United States; PBVC = percentage brain volume change; PLVVC = percentage lateral ventricle volume change; RRMS = relapsing-remitting MS; SIENA = structural image evaluation with normalization of atrophy

Multiple sclerosis is an inflammatory and neurodegenerative autoimmune disease of the central nervous system characterized by demyelination and axonal degeneration. The measurement of brain atrophy has become one of the most important biomarkers for assessing the extent of neurodegenerative pathology in patients with MS. Development of brain atrophy in patients with MS is 3–5 times accelerated compared with the healthy aging population, and correlates with physical and cognitive disability from the earliest disease stages, and continues throughout the course of the disease.
Evidence is mounting that the development of disability progression in patients with MS is partially independent of accumulation of active MR imaging lesions and substantially dependent on the development of brain atrophy.\textsuperscript{14} Hence, there is an increasing interest in evaluating the effect of disease-modifying therapy on decelerating brain volume loss in clinical trials and consequently making personalized, patient-centric treatment choices.\textsuperscript{2,4,15,16} Therefore, there is an urgent need for translation and incorporation of brain atrophy measurement into clinical routine and patient-level treatment monitoring.\textsuperscript{3,4}

While the need for assessment of brain atrophy on the individual patient level for the effectiveness of treatment monitoring has become a hot topic in the literature recently,\textsuperscript{15} there is little understanding of how major obstacles are to be overcome. The feasibility of brain atrophy measurement for the short term, midterm, and long term,\textsuperscript{2,4,17} using MR imaging sequences available in a clinical routine is currently unknown.

Against this background, we aimed to determine whether it is feasible to measure brain atrophy in patients with MS in clinical routine without prior standardization of the MR imaging protocol, using a multicenter study design that included academic and nonacademic centers in the United States to resemble the real-world situation. The study evaluated MR imaging scanner strength, type, and the quality of pulse sequence characteristics and investigated whether MR imaging changes influence the measurement of brain atrophy on the most commonly used pulse sequences in a clinical routine on a group level over the midterm.

**MATERIALS AND METHODS**

The Multiple Sclerosis and Clinical Outcome and MR Imaging in the United States (MS-MRIUS) study is a multicenter, longitudinal, retrospective, observational chart review of patients with MS treated with fingolimod (Gilenya) in clinical routine practice. The methodologic approach and design of this retrospective study have been previously reported.\textsuperscript{18}

Briefly, clinical information and digital brain MR imaging image data were retrospectively collected from 33 participating academic and nonacademic MS centers across the United States and integrated into a central research data base.

The inclusion criteria were the following: 1) adult patients younger than 65 years of age with relapsing-remitting MS (RRMS) able to walk 20 m with or without assistance at the index date (defined as the date the patient first received treatment with fingolimod), equivalent to an Expanded Disability Status Scale\textsuperscript{19} score of \( \leq 6.5 \); 2) starting fingolimod at the index date and remaining on fingolimod for at least 28 days; 3) the availability of clinical data \( \pm 12 \) months from baseline (index date); and 4) a minimum of index and follow-up (postindex date) MR imaging examinations performed from 6 months before to 1 month after the index date and 9–24 months after index, respectively. Key exclusion criteria were the following: 1) neurologic disease other than MS affecting CNS structure or function; 2) a history of alcohol or substance abuse; 3) participation in an interventional trial during the study period; 4) prior use of fingolimod or natalizumab; and 5) steroid treatment in the 30 days prior to the scan dates.

A subgroup of 184 patients had preindex scans performed 9–24 months before fingolimod initiation (median, 13.9 months), but this was not a required inclusion criterion. This subgroup of patients was used only to investigate serial longitudinal changes (3 time points) of brain volume, measured for approximately 30 months.

All demographic and clinical data required for this study were collected from patient medical records into a study-specific electronic clinical research form. For each patient, clinical information was collected for a 48-month period, including 12–24 months’ data in the pre- and postindex period.

MR imaging examinations should have been performed on 1.5T or 3T scanners, and no standardization was expected. Patients did not need to have study examinations performed on the same scanner type and strength to resemble a real-world situation. 2D- or 3D-T2 fluid-attenuated inversion recovery and 2D- or 3D-T1-weighted images were collected. The participating sites transferred digital images using the standard DICOM format. To ensure that patient privacy was protected and that we adhered to relevant regulations, the centralized imaging center followed guidance from DICOM PS3.15 2015b, Security and System Management Profiles, Annex E: Attribute Confidentiality Profiles (http://dicom.nema.org/medical/Dicom/2015b/output/html/part15.html#chapter_E).\textsuperscript{20} Automatic de-identification via the on-line transfer portal was performed for all study scans. This pathway was the simplest and least burdensome for the sites, because all sites had digital transfer capability. DICOM images were automatically anonymized before transmission to the centralized imaging center via encrypted channels, and there was no “burned-in” information on the images.

All scans were inspected by an experienced rater at the centralized imaging center. We evaluated the following quality metrics: section thickness, excessive patient motion (“yes” or “no”), image contrast (“bad,” “acceptable,” or “good”), and overall quality (“bad,” “acceptable,” or “good”). The overall quality metric reflected anatomic coverage, the presence of imaging artifacts, noise level, and contrast. Examinations with excessive patient motion or bad image contrast automatically received a bad rating for overall quality. Additionally, for each MR imaging examination, differences in hardware model, scanner software, and the coil between index and postindex were evaluated. For each MR imaging sequence (2D- or 3D-T2-FLAIR, 2D-T1WI, and 3D-T1WI), differences in orientation, thickness, and protocol changes were examined. Then, overall hardware, software, coil, or protocol differences between time points were determined. Hardware change was defined as a change in the MR imaging scanner. Software was defined as an upgrade to a different software version using the same scanner. Coil change was defined as a change of the coil. Protocol change was defined as a meaningful change in TR/TE/TI/flip angle/geometry. When the hardware changes occurred, in almost all instances, software, coil, and protocol changes were noticed; therefore, we refer to those as the MR imaging scanner change. When software, coil, or protocol changes occurred without hardware change, we refer to those as software/coil/protocol MR imaging changes.

Longitudinal brain atrophy outcomes included percentage brain volume change (PBVC) measured by structural image evaluation, with normalization of atrophy (SIE/NA)\textsuperscript{21} on 2D-T1WI and 3D-T1WI and the percentage lateral ventricle volume change.
(PLVVC) measured by VIENA on 2D-T1WI and 3D-T1WI examinations and measured by NeuroSTREAM on T2-FLAIR. Lesions were inpainted before segmentation to reduce the impact of T1 hypointensities. All outcomes of brain atrophy analyses were assessed by an experienced rater. Because hardware changes can affect longitudinal measurements, we explored the stability of this measure in T2-FLAIR.

To investigate correlations between PLVVC on T2-FLAIR periods during the follow-up, and baseline volume, as appropriate. To investigate correlations between PLVVC on T2-FLAIR and baseline volume, as appropriate. We performed Spearman rank correlations. A nominal P value of ≤ .05 was considered statistically significant, using 2-tailed tests.

**RESULTS**

**Demographic and Clinical Characteristics at Index and Postindex**

Table 1 outlines demographic and clinical characteristics of the study subjects, according to the MR imaging scanner changes between index and postindex. Of the 590 patients with RRMS included in the study, 464 (78.6%) were women. Of the 33 centers participating in the study, 25 (75.8%) centers were nonacademic and 8 (24.2%) were academic; 398 (67.5%) patients with RRMS were collected in nonacademic, and 192 (32.5%), in academic centers. Between index and postindex follow-ups, MR imaging scanner changes occurred in 284 (48.1%) patients with RRMS. The median follow-up between index and postindex was approximately 16 months.

The median age was somewhat lower in patients with RRMS with MR imaging scanner changes compared with those without (P = .007). The number of relapses in 2 years before index was significantly higher in patients with RRMS with MR imaging scanner changes compared with those without (P = .001). Patients with RRMS without MR imaging scanner changes had higher Expanded Disability Status Scale scores (P = .002). No significant differences between patients with RRMS with and without MR imaging scanner changes were detected for disease duration or previous type of disease-modifying treatment at index.

**MR Imaging Acquisition Characteristics at Index and Postindex**

At postindex, there was a higher proportion of patients examined on 3T scanners (34.1%) compared with the index (29.5%, Online Table 1). All except 3 patients at index (99.5%) and 4 patients...
Table 2: Brain atrophy measures of the study subjects between index and postindex, according to MRI changesa

| Total Patients with RRMS | Patients with RRMS | Patients with RRMS without MRI | P Valueb |
|-------------------------|-------------------|-------------------------------|----------|
| PBVC on 2D-T1WI         | n = 259           | n = 89                        | n = 170  | .233 |
|                         | −0.31 (−0.74−0.13) | −0.35 (−0.97−0.11)            | −0.28 (−0.67−0.14) | |
| PBVC on 3D-T1WI         | n = 110           | n = 37                        | n = 73   | .740 |
|                         | −0.38 (−0.88 to −0.04) | −0.38 (−0.93−0.02)           | −0.37 (−0.81 to −0.08) | |
| PLVVC on 2D-T1WI        | n = 257           | n = 89                        | n = 168  | .848 |
|                         | 0.95 (−1.77−3.99) | 0.95 (−1.92−4.1)              | 0.96 (−1.63−3.94) | |
| PLVVC on 3D-T1WI        | n = 110           | n = 37                        | n = 73   | .940 |
|                         | 1.47 (−1.26−4.39) | 1.5 (−1.36−4.15)              | 1.01 (−1.17−4.46) | |
| PLVVC on T2-FLAIR       | n = 388           | n = 130                       | n = 258  | .79 |
|                         | 0.90 (−2.91−5.94) | 0.21 (−4.67−7.2)              | 1.1 (−2.1−5.51) | |
| PLVVC on T2-FLAIRc      | n = 554           | n = 291                       | n = 263  | .007 |
|                         | 0.51 (−3.95–6.01) | −0.31 (−6.4–6.9)              | 1.13 (−2.15–5.53) | |

aPBVC and PLVVC on 2D-, 3D-T1WI and T2-FLAIR were obtained in patients with software/coil/protocol scanner changes. All values in the table are presented as the number of patients for a particular analysis, median, and interquartile range. The longitudinal follow-up values are annualized.
bP values represent patients with MS with and without MRI changes during the follow-up group comparisons and were derived using analysis of covariance adjusted for age, sex, days between index and postindex, and baseline volume as appropriate.
cPLVVC on T2-FLAIR was performed in patients with hardware/software/coil/protocol changes.

at postindex (99.3%) had T2-FLAIR examinations, with most being 2D acquisitions, except for 9 examinations at index and 16 examinations at postindex, which were 3D. The use of 2D-T1WI decreased during follow-up from 79.7% at index to 75.6% at postindex, whereas the use of 3D-T1WI increased from 31.4% at index to 39.7% at postindex. Only 16.6% index and 20.7% postindex examinations had both 2D- and 3D-T1WI. A section thickness of ≤5 mm was present in 40%–50% of the T2-FLAIR and 35%–40% of the 2D-T1WI examinations. More than 85% of 3D-T1WI examinations were ≤2 mm thick. There was minimal-to-no excessive patient motion for all sequences, and scanner contrast and overall quality of the examinations were generally acceptable or good. 3D-T1WI examinations had superior quality to 2D-T1WI, with 80%–90% having good quality compared with 60%–70% on T2-FLAIR and 40%–45% on 2D-T1WI.

At index and postindex, there was a significantly higher proportion of 3T scanners (P < .001) and more 2D- and 3D-T1WI examinations in academic centers (On-line Table 2). The section thickness for various pulse sequences was lower in academic centers (≤5-mm section thickness for T2-FLAIR and 2D-T1WI and ≤2-mm section thickness for 3D-T1WI) (On-line Table 2). No differences in patient motion were found between academic and nonacademic centers, but scanner contrast and overall scan quality were better in the academic centers (On-line Table 2).

Eligibility of MR Imaging Examinations for Calculation of Brain Atrophy Outcomes

On-line Table 3 shows the eligibility of a longitudinal pair of examinations for calculation of PBVC and PLVVC measures between index and postindex. Among index-to-postindex examinations, hardware changes were identified in 29.5%; software, in 27.3%; and coil, in 31.9% of the longitudinal pairs. When pooled together, hardware/software/coil changes were identified in about 50% of the examinations, more frequently in the academic than in nonacademic centers (59.9% versus 42.2%, P = .002). A similar frequency of MR imaging protocol changes during the follow-up occurred on a pulse sequence basis for T2-FLAIR, 2D-T1WI, and 3D-T1WI, with higher rates of change in the academic than in nonacademic centers.

There were 583 (99.2%) T2-FLAIR, 425 (72%) 2D-T1WI, and 166 (28.1%) 3D-T1WI examinations eligible for calculation of brain atrophy outcomes. If we excluded longitudinal pairs of examinations that underwent hardware changes (174, 29.5%), analyses of brain atrophy outcomes were possible for 388 (65.8%) patients using PLVVC on T2-FLAIR, 259 and 257 (43.9% and 43.6%, respectively) patients using PBVC and PLVVC on 2D-T1WI, and 110 (18.6%) patients using PBVC and PLVVC on 3D-T1WI.

Changes in Brain Atrophy Measures during Follow-Up

At index (data not shown), no differences were seen in brain atrophy measures for any pulse sequence when stratified according to MR imaging software/coil/protocol changes occurring during the follow-up, except for the 3D-T1WI lateral ventricle volume, which was significantly higher in the group that underwent MR imaging changes (P = .05).

There were no significant differences in PBVC or PLVVC measures between patients with RRMS with and without software/coil/protocol MR imaging changes from index to postindex for any of the MR imaging sequence types explored. Median annualized PBVC was −0.31% on 2D-T1WI and −0.38% on 3D-T1WI. Median annualized PLVVC was 0.95% on 2D-T1WI, 1.47% on 3D-T1WI, and 0.90% on T2-FLAIR (Table 2). In a subgroup of 91 patients with RRMS with available preindex, index, and postindex MR imaging examinations, no significant differences in brain atrophy measures were found between patients with RRMS with and without software/coil/protocol MR imaging changes (On-line Table 4).

Using all patients with RRMS, who had analysis for PLVVC on T2-FLAIR examination (n = 554) independent of hardware changes, we found a significant index-to-postindex difference when comparing those with and without MR imaging scanner changes (−0.31% versus 1.13%, P = .007, Table 2), but not in a subgroup of 184 patients with RRMS with available preindex,
index, and postindex MR imaging examinations (On-line Table 4).

When only MR imaging hardware change was considered, a significantly different PLVVC was observed for patients imaged with hardware changes between index and postindex \( (n = 166, 30\%) \) compared with those who were not \( (n = 388, 70\%) \) \((-1.23\% \text{ versus } 0.9\%, P = .02)\).

**Correlation among Different Brain Atrophy Measures Using Different MR Imaging Pulse Sequences**

On-line Table 5 shows the correlations between PLVVC on T2-FLAIR using NeuroSTREAM and PLVVC and PBVC on 2D-T1WI and 3D-T1WI using VIENA and SIENA, respectively, in patients with MS with and without software/coil/protocol MR imaging changes. PLVVC on 3D-T1WI, 2D-T1WI, and T2-FLAIR was significantly associated with PBVC on 3D-T1WI and 2D-T1WI when all patients with MS were considered. As expected, correlations were stronger in patients with MS who were not imaged with software/coil/protocol changes, compared with those who were. PBVC on 3D-T1WI was associated with PLVVC on 3D-T1WI \( (r = −0.7, P < .0001) \), PLVVC on T2-FLAIR \( (r = −0.5, P < .0001) \), and PLVVC on 2D-T1WI \( (r = −0.39, P < .0001) \) in patients with MS imaged without MR imaging software/coil/protocol changes during the follow-up.

**DISCUSSION**

This is one of the first multicenter, retrospective, real-world studies that investigated the feasibility of brain atrophy measurement in a clinical routine without prior standardization of the MR imaging protocol, using academic and nonacademic centers specialized in treatment and monitoring of MS in the United States. The main findings of the study can be summarized as follows: 1) The quality of the MR images used for brain atrophy analyses were mostly acceptable or good in all centers; 2) about 70% of the centers used 1.5T field strength, and there was a tendency for higher use of 3T scanners during the follow-up; 3) academic centers performed MR imaging examinations with thinner sections, better contrast, and overall scan quality; they also used 3T scanners more frequently, with >50% of postindex examinations being performed on 3T, and had a higher proportion of 3D-T1WI examinations; 4) T2-FLAIR examinations were used in >99% of patients, while the figures were 72% for 2D-T1WI and 28% for 3D-T1WI examinations longitudinally; 5) scanner changes occurred in >50% of the patients during the follow-up, and the changes occurred more frequently in the academic, compared with nonacademic centers; 6) measurement of PLVVC on T2-FLAIR was feasible in 66% of patients, making it the most suitable measure of brain atrophy for clinical routine, while PBVC was obtained in 44% of patients on 2D-T1WI and 18% on 3D-T1WI; 7) excluding MR imaging hardware changes, there were no significant differences of brain atrophy outcomes in patients with and without MR imaging software/coil/protocol changes from index to postindex; 8) hardware changes resulted in significant PLVVC differences on T2-FLAIR, though this was not evident in the subgroup of patients with 3 serial MR imaging examinations during 30 months; and 9) finally, PBVC changes on 2D- and 3D-T1WI in patients treated with fingolimod for 16 months were similar to those reported in pivotal and/or open-label observational studies.

Clinical routine MR imaging examinations pose many unique challenges. They have a lower signal-to-noise ratio and/or spatial resolution due to trade-offs in scanning time. They lack standardization, which, in turn, is compounded by changes in MR imaging hardware and/or software upgrades. In the MS-MRIUS study, we confirmed that the spatial resolution was lower than that used in MS research clinical trials; however, most of the scans were of acceptable or good quality with negligible patient motion.

The MS-MRIUS study detected important differences in the type of MR imaging scanners used between academic and nonacademic centers in the United States. On the basis of some previous reports of the use of MR imaging in the referral of patients with MS to academic centers in the United States, it could be expected that the adherence to the Consortium of MS Centers and Magnetic Resonance Imaging in MS (MAGNIMS) MR imaging acquisition protocol guidelines for MS is somewhat lower in nonacademic compared with academic centers, which was exactly what we found in the current study. However, MR imaging scanner changes during the follow-up occurred even more frequently in academic than in nonacademic centers, which could be attributed to academic centers upgrading software more frequently or replacing their scanning technology more rapidly.

Image resolution and image contrast are important for a reliable and optimal segmentation of brain volume, and 3D pulse sequences are preferred for measurement of brain atrophy as the criterion standard for brain volumetric imaging because of the reduction of partial volume effects and more accurate coregistration, especially for serial imaging with time, compared with 2D-pulse sequences. Although 3D-T1WI is the recommended sequence for the calculation of brain volume measures, it is considered only as an optional sequence in the current MR imaging acquisition guidelines. The MS-MRIUS study showed that less than one-quarter of patients with MS had eligible 3D-T1WI for estimation of brain volume measures during the follow-up, which substantially impacted applicability for use in clinical routine. On the other hand, the MS-MRIUS study showed that T2-FLAIR was eligible in 99% of patients, allowing reliable PLVVC analysis in 66% of patients during the follow-up. Therefore, measurement of PLVVC on T2-FLAIR increased by 33% the proportion of patients who obtained reliable brain atrophy measurements compared with 2D-T1WI and by 73% compared with 3D-T1WI, respectively.

Given that more than half of the patients underwent changes in the MR imaging scanner during the follow-up, there is a strong need for a universally applicable panel of simple brain volume measures that are resistant to MR imaging scanner changes. For proper estimation of brain volume changes with time, patients should undergo an acquisition with the same hardware and without software/coil/protocol changes. However, this is very difficult to achieve in clinical routine because of a number of different logistic factors. The MS-MRIUS study showed no significant impact of software/coil/protocol changes on the measurement of PBVC and PLVVC in the clinical routine during 16 months, as evaluated using 3 different types of software on 3 different types of sequences. In particular, PLVVC is of interest as a choice for brain...
atrophic assessment in the clinical routine because it is relatively robust to the negative impact of imprecise positioning, gradient distortions, incomplete head coverage, and other motion and wraparound artifacts. In addition, PLVVC occurs at a 5-fold greater rate compared with PBVC, and the effect size of PBVC and PLVVC for the estimation of disability progression from baseline to 10-year follow-up is similar. In a previous study, the correlation analysis between PLVVC on T2-FLAIR using NeuroSTREAM and PLVVC and PBVC on 3D-T1WI using VIENA and SIENA, respectively, was described. The correlation analysis from the present study showed similar associations between PBVC on 3D-T1WI and PLVVC on T2-FLAIR in patients with MS who did not experience software/coil/protocol MR imaging changes during the follow-up, corroborating that PLVVC on T2-FLAIR can be used reliably in retrospective and prospective observational studies without prior standardization of the MR imaging protocol. However, measurement of brain atrophy on T2-FLAIR using PLVVC is still suboptimal compared with PBVC on 3D-T1WI, which should prompt clinicians in the more extensive use of the latter in clinical practice. The correlation analysis from the present study showed similar associations between PBVC on 3D-T1WI and PLVVC on T2-FLAIR in patients with MS who did not experience software/coil/protocol MR imaging changes during the follow-up, corroborating that PLVVC on T2-FLAIR can be used reliably in retrospective and prospective observational studies without prior standardization of the MR imaging protocol. However, measurement of brain atrophy on T2-FLAIR using PLVVC is still suboptimal compared with PBVC on 3D-T1WI, which should prompt clinicians in the more extensive use of the latter in clinical practice. In line with other recent studies, the MS-MRIUS study showed that scanner changes had an impact on brain volume estimates. While it was previously shown that NeuroSTREAM-derived PLVVC is relatively robust to different field strengths when imaged during a short time (approximately 2% coefficient of variation in the 1.5T versus 3T scan-rescan test for 72 hours), the current study showed that PLVVC on T2-FLAIR was significantly different in patients with RRMS with hardware changes, compared with those without. If the measurement of brain atrophy is to become an additional assessment tool for monitoring individual patients with MS, the annualized rate of pathologic cutoff will have to be lower than the error rate introduced by hardware changes or the analyses will be considered invalid for those patients. Although the MS-MRIUS study was not designed to answer this important question, it provides the first evidence of the feasibility of brain atrophy measurement in clinical routine without prior standardization of the MR imaging protocol. Future studies should investigate in greater detail the influence of individual components of scanner changes on a variety of brain atrophy measures, applicable to clinical routine, over the short term, midterm, and long term.

CONCLUSIONS
We showed, in this retrospective observational study, that T2-FLAIR was the most frequent sequence in the clinical routine. To increase general applicability of brain atrophy measurement in observational studies in the clinical routine, one can more feasibly estimate brain atrophy by assessing PLVVC on T2-FLAIR compared with PBVC or PLVVC using 2D- or 3D-T1WI. As the most accurate and well-established measurement of brain atrophy, PBVC assessment on 3D-T1WI is, and should remain, the criterion standard of brain volumetric imaging research. However, T2-FLAIR-derived PLVVC may be a more feasible surrogate when historical or other practical constraints limit the availability of PBVC on 3D-T1WI.

APPENDIX
MS-MRIUS study group:
Principal Investigator: R. Zivadinov.
Co-Principal Investigator: B. Weinstock-Guttman.
Study Steering Committee: R. Zivadinov, J. Medin, D. Silva, B. Weinstock-Guttman.
QuintilesIMS: N. Khan, J.R. Korn, J. Price.
Novartis: J. Medin, D. Silva.
MRI Analysis Center: R. Zivadinov, N. Bergsland, M.G. Dwyer, D.P. Ramasamy, E. Carl, Buffalo Neuroimaging Analysis Center, Department of Neurology, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo, State University of New York, Buffalo, New York.
Study sites:
Advanced Neurosciences Institute, Franklin, Tennessee: S. Hunter; Allegheny-Singer Research Institute, Pittsburgh, Pennsylvania: T. Scott; Barnabas Health, Livingston, New Jersey: K. Pandey; Central Texas Neurology Consultants, Round Rock, Texas: E. Fox; Central State Medical Center, Freehold, New Jersey: A. Katz; Christiana Care Health Systems, Newark, Delaware: J. Silversteen; College Park Family Care Center, Overland Park, Kansas: J. Kaplan; Denver Health and Hospital Authority, Denver, Colorado: E. Maa; Evergreen Health, Kirkland, Washington: V. Simnad; Georgetown University, Washington, DC: R. Shin; Island Neurological Associates, Plainview, New Jersey: S. Newman; Kinkel Neurologic, Williamsville, New York: P. Kinkel; Mercy Clinic Neurology, Chesterfield, Missouri: B. Green; Minneapolis Clinic of Neurology, Golden Valley, Minnesota: J. Calkwood; MS Center of Northeastern NY, Latham, New York: K. Edwards; Neurological Services of Orlando, Orlando, Florida: D. Jacobs; Neurology & Neurosurgery Associates, Akron, Ohio: D. Huang; Neurology Group of San Antonio, San Antonio, Texas: A. Bass; Neuroscience Group of Neenah, Neenah, Wisconsin: S. Hibbs; NYU Langone Medical Center, New York City, New York: I. Kister; Ohio Health, Columbus, Ohio: G. Eubank; OMRF MS Center of Excellence, Oklahoma City, Oklahoma: G. Pardo; Partners MS Center, Harvard, Boston, Massachusetts: C. Chitnis; Phoenix Neurological Associates, Phoenix, Arizona: B. Hendin; Providence St. Vincent Medical Center, Portland, Oregon: S. Cohan; Raleigh Neurology Associates, Raleigh, North Carolina: M. Freedman; South Shore Neurologic Associates, Patchogue, New York: M. Gudesblatt; SUNY University of Buffalo, Buffalo, New York: B. Weinstock-Guttman; The Elliot Lewis Center for MS Care, Boston, Massachusetts: E. Lathi; University of Colorado, Denver, Colorado: E. Alvarez; University of Rochester, Rochester, New York: A. Goodman; University of Tennessee, Knoxville, Tennessee: R. Trudell; Washington University, St Louis, Missouri: R. Naismith.

Disclosures: Robert Zivadinov—RELATED: Grant: QuintilesIMS; UNRELATED: Consultancy: Novartis, Sanofi Genzyme, Celgene; Grants/Grants Pending: Novartis, Sanofi Genzyme, QuintilesIMS; Payment for Lectures Including Service on Speaking Bureaus: Novartis, Sanofi Genzyme. Jonathan R. Korn—RELATED: Consulting Fee or Honorarium: QuintilesIMS; Comments: QuintilesIMS received funding from Novartis to perform this study.* Michael G. Dwyer—RELATED: Grant: Novartis.* UNRELATED: Consultancy: Claret Medical, EMD Serono. Nasreen Khan—RELATED: Consulting Fee or Honorarium: QuintilesIMS; Comments: The work performed was paid by Quintiles IMS. The QuintilesIMS study was supported by Novartis; UNRELATED: Employment: Real-World Economics Consultancy; Comments: I am the owner of the consultancy firm Real-World Economics Consultancy, which provides consulting services for pharmaceutical and other consulting companies. In the past, I have
brain atrophy correlates with treatment effect on disability in multiple sclerosis. Ann Neurol 2014;75:43–49 CrossRef Medline
15. Zivadinov R, Dwyer MG, and Bergland N. Brain atrophy measurements should be used to guide therapy monitoring in MS: YES. Mult Scler 2016;22:1522–24 CrossRef Medline
16. Azevedo C, Pelletier D. Whole-brain atrophy: ready for implementation into clinical decision-making in multiple sclerosis? Curr Opin Neurol 2016;29:237–42 CrossRef Medline
17. Rocca MA, Battaglini M, Benedict RH, et al. Brain MRI atrophy quantification in MS: from methods to clinical application. Neurology 2017;88:403–13 CrossRef Medline
18. Zivadinov R, Khan N, Medin J, et al. An observational study to assess brain MRI change and disease progression in multiple sclerosis clinical practice: the MS-MRIUS study. J Neuroimaging 2017;27:339–47 CrossRef Medline
19. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33:1444–52 CrossRef Medline
20. Fetzer DT, West OC. The HIPAA privacy rule and protected health information: implications in research involving DICOM image databases. Acad Radiol 2008;15:590–95 CrossRef Medline
21. Smith SM, Zhang Y, Jenkinson M, et al. Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage 2002;17:479–89 CrossRef Medline
22. 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;35:1101–10 CrossRef Medline
23. Dwyer MG, Silva D, Bergland N, et al. Neurological software tool for reliable atrophy measurement (NeuroSTREAM) of the lateral ventricles on clinical-quality T2-FLAIR MRI scans in multiple sclerosis. Neuroimage Clin 2015;8:759–79 CrossRef Medline
24. Gelineu-Morel R, Tomassini V, Jenkinson M, et al. The effect of hypointense white matter lesions on automated gray matter segmentation in multiple sclerosis. Hum Brain Mapp 2012;33:2802–14 CrossRef Medline
25. Chu R, Tahuvid S, Glanz BI, et al. Whole brain volume measured from 1.5T versus 3T MRI in healthy subjects and patients with multiple sclerosis. J Neuroimaging 2016;26:62–67 CrossRef Medline
26. Lysandropoulos AP, Abisal J, Metens T, et al. Quantifying brain volumes for multiple sclerosis patients follow-up in clinical practice: comparison of 1.5 and 3 Tesla magnetic resonance imaging. Brain Behav 2016;6:e00422 CrossRef Medline
27. Biberacher V, Schmidt P, Keshavan A, et al. Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. Neuroimage 2016;142:188–97 CrossRef Medline
28. De Stefano N, Silva DG, Barnett MH. Effect of fingolimod on brain volume loss in patients with multiple sclerosis. CNS Drugs 2017;31:289–305 CrossRef Medline
29. Carmosino MJ, Brousseau KM, Arciniegas DB, et al. Initial evaluations for multiple sclerosis in a university multiple sclerosis center: outcomes and role of magnetic resonance imaging in referral. Arch Neurol 2005;62:585–90 CrossRef Medline
30. Traboulsee A, Simon JH, Stone L, et al. Revised Recommendations of the Consortium of MS Centers Task Force for a Standardized MRI Protocol and Clinical Guidelines for the Diagnosis and Follow-Up of Multiple Sclerosis. AJNR Am J Neuroradiol 2016;37:394–401 CrossRef Medline
31. Rovira A, Watts MP, Tintoré M, et al; MAGNIMS study group. Evidence-based guidelines: MAGNIMS consensus guidelines on the use of MRI in multiple sclerosis-clinical implementation in the diagnostic process. Nat Rev Neurol 2015;11:471–82 CrossRef Medline