



Magnetic resonance disease severity scale (MRDSS) for patients with multiple sclerosis: A longitudinal study

Jennifer Moodie ^a, Brian C. Healy ^b, Guy J. Buckle ^b, Susan A. Gauthier ^c, Bonnie I. Glanz ^b, Ashish Arora ^b, Antonia Ceccarelli ^b, Shahamat Tauhid ^b, Xue-Mei Han ^b, Arun Venkataraman ^b, Tanuja Chitnis ^b, Samia J. Khoury ^b, Charles R.G. Guttmann ^d, Howard L. Weiner ^b, Mohit Neema ^{b,*}, Rohit Bakshi ^{b,1}

^a Department of Neurology, University of Massachusetts, Boston, MA, USA

^b Department of Neurology, Brigham and Women's Hospital, Laboratory for Neuroimaging Research, Partners MS Center, Harvard Medical School, Boston, MA, USA

^c Judith Jaffe Multiple Sclerosis Center, Weill Cornell Medical College, New York, USA

^d Department of Radiology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

ARTICLE INFO

Article history:

Received 31 October 2011

Accepted 30 November 2011

Available online 28 December 2011

Keywords:

MRI
Multiple sclerosis
Neurologic disability
T1 hypointense lesions
T2 hyperintense lesions
Brain atrophy
Measurement scales

ABSTRACT

Background: We previously described a composite MRI scale combining T1-lesions, T2-lesions and whole brain atrophy in multiple sclerosis (MS): the magnetic resonance disease severity scale (MRDSS).

Objective: Test strength of the MRDSS vs. individual MRI measures for sensitivity to longitudinal change.

Methods: We studied 84 MS patients over a 3.2 ± 0.3 year follow-up. Baseline and follow-up T2-lesion volume (T2LV), T1-hypointense lesion volume (T1LV), and brain parenchymal fraction (BPF) were measured. MRDSS was the combination of standardized T2LV, T1/T2 ratio and BPF.

Results: Patients had higher MRDSS at follow-up vs. baseline ($p < 0.001$). BPF decreased ($p < 0.001$), T1/T2 increased ($p < 0.001$), and T2LV was unchanged ($p > 0.5$). Change in MRDSS was larger than the change in MRI sub-components. While MRDSS showed significant change in relapsing–remitting (RR) ($p < 0.001$) and secondary progressive (SP) phenotypes ($p < 0.05$), BPF and T1/T2 ratio changed only in RRMS ($p < 0.001$). Longitudinal change in MRDSS was significantly different between RRMS and SPMS ($p = 0.0027$); however, change in the individual MRI components did not differ. Evaluation with respect to predicting on-study clinical worsening as measured by EDSS revealed a significant association only for T2LV ($p = 0.038$).

Conclusion: Results suggest improved sensitivity of MRDSS to longitudinal change vs. individual MRI measures. MRDSS has particularly high sensitivity in RRMS.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Conventional magnetic resonance imaging (MRI) measures of brain atrophy, T1 hypointense and T2 hyperintense lesion volume, provide a qualitative estimate of damage that has accumulated in patients with multiple sclerosis (MS). However, often there is a clinical and MRI dissociation, whereby conventional MRI findings show a weak relationship to clinical status, such as disability as measured by the Expanded Disability Status Scale (EDSS) [1]. Furthermore, ideal treatment planning and risk stratification would require the availability of accurate and reliable surrogates that predict disease course. While current MRI measures have a role in providing clinically relevant disease activity and severity to some extent, there remains a

need for improved biomarkers which are sensitive and comprehensive in their predictive ability.

Composite MRI measurements offer a new approach to assess various aspects of disease involvement which are reflected in a myriad of MRI findings, which demonstrate both inflammatory and neurodegenerative changes [2–6]. Previously, we developed a composite scale to define the severity of cerebral damage in MS, known as the magnetic resonance disease severity scale (MRDSS) [7], which combines three continuous measures: T2 hyperintense lesions, the ratio of T1 hypointense to T2 lesion volume, and whole brain atrophy (normalized whole brain volume). MRDSS showed a larger effect size than any of the individual MRI components in distinguishing among clinical phenotype groups and was associated with disability progression occurring during the subsequent three years. However the longitudinal sensitivity of MRDSS is not known.

The purpose of this study was to longitudinally assess the sensitivity and validity of the MRDSS over three years in a cohort of MS patients representing a wide range of disability and disease duration and all three major clinical phenotype groups. We focused on the sensitivity to change in MRDSS and its relationship to clinical disability in

* Corresponding author at: Brigham & Women's Hospital, Harvard Medical School, Laboratory for Neuroimaging Research, One Brookline Place, Suite 602, Brookline, MA 02445, USA. Tel.: +1 617 732 6568; fax: +1 617 732 6540.

E-mail address: mneema@bwh.harvard.edu (M. Neema).

¹ Both authors contributed equally to the project.

comparison to the change in the individual components of MRDSS assessed independently.

2. Methods

2.1. Subjects

Demographic and clinical characteristics of the subjects are summarized in Table 1. We retrospectively identified 84 patients with MS from a consecutive sample being prospectively enrolled and monitored as part of the Comprehensive Longitudinal Investigation of MS at the Brigham and Women's Hospital and Partners MS Center (CLIMB). CLIMB is an ongoing prospective observational cohort study that began following patients in 2000 [8]. Inclusion into the current study was based on the following criteria: 1) age 18–60 at baseline; 2) brain MRI at baseline and follow-up scan 2.5 to 3.5 years later performed at the Brigham and Women's Hospital on the 1.5 T unit dedicated for MS care using the scanning protocol established for the CLIMB study; 3) baseline and follow-up examination with EDSS scoring [9] performed by an MS specialist neurologist at the Partners MS Center; 4) baseline and follow-up EDSS testing performed within three months of the MRI; and 5) established MS diagnosis [10] at baseline of either relapsing–remitting (RR), secondary progressive (SP) or primary progressive (PP) by established criteria [11]. These patients were part of our initial study which focused on the relationship between baseline MRDSS only and clinical status; more details on the study population, study design, medication use, etc. have been previously published [7]. The current sample represents a subset of patients from the previous study in which follow-up MRI was available. The current study was approved by our institutional review board.

2.2. Progression of disability

Progression of neurologic disability at follow-up was defined as 1-point progression on EDSS if the baseline score was less than 6 or 0.5 point progression if the baseline score was 6 or higher [7]. EDSS worsening at follow-up had to be sustained for at least three months to be considered progression. Patients were thus classified as stable or progressed on EDSS at the three-year follow-up. These patients were followed for (mean \pm SD) 3.2 ± 0.3 years (see Table 1).

2.3. MR imaging

All patients underwent baseline and follow-up brain MRI on the same scanner using the same scanning protocol [7]. MRI was obtained on a Signa 1.5-T unit (GE Signa, General Electrics, Milwaukee, WI) using a quadrature head coil. Axial brain imaging included T1-

weighted spin-echo (TR/TE: 725/20) and dual-echo spin-echo T2-weighted (3000/80/30) images with $256 \times 256 \times 54$ voxels and a nominal voxel size of $0.9375 \times 0.9375 \times 3$ mm, without inter-slice gaps. After infusion of intravenous gadolinium (0.1 mM/kg), and a five-minute delay, axial T1-weighted spin-echo imaging was repeated.

2.4. MRI analysis

2.4.1. T2 hyperintense lesion volume and whole brain atrophy

Using automated template-driven segmentation (TDS+) [12] from the dual echo images, T2 hyperintense lesion volume (T2LV) and normalized whole brain volume [brain parenchymal fraction (BPF)] were determined, the latter of which was an estimate of whole brain atrophy [13].

2.4.2. T1 hypointense lesion volume

Analysis of hypointense lesions (black holes) on T1-weighted images was performed using a computer based semiautomated edge-finding tool [14]. A black hole was defined as a lesion appearing visibly hypointense to the surrounding white matter and showing at least partial hyperintensity on dual echo T2 images, but non-enhancement on post-gadolinium studies (to reduce the likelihood of including transient black holes) [15]. The presence of T1 hypointense lesions (black holes) was determined by consensus of two trained observers as part of a reading panel.

2.4.3. T1/T2 ratio

To describe the destructive potential of lesions, a ratio of T1 hypointense to T2 hyperintense lesion volume was created for each patient.

2.4.4. Reliability of MRI measurements

Template-driven segmentation with partial volume correction achieves an intraclass correlation of 0.994, interscan coefficient of variation of 4.98%, and mean \pm SD volume bias of 0.01 ± 0.68 mL [12]. The T1LV measurement showed intraobserver and interobserver coefficients of variation of 1.7% and 4.5%, respectively [7].

2.5. Creation of the MRDSS

A detailed description of the calculation of the original MRDSS was provided in a previous manuscript [7]; a brief summary of the calculation and changes to the model is provided here. All MRI data were rounded to 2 decimal places prior to analysis. Because the distributions of T2LV and T1:T2 were skewed, log (T2LV) or logistic (T1:T2) transformation was performed. The T2LV, BPF, and T1:T2 were then standardized (z) by subtracting the mean and dividing by the standard deviation of our sample at baseline. Patients with zero T1:T2 were not included in the normalization but were assigned a value more extreme (-2) than the most extreme nonzero standardized T1:T2 (-1.92), similar to the MS Functional Composite scale [16]. In the original paper, a more extreme value of -2.5 was used, but we found that -2 led to better performance when investigating changes over time in the MRDSS. Results using arcsin-square root transformation for the T1:T2 were similar and did not alter conclusions (data not shown). The individual standardized scores were equally weighted and summed for each patient: $zMRDSS = zT2LV + [zT1LV:T2LV] - zBPF$ ($z = [\text{raw score} - \text{mean}] / \text{standard deviation}$). Each subject's composite value was then transformed to a continuous 0 to 10 MRDSS score (zero is lowest severity) based on the highest and lowest $zMRDSS$ values.

2.6. Statistical analysis

A Wilcoxon signed rank test was used to investigate the longitudinal change in the MRDSS and each individual MRI marker in the

Table 1
Summary of clinical data.

Number of patients	84
Age (mean \pm SD, years)	42.6 \pm 8.6
Number of females (%)	56 (67%)
EDSS baseline score (mean \pm SD)	3.2 \pm 2.0
EDSS follow up score (mean \pm SD)	3.4 \pm 2.2
Number of patients showing sustained EDSS progression during the observation period	21 (25%)
Baseline disease category	
Relapsing–remitting n (%)	57 (68%)
Secondary progressive n (%)	20 (24%)
Primary progressive n (%)	7 (8%)
Length of follow-up (mean \pm SD, years)	3.2 \pm 0.3
Number of patients on disease-modifying therapy during the observation period (%)	65 (77%)

Legend: EDSS = Expanded Disability Status Scale.

entire sample and in the RRMS and SPMS patients separately. The difference in the longitudinal change between the RRMS and SPMS patients was investigated using a Wilcoxon rank sum test. The association between baseline EDSS and MRI measures and the change in the MRDSS was assessed using Spearman's (EDSS vs. change in MRDSS) and Pearson's correlation coefficient (MRI factors vs. change in MRDSS). Logistic regression was used to estimate the effect of baseline MRI variables and change in MRI variables on the probability of progression. Baseline EDSS was included as a categorical covariate in all prediction models with categories (0, 1–1.5, 2–2.5, 3–3.5, 4–5.5 and ≥6) for the entire group, categories (0, 1–1.5, 2–2.5, 3–3.5 and ≥4) for RRMS, and categories (0–3.5, 4–5.5, ≥6) for progressive patients. EDSS was also treated as a continuous covariate in secondary analyses, and the results were largely unchanged. Models controlling for length of follow-up and baseline age were also fit. Finally, the MRI characteristics of RRMS and SPMS patients at last visit were compared using a Wilcoxon rank sum test for each measure. A p-value less than 0.05 was considered significant; a p value less than 0.1 but greater than 0.05 was considered a trend.

3. Results

3.1. Clinical and MRI disability measures

The clinical data are summarized in Table 1. The EDSS scores for the group during the observation period remained stable (from 3.2 ± 2.0 to 3.4 ± 2.2), although 21 (25%) patients developed progression of physical disability during the 3 year follow up period.

3.2. Longitudinal change in MRI data

Table 2 shows the change in the MRDSS as well as the individual MRI measures over the three year follow-up in the whole cohort and within the clinical phenotype groups. For the BPF, T1/T2 and MRDSS, a statistically significant worsening of disease severity was observed on each (decreasing BPF and increasing T1/T2 or MRDSS). The MRDSS demonstrated the largest magnitude of change in terms of effect size over time in comparison to individual variables and lowest p value in the whole patient cohort (estimated mean change ± SD, 0.64 ± 0.80; p = 1.5 × 10⁻¹⁰). Within the MS clinical phenotypes, a larger magnitude of worsening in the MRDSS was observed in the RR compared to the SP patients (0.74 ± 0.60 vs. 0.27 ± 0.48; p = 0.0027), but the change over time in the MRDSS was statistically significant in both groups (p_{RRMS} = 1.2 × 10⁻⁹, p_{SPMS} = 0.021). In contrast, BPF and T1/T2 worsened significantly in the RR but not the SP group. Furthermore, when comparing the MRI change over time between the RR and SP groups, only the MRDSS showed a significant difference between groups among all MRI

variables. Throughout every aspect of this analysis, T2 lesion volume did not significantly change during the observation period in the entire population or in either subgroup. Analysis of individual raw vs. standardized MRI measures had little effect on these results (Table 2). In terms of the change relative to the baseline value, the distribution of the change in the standardized BPF and T2 lesion volume was approximately equal for all baseline values. Conversely, patients with low baseline values showed greater change in the T1/T2 and MRDSS compared to patients with high baseline values (data not shown).

3.3. Prediction of change in MRDSS

We tested the relationship between baseline data (both clinical and MRI measures) and subsequent on-study change in MRDSS (Table 3). This was performed in the whole cohort and then in specific MS clinical subgroups. Significant associations were observed in the whole cohort, and a lower baseline MRDSS, lower baseline T2LV, and lower T1/T2 were associated with larger worsening on MRDSS. However, these correlations remained in the weak to moderate range. It appeared that clinical phenotype group influenced these associations. When examining whether the RR or SP phenotype groups drove these significant correlations, the baseline MRDSS showed a significant but weak association with the change in MRDSS in the RR group only. The baseline MRDSS had similar correlation in the PP/SP group although the correlation was not statistically significant due to smaller sample sizes. Baseline T1/T2 was significantly associated with the change in both the RR and the combined PP/SP groups. Furthermore, a lower baseline BPF was associated with a higher risk of subsequent progression on MRDSS but only in the SP group. Thus, these results were mixed in that for the RR group less MRI-defined disease at baseline, whether measured by the MRDSS or its subcomponents, predicts a higher risk of on-study worsening of MRDSS. Conversely, for SP patients, an opposite effect was seen for baseline BPF in that worse disease (lower BPF) at baseline was congruent with the subsequent direction of change in MRDSS.

3.4. Prediction of change in disability

While these were reported from a larger cohort previously [7], we again tested the relationship between baseline MRI data (MRDSS and all individual subcomponent MRI measures) and subsequent on-study risk for developing sustained progression in disability on the EDSS scale in this smaller cohort. The current cohort was smaller than the previously published cohort because of the requirement for inclusion in the present study that patients have both baseline and 3 year follow-up MRI scans available from the same scanning platform and protocol. This was performed in the whole cohort and then on an exploratory basis among various clinical phenotype subgroups. Table 4 shows the results

Table 2
Change in MRDSS and individual MRI measures.

	All patients Mean ± SD (p-value) [§]	RR Mean ± SD (p-value) [§]	SP Mean ± SD (p-value) [§]	RR vs. SP, p value ⁺
MRDSS	0.64 ± 0.80 (1.5 × 10⁻¹⁰)	0.74 ± 0.60 (1.2 × 10⁻⁹)	0.27 ± 0.48 (0.021)	0.0027
BPF	-0.0078 ± 0.016 (8.8 × 10⁻⁵)	-0.0096 ± 0.016 (0.00012)	-0.0019 ± 0.014 (0.61)	0.16
T2LV	0.26 ± 2.61 (0.31)	0.27 ± 2.65 (0.28)	0.17 ± 2.93 (0.87)	0.52
T1/T2	0.059 ± 0.12 (1.6 × 10⁻⁶)	0.055 ± 0.097 (9.6 × 10⁻⁶)	0.052 ± 0.11 (0.070)	0.67
zBPF	-0.15 ± 0.30 (8.9 × 10⁻⁵)	-0.18 ± 0.31 (0.00012)	-0.035 ± 0.26 (0.62)	0.16
zT2LV	0.026 ± 0.28 (0.38)	0.039 ± 0.28 (0.33)	0.0023 ± 0.26 (0.90)	0.57
zT1/T2	0.47 ± 0.82 (6.8 × 10⁻⁷)	0.53 ± 0.75 (6.1 × 10⁻⁶)	0.24 ± 0.45 (0.064)	0.098

Legend: Values in the table are mean ± standard deviation; [§] = p-values from a Wilcoxon signed rank test report the change over three years within the group; ⁺ = p-values reported from a Wilcoxon rank sum test comparing change in RR and SP; MRDSS = magnetic resonance disability severity scale; BPF = brain parenchymal fraction; T2LV = T2 hyperintense lesion volume; T1/T2 = ratio of T1 lesion volume and T2 lesion volume; zBPF = standardized BPF; zT2LV = standardized T2 lesion volume; zT1/T2 = standardized ratio of T1 lesion volume and T2 lesion volume; RR = relapsing remitting; SP = secondary progressive.

Table 3
Relationship between baseline clinical or MRI variable and longitudinal change in MRDSS.

	r Value	p Value
<i>Baseline EDSS vs. change in MRDSS</i>		
All MS	−0.18	0.10
RR	0.08	0.54
SP	0.02	0.95
PP/SP	−0.07	0.74
<i>Baseline MRDSS vs. change in MRDSS</i>		
All MS	−0.34	0.002
RR	−0.32	0.014
SP	0.12	0.62
PP/SP	−0.32	0.099
<i>Baseline zBPF vs. change in MRDSS</i>		
All MS	0.10	0.37
RR	0.12	0.38
SP	−0.49	0.028
PP/SP	−0.03	0.90
<i>Baseline zT2LV vs. change in MRDSS</i>		
All MS	−0.26	0.016
RR	−0.18	0.17
SP	−0.04	0.86
PP/SP	−0.34	0.083
<i>Baseline zT1/T2 vs. change in MRDSS</i>		
All MS	−0.45	0.00002
RR	−0.47	0.0003
SP	−0.20	0.40
PP/SP	−0.41	0.034

Legend: Results are from correlation analysis of the baseline value and the change in MRDSS score. A positive correlation indicates that a higher baseline value was associated with higher change in the direction of worsening disease. EDSS = Expanded Disability Status Scale; MRDSS = magnetic resonance disease severity scale; zBPF = standardized brain parenchymal fraction; zT1/T2 = standardized ratio of T1 lesion volume and T2 lesion volume; zT2LV = standardized T2 hyperintense lesion volume; RR = relapsing remitting MS; SP = secondary progressive MS; PP = primary progressive MS.

comparing baseline MRDSS and individual standardized MRI measures as candidate predictors of on-study clinical worsening. When examining the whole cohort or subgroups of patients, the only significant

Table 4
Relationship between baseline MRI variables and EDSS-defined clinical progression.

Measure	Group	OR	p-Value
MRDSS	All	0.83	0.16
MRDSS	RR	1.01	0.94
MRDSS	SP	0.75	0.32
MRDSS	PP/SP	0.61	0.054
zBPF	All	0.93	0.82
zBPF	RR	0.71	0.41
zBPF	SP	1.07	0.91
zBPF	PP/SP	1.20	0.75
zT2	All	0.50	0.038
zT2	RR	0.78	0.53
zT2	SP	0.33	0.16
zT2	PP/SP	0.29	0.048
zT1/T2	All	0.68	0.19
zT1/T2	RR	1.02	0.97
zT1/T2	SP	0.64	0.50
zT1/T2	PP/SP	0.32	0.055

Legend: Results are from logistic regression between the baseline value and the outcome (whether or not the patient developed sustained progression of change in EDSS score during the observation period) controlling for baseline EDSS as a categorical variable. Similar results were observed when EDSS was treated as a continuous variable. An odds ratio greater than 1.0 indicates that a higher baseline value was associated with higher probability of progression – the opposite is true if the odds ratio is less than 1.0. OR = odds ratio; EDSS = Expanded Disability Status Scale; MRDSS = magnetic resonance disease severity scale; zBPF = standardized brain parenchymal fraction; zT1/T2 = standardized ratio of T1 lesion volume and T2 lesion volume; zT2LV = standardized T2 hyperintense lesion volume; RR = relapsing remitting; PP = primary progressive; SP = secondary progressive.

association was seen for a lower baseline T2LV and a higher risk of progressing on EDSS. However, this relationship was in the weak to moderate range. It appeared that clinical phenotype influenced this association as it was driven by patients with SP and PP MS (particularly PP). In agreement with these findings, the pooled PP/SP group also showed trends towards a lower baseline MRI disease severity (for MRDSS and other individual MRI subcomponents) predicting a higher risk of progressing on EDSS. Overall, these results were in agreement with models predicting change in MRDSS in that for particular clinical phenotypes, less cerebral MRI-defined disease at baseline predicted a higher risk of on-study disease progression on EDSS.

3.5. On-study relationship between change in MRI and disability

Table 5 demonstrates the association between the change in the MRI measures and the probability of developing sustained progression of physical disability during the observation period. There were no significant findings or trends in the whole cohort (Table 5) or when examining just RR or SP patients (all $p > 0.1$; data not shown). All results were similar after controlling for age and length of follow-up and when considering non-standardized (raw) individual measures (data not shown).

3.6. MRI–clinical relationships at exit

Data are summarized in Table 6. There were statistically significant differences between RRMS and SPMS with regard to MRDSS ($p = 0.0069$), zBPF ($p = 0.038$), and zT1:T2 ratio ($p = 0.0059$), all in the expected direction – SP had higher severity of these MRI measures than RR patients. However, there were no statistically significant differences in the zT2 lesion volume between groups ($p = 0.13$). All results were similar when considering non-standardized (raw) individual measures (data not shown). The significant differences between the groups on the MRI measures confirmed the findings in our previous article [7] in which similar significant MRI–clinical relationships were observed at baseline, thus demonstrating internal consistency.

4. Discussion

In the present study, we tested the strength of the MRDSS against individual MRI measures for sensitivity to longitudinal change in a relatively small group of patients with MS. The major finding of our study was that MRDSS showed improved sensitivity to longitudinal change vs. individual MRI measures of lesions and atrophy. MRDSS was better able to differentiate RR and SP patients than individual MRI components longitudinally and was the only variable that significantly changed in SP patients over time. Our data further support the potential utility of our

Table 5
Association between change in MRI measures and probability of sustained progression during the three year observation period.

Measure	Odds ratio (95% CI)	p-value
MRDSS	1.55 (0.80, 3.00)	0.19
zBPF	0.46 (0.08, 2.73)	0.39
zT2LV	0.99 (0.14, 7.01)	0.99
zT1/T2	1.41 (0.72, 2.74)	0.31

Legend: Only standardized MRI measures are shown because these ensure the scale is common across the individual variables. The analyses were controlled for baseline EDSS score as a categorical variable. MRDSS = magnetic resonance disease severity scale; zBPF = standardized brain parenchymal fraction; zT2LV = standardized T2 lesion volume; zT1/T2 = standardized ratio of T1 hypointense lesion volume and T2 hyperintense lesion volume.

Table 6
Comparison of exit MRI factors in RR and SP patients.

Measure	RRMS	SPMS	p-Value
MRDSS	5.15 ± 2.29	6.70 ± 1.90	0.0069
zBPF	−0.049 ± 0.99	−0.56 ± 0.98	0.038
zT2	0.0016 ± 0.93	0.42 ± 0.87	0.13
zT1/T2	−0.032 ± 1.01	0.61 ± 0.79	0.0059

Legend: Only standardized MRI measures are shown because these ensure the scale is common across the individual variables. Values in the table are mean ± standard deviation; MRDSS = magnetic resonance disease severity scale; zBPF = standardized brain parenchymal fraction; zT2LV = standardized T2 lesion volume; zT1/T2 = standardized ratio of T1 hypointense lesion volume and T2 hyperintense lesion volume; RR = relapsing–remitting; SP = secondary progressive.

neuroimaging-based composite scale to comprehensively evaluate disease severity in MS.

Our study directly compared the longitudinal performance of the MRDSS with two brain MRI measures of lesions (T1 hypointense and T2 hyperintense) and whole brain atrophy. While the individual MRI measures also showed some degree of longitudinal change, the MRDSS showed significant change in the entire patient cohort and within the patient subgroups to a better extent than the individual MRI measures, suggesting its sensitivity in both the inflammatory and neurodegenerative stages of the disease. Unlike the individual MRI measures, MRDSS was the only variable that significantly worsened in both RR and SP groups. While BPF and T1 hypointense lesions also showed a statistically significant change over time, there was no significant difference in the amount of change between the RR and SP phenotypes. Furthermore, T2 lesions did not significantly change over time in the entire patient cohort or within the patient subgroups. Similar results regarding lesions (T1 hypointense and T2 hyperintense) and whole brain atrophy have been reported in prior studies [17–19].

With regard to the relationship between baseline data (both clinical and MRI measures) and subsequent on-study change in MRDSS, there was a tendency towards larger worsening on MRDSS with a lower baseline MRDSS, lower baseline T2LV, and lower T1/T2. This result was driven in part by the upper limit imposed by the scale; however, even when the zMRDSS was used rather than the 0–10 scale, a similar correlation was observed. The patients with the largest changes in the MRDSS were patients with no T1 lesion volume at baseline. To address this limitation, we adjusted the z-score for these patients compared to our original paper, but the impact of these patients remained. When the alternative transformation for the T1/T2 (arc-sine square root) was investigated, the correlation between baseline MRDSS and change in MRDSS was reduced, but it remained statistically significant. Future work should involve considering other statistical transformations of the T1 lesion volume that may further reduce the impact of patients who change from no T1 lesion volume.

We also evaluated the relationship between clinical and MRI data. The pattern of MRDSS–clinical relationships at exit in the current study is consistent with our previously reported baseline data suggesting that MRDSS has good internal consistency and concurrent validity [7]. On the other hand, we found some unexpected results in terms of its predictive validity for disability progression. Our previous study found that baseline MRDSS was associated with the risk of developing sustained progression of physical disability three years later [7]. However, this association was not statistically significant in our reduced sample. In addition, the on-study change in MRDSS was also not related to clinical progression. Thus, it seems that our scale has an unreliable strength for predicting clinical change in terms of disability in the current form and needs further refinements. Similar to MRDSS, individual MRI measures also showed unreliable strength for predicting clinical change. This is in agreement with prior studies demonstrating that these measures show relatively weak correlations

with clinical progression [1,17–19]. Of note, we found an inverse correlation between baseline T2LV and a higher risk of progressing on EDSS, driven largely by the SPMS/PPMS patients. One potential explanation for this observation is that SPMS/PPMS patients with low T2 lesion volume in our sample may have lesions in clinical eloquent areas such that even though the total volume is low, the impact on clinical features may be high. Of note, there is a well known dissociation between T2 lesion volume and clinical findings [1,18–21], which is not entirely surprising given that T2 lesion volume is not completely representative of all features of disease progression, such as neurodegeneration. Nonetheless, given the lack of association with change in T2 lesion volume, our results must be considered preliminary and require validation in larger samples. The tendency for SP and PP patients with lower MRI disease burden to progress was also seen with the MRDSS and T1/T2 measures.

Several other limitations were also evident in our study. A more even balance of primary and secondary progressive patients with a larger sample size would afford the opportunity to continue to evaluate the MRDSS across a wide spectrum of disease states. This would strengthen any conclusions reached regarding its utility in the general MS population as well as enabling it to be used as a potential marker in clinical trials with smaller sample sizes. Additional measurements detailing diffuse damage in the normal appearing brain tissue (such as magnetization transfer, diffusion imaging, or MR spectroscopy) [18,19,21], would help refine the scale and potentially allow for greater predictability even earlier in the disease course prior to accumulation of atrophy. The inclusion of gray matter atrophy and spinal cord damage may also be a very informative addition to our scale, given that involvement of these areas of the CNS are common and clinically relevant in MS [22,23]. Furthermore, the use of 3T MRI [19] may also increase the sensitivity of structural changes detected and may strengthen the validity of the composite scale. We are currently collecting data to refine the MRDSS scale along these lines, which we will present in future publications.

In further considering directions for future studies, investigating correlation of the MRDSS with other disability and quality of life variables, which are not heavily represented in the EDSS score, such as cognition and fatigue, would help to better understand the validity of the scale. Unfortunately, quality of life measures, including cognitive and fatigue scores, were not available for these cases. Assessment of the MRDSS over shorter (less than 2 years) and longer (greater than 6 years) periods of time would be useful to determine the dynamic longitudinal range of the scale. It would be particularly interesting to test if the MRDSS can predict the conversion from RR to SP MS in individual cases.

In conclusion, after defining the MRDSS scale in our original work and obtaining some information on its validity [7], we now have shown the longitudinal characteristics of the scale, including the sensitivity to change over time. Our study suggests that the composite MRDSS scale may be able to capture the destructive aspects of the disease with more longitudinal sensitivity than derived from conventional MRI lesion and whole brain atrophy measurements. Further studies are warranted to confirm and extend our findings regarding the potential utility of the MRDSS.

Conflict of interest

The authors have no conflict of interest.

Acknowledgements

This study was supported by the National Institutes of Health (1R01NS055083-01) and National Multiple Sclerosis Society (RG3705A1; RG3798A2). We thank Ms. Sophie Tamm for assistance with manuscript preparation. These findings were presented in preliminary form at the 1st joint meeting of ACTRIMS (the Americas Committee on Treatment and Research in Multiple Sclerosis) and its

counterparts in Europe and Latin America: ECTRIMS and LACTRIMS, Montreal, September 17–20, 2008 and the 62nd annual meeting of the American Academy of Neurology, Toronto, Canada, April 10–17, 2010.

References

- [1] Zivadinov R, Leist TP. Clinical–MRI correlations in multiple sclerosis. *J Neuroimaging* 2005;15:10S–21S.
- [2] Wolinsky JS, Narayana PA, Noseworthy JH, Lublin FD, Whitaker JN, Linde A, et al. Linomide in relapsing and secondary progressive MS: Part II: MRI results. MRI Analysis Center of the University of Texas – Houston, Health Science Center, and the North American Linomide Investigators. *Neurology* 2000;54:1734–41.
- [3] Mainero C, De Stefano N, Iannucci G, Sormani MP, Guidi L, Federico A, et al. Correlates of MS disability assessed in vivo using aggregates of MR quantities. *Neurology* 2001;56:1331–4.
- [4] Wolinsky JS, Narayana PA, Johnson KP. Multiple Sclerosis Study Group and the MRI Analysis Center. United States open-label glatiramer acetate extension trial for relapsing multiple sclerosis: MRI and clinical correlates. *Mult Scler* 2001;7:33–41.
- [5] Sriram S, Yao SY, Stratton C, Moses H, Narayana PA, Wolinsky JS. Pilot study to examine the effect of antibiotic therapy on MRI outcomes in PRMS. *J Neurol Sci* 2005;20:1–5.
- [6] Poonawalla AH, Datta S, Juneja V, Nelson F, Wolinsky JS, Cutter G, et al. Composite MRI scores improve correlation with EDSS in multiple sclerosis. *Mult Scler* 2010;16:1117–25.
- [7] Bakshi R, Neema M, Healy BC, Liptak Z, Betensky RA, Buckle GJ, et al. Predicting clinical progression in multiple sclerosis with the magnetic resonance disease severity scale. *Arch Neurol* 2008;65:1449–53.
- [8] Gauthier SA, Glanz BI, Mandel M, Weiner HL. A model for the comprehensive investigation of a chronic autoimmune disease: the multiple sclerosis CLIMB study. *Autoimmun Rev* 2006;5:532–6.
- [9] Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* 1983;33:1444–52.
- [10] Polman CH, Reingold SC, Edan G, Filippi M, Hartung HP, Kappos L, et al. Diagnostic criteria for multiple sclerosis: 2005 revisions to the “McDonald Criteria”. *Ann Neurol* 2005;58:840–6.
- [11] Lublin FD, Reingold SC, for the National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. Defining the clinical course of multiple sclerosis: results of an international survey. *Neurology* 1996;46:907–11.
- [12] Wei X, Warfield S, Zou KH, Wu Y, Li X, Guimond A, et al. Quantitative analysis of MRI signal abnormalities of brain white matter with high reproducibility and accuracy. *J Magn Reson Imaging* 2002;15:203–9.
- [13] Pelletier D, Garrison K, Henry R. Measurement of whole-brain atrophy in multiple sclerosis. *J Neuroimaging* 2004;14:11S–9S.
- [14] Bermel RA, Sharma J, Tjoa CW, Puli SR, Bakshi R. A semiautomated measure of whole-brain atrophy in multiple sclerosis. *J Neurol Sci* 2003;208:57–65.
- [15] Bagnato F, Jeffries N, Richert ND, Stone RD, Ohayon JM, McFarland HF, et al. Evolution of T1 black holes in patients with multiple sclerosis imaged monthly for 4 years. *Brain* 2003;126:1782–9.
- [16] Cutter GR, Baier ML, Rudick RA, Cookfair DL, Fischer JS, Petkau J, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. *Brain* 1999;122:871–82.
- [17] Bermel RA, Bakshi R. The measurement and clinical relevance of brain atrophy in multiple sclerosis. *Lancet Neurol* 2006;5:158–70.
- [18] Neema M, Stankiewicz J, Arora A, Guss ZD, Bakshi R. MRI in multiple sclerosis: what's inside the toolbox? *Neurotherapeutics* 2007;4:602–17.
- [19] Bakshi R, Thompson AJ, Rocca MA, Pelletier D, Dousset V, Barkhof F, et al. MRI in multiple sclerosis: current status and future prospects. *Lancet Neurol* 2008;7:615–25.
- [20] Miki Y, Grossman R, Udupa JK, Wei L, Polansky M, Mannon LJ, et al. Relapsing–remitting multiple sclerosis: longitudinal analysis of MR images – lack of correlation between changes in T2 lesion volume and clinical findings. *Radiology* 1999;213:395–9.
- [21] Filippi M, Rocca MA, Rovaris M. Clinical trials and clinical practice in multiple sclerosis: conventional and emerging magnetic resonance imaging technologies. *Curr Neurol Neurosci Rep* 2002;2:267–76.
- [22] Pirko I, Lucchinetti CF, Sriram S, Bakshi R. Gray matter involvement in multiple sclerosis. *Neurology* 2007;68:634–64.
- [23] Horsfield MA, Sala S, Neema M, Absinta M, Bakshi A, Sormani MP, et al. Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: application in multiple sclerosis. *Neuroimage* 2010;50:446–55.