

Tissue damage detected by quantitative gradient echo MRI correlates with clinical progression in non-relapsing progressive MS

Biao Xiang, PhD,^{1†} Matthew R. Brier, MD, PhD,^{2†} Manasa Kanthamneni, BS,^{1,3} Jie Wen, PhD,¹ Abraham Z. Snyder, MD, PhD,^{1,2} Dmitriy A. Yablonskiy, PhD,¹ and Anne H. Cross, MD^{2*}

1. Department of Radiology, Washington University, St. Louis, MO, 63110, USA
2. Department of Neurology, Washington University, St. Louis, MO, 63110, USA
3. School of Medicine, St. George's University, St. George, Grenada

† These authors contributed equally to this study.

* Corresponding Author:

Anne H. Cross

Department of Neurology

Washington University

660 South Euclid, Campus Box 8111

St. Louis, MO 63110, USA

Email: crossa@neuro.wustl.edu

Tel: +1(314)362-3293

Fax: +1(314)747-1345

Abstract

Background:

Imaging biomarkers of progressive MS are needed. Quantitative gradient recalled echo (qGRE) MRI technique allows evaluation of tissue damage associated with microstructural damage in multiple sclerosis (MS).

Objective:

To evaluate qGRE-derived R2t* as an imaging biomarker of MS disease progression as compared to atrophy and lesion burden.

Methods:

Twenty-three non-relapsing progressive MS (PMS), twenty-two relapsing-remitting MS (RRMS) and eighteen healthy control participants were imaged with qGRE at 3T. PMS subjects were imaged and neurologically assessed every nine months over five sessions. In each imaging session, lesion burden, atrophy and R2t* in cortical grey matter (GM), deep GM, normal-appearing white matter (NAWM) were measured.

Results:

R2t* reductions correlated with neurological impairment cross-sectionally and longitudinally. PMS patients with clinically defined disease progression showed significantly faster decrease of R2t* in NAWM and deep GM compared with the clinically stable PMS group. Importantly, tissue damage measured by R2t* outperformed lesion burden and atrophy as a biomarker of progression during the study period.

Conclusion:

Clinical impairment and progression correlated with accumulating R2t*-defined microstructural tissue damage in deep GM and NAWM. qGRE-derived R2t* is a potential imaging biomarker of MS progression.

1. Introduction

Magnetic resonance imaging (MRI) plays an important role in multiple sclerosis (MS) diagnosis and monitoring ¹⁻³. Quantification of new or gadolinium enhancing white matter (WM) lesions (WMLs) by means of T1- or T2-weighted clinical images represents disease activity and serves as an end-point for trials of disease-modifying therapies ⁴. However, this approach to disease monitoring applies primarily to relapsing remitting MS (RRMS). Progressive MS (PMS) subtypes [non-relapsing secondary progressive (SP) MS and primary progressive (PP) MS] lack highly sensitive imaging biomarkers of progression due in part to the less inflammatory pathology of PMS ^{5,6}. The best available biomarker of PMS currently is atrophy, especially of grey matter (GM) ⁷. Importantly, atrophy represents the end-result of multiple pathological processes. More specific biomarkers which measure pathology prior to the manifestation of frank atrophy are needed ⁸.

Here, we assess quantitative gradient recalled echo (qGRE) imaging to measure $R2t^*$ ($=1/T2t^*$) as a potential biomarker of progression in MS. $R2t^*$ is a component of the more commonly reported $R2^*$ ($=1/T2^*$) measure and is more specific to tissue microstructure. This study addresses the ability of $R2t^*$ to reflect clinical progression and compares the utility of $R2t^*$ to atrophy and lesion burden imaging biomarkers. We compare these measures in PMS, RRMS and age-matched healthy control subjects and longitudinally in relation to advancing impairment in PMS patients.

2. Materials and methods

2.1 Study Participants

Eighteen healthy control (HC), twenty-three PMS (including four PPMS and nineteen SPMS), and twenty-two RRMS subjects were enrolled after providing informed consent. The study was approved by the Institutional Review Board. PMS patients did not have superimposed relapses or gadolinium enhancing lesions within twenty-four months preceding enrollment. RRMS and HC were recruited to reflect the age and sex distribution of the PMS patients. RRMS subjects had no clinical worsening except related to MS attacks. The HC and RRMS subjects were scanned one time at the beginning of the study. PMS subjects were scanned every nine months for five sessions with the goal of identifying imaging markers that correlated best with or predicted clinical progression in PMS patients.

One PMS patient was lost to follow up after the second session, two following the third session, and one following the fourth session. Six imaging sessions were excluded due to severe motion artifacts and six sessions due to technical difficulties. In total, ninety-two sessions (out of 115 possible) in the PMS cohort were used for data analysis.

2.2 Clinical testing

Clinical evaluation was performed on the day of MRI exam and included the Expanded Disability Status Scale (EDSS) standardized neurological examination, 25-foot timed walk (25FTW) assessment of gait, nine-hole peg test (9HPT) assessment of bilateral upper extremity function, paced auditory serial addition test (PASAT) and symbol digit modalities test (SDMT) assessments

of cognitive function. Clinical assessment was performed by examiners blinded to the imaging results. For data analyses, the 25FTW and 9HPT raw scores were converted to Z-scores according to the Multiple Sclerosis Functional Composite (MSFC) guidelines ⁹.

2.3 Definition of clinical progression during the study interval

The PMS subjects were classified as either clinically stable or progressive during the study period based on EDSS, 25FTW and 9HPT scores ¹⁰. EDSS score increases by one point were defined as progression if the subject's baseline EDSS was less than 5.5; for baseline EDSS 5.5 or greater, one-half point increases defined progression. For the 25FTW score, $\geq 20\%$ increase or converting from walking to not walking defined progression. For the 9HPT score, increasing by $\geq 20\%$ or converting from able to unable to complete the test defined progression. Those PMS subjects who did not show progression in EDSS, 25FTW or 9HPT were defined as clinically stable.

2.4 Image acquisition

All MRI was performed with a 3.0 Tesla Trio MRI scanner (Siemens, Erlangen, Germany) using a 32-channel phased-array radiofrequency head coil. qGRE data were acquired using three-dimensional multi-gradient-echo sequence with flip angle 30° , TR=50ms, voxel size $1 \times 1 \times 2 \text{ mm}^3$ and acquisition time 12 minutes. 10 gradient echoes, with first echo time TE1=4ms and echo spacing $\Delta \text{TE}=4\text{ms}$ were collected. In addition, fluid-attenuated inversion recovery (FLAIR) images (voxel size $1 \times 1 \times 3 \text{ mm}^3$) and magnetization-prepared rapid gradient-echo (MP-RAGE) (voxel size $1 \times 1 \times 1 \text{ mm}^3$) were acquired for purposes of atlas registration and tissue segmentation.

2.5 Image processing and segmentation

R2t* maps were created using the qGRE approach described in a previous study ¹¹. In brief, multi-channel data were combined using a previously published algorithm ¹². Voxel-wise analysis was performed on combined data using the theoretical model of GRE signal relaxation ¹³ and a set of post-processing algorithms that minimize adverse artifacts related to macroscopic magnetic field inhomogeneities ¹⁴ and physiological fluctuations ¹⁵ were employed.

Image segmentation was achieved via a two-step process. First, MP-RAGE and FLAIR images were subjected to FreeSurfer analysis (version 6.0) ¹⁶ and segmentations were reviewed and errors corrected by a trained technician (MK). FreeSurfer does not accurately segment WMLs. Hence, a second approach included application of an intensity-based approach which we recently developed and validated ¹⁷. For subsequent analyses, voxels were labeled as CSF or GM if both segmentation strategies agreed. Cortical and deep GM were separated based on FreeSurfer segmentation¹. Normal Appearing White Matter (NAWM) was defined as voxels in which both segmentations returned a WM label. Lesions were defined as voxels identified as lesions by both approaches or where FreeSurfer labeled the voxel as WM and the intensity-based approach labeled the voxel as lesion.

Mean R2t* in cortical grey matter (GM), deep GM, normal appearing white matter ($R2t_{NAWM}^*$), and individual lesions ($R2t_{Les}^*$) were computed for each subject. In MS patients, to quantify the

¹ Deep GM included the basal ganglia (caudate, putamen, globus pallidus), and thalamic nuclei.

damage specific to WM lesions, controlling for the damage in NAWM, we defined the Tissue Damage Score (TDS) in each WM lesion as

$$TDS = \frac{R2t_{NAWM}^* - R2t_{Les}^*}{R2t_{NAWM}^*}$$

TDS measured in this way quantifies the damage within lesions relative to the NAWM which, in MS patients, may also be abnormal ¹⁷.

2.6 Statistical analysis

Longitudinal analysis of clinical test performance and imaging biomarkers in PMS subjects was conducted via ANOVA with subject as repeated, within-subject factor; the Dunnett test was used to correct for multiple comparison. Baseline clinical and imaging data were compared between subject groups using a one-way ANOVA with group (HC, RRMS, PMS) as factor; the Tukey test was used to correct for multiple comparison.

Spearman rank correlation tests were performed to assess the relationship between imaging biomarkers and clinical test scores. Imaging biomarkers included lesion volume, R2t* measured in cortical and deep GM, NAWM, and TDS. Clinical test performance measures included EDSS, 25FTW, 9HPT, PASAT and SDMT. This cross-sectional analysis used only the first imaging session in the PMS patients. Age and gender were controlled as covariates. False discovery rate was used to correct for multiple comparisons ¹⁸.

Linear fitting was used to calculate the rate of change of imaging biomarkers in each PMS patient across available visits. Two-sample t-tests were used to compare the rate of change and baseline measurements for stable PMS versus PMS with clinically defined progression.

A major objective of the present work is to isolate imaging biomarkers that predict clinical progression. However, these biomarkers are characterized by high levels of shared variance. Accordingly, to isolate the most predictive biomarkers, LASSO regression was applied to identify the most salient imaging biomarkers of clinical progression ¹⁹. LASSO is a modification to linear regression wherein regression coefficients (β) values are penalized such that small or redundant β values are forced to zero. Thus, LASSO models are sparse: several β values are large while many are identically 0 indicating those variables do not contribute to the model. Sparsity is controlled by a tuning variable, λ , chosen by cross-validation.

3. Results

Table 1. Demographic and clinical information of healthy controls and multiple sclerosis subjects.

		PMS	RRMS	Healthy Control
Number of Subjects		23	22	18
Number of imaging sessions		92	22	18
mean Age $\pm$ SD (years) (range)		56.8 $\pm$ 8.6 (31-75)	54.9 $\pm$ 10.1 (22-72)	48.8 $\pm$ 14.5 (23-76)
Female/Male		15/8	18/4	13/5
EDSS mean $\pm$ SD (range)		6.0 $\pm$ 1.5 (2.5-8.5)	2.4 $\pm$ 1.1 (1-5)	N/A
25FTW mean $\pm$ SD (second) (range)		46.1 $\pm$ 63.7 (4.2-165.8)	4.5 $\pm$ 1.1 (2.7-7.7)	N/A
9HPT mean $\pm$ SD (second) (range)	Dominant	116.1 $\pm$ 231.3 (18.7-777)	22.8 $\pm$ 6.2 (15.2-45.5)	N/A
	Non-dominant	127.4 $\pm$ 251.1 (21.4-777)	22.0 $\pm$ 3.6 (16.3-31.8)	N/A
SDMT mean $\pm$ SD (range)		41.5 $\pm$ 14.4 (10-63)	54.7 $\pm$ 8.2 (37-80)	N/A
3 sec PASAT mean $\pm$ SD (range)		43.5 $\pm$ 13.4 (14-60)	44.9 $\pm$ 10.9 (27-58)	N/A

Note: The clinical tests scores of PMS subjects were from session one. Tests scores of subsequent sessions are shown in figure 3.

Demographic information and clinical test performance are summarized in Table 1. Age was well matched between cohorts. The RRMS patients were chosen to match the PMS patients, and thus had an average age of 54.9 years with a range spanning 5 decades. PMS subjects had significantly higher baseline EDSS, 25FTW and 9HPT (dominant and non-dominant) scores compared to RRMS subjects ($p < 0.001$, $p < 0.01$, $p < 0.001$, $p < 0.001$, respectively). The RRMS cohort performed significantly better on SDMT than the PMS cohort ($p < 0.001$). No differences were found in PASAT performance.

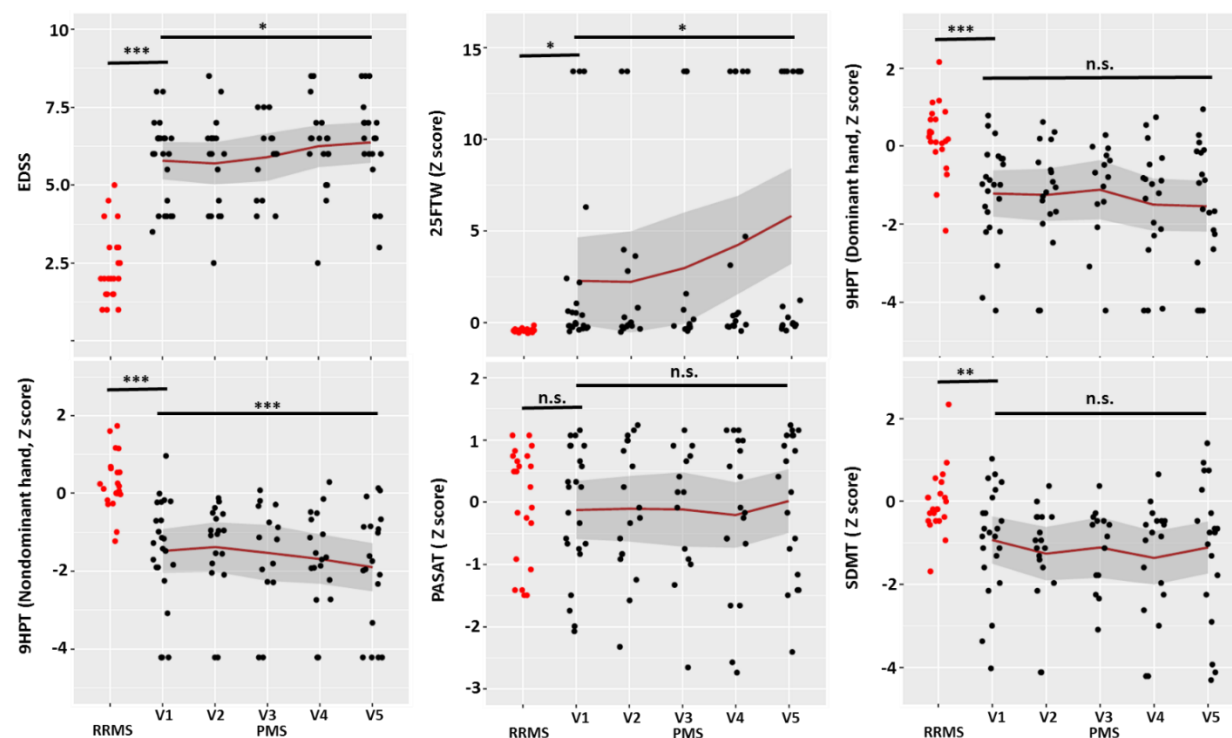


Figure 1. Clinical assessments of the PMS cohort over four years. Compared with the PMS cohort at baseline, the RRMS cohort (red dots) showed significantly better clinical test scores. EDSS, 25-foot walk and 9-hole peg test showed an increasing trend of disability and motor dysfunction for the longitudinal PMS cohort over five imaging sessions, suggesting MS progression within that group. For EDSS, 25FTW and 9HPT (non-dominant hand), there were significant differences between scores in visit five and visit one ($p < 0.05$, $p < 0.05$, $p < 0.001$, respectively). 25FTW, 9HPT, PASAT and SDMT were converted to Z scores. Each dot represents one subject. The gray area represents the standard error. One-way ANOVA with repetitive measurements was used to compare the difference between visit one and each subsequent visit of the PMS group. Dunnett test was used to correct for multiple comparisons. Two-sample t-test was used to compare the RRMS cohort to the first visit of PMS group. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, n.s. $p > 0.05$.

The PMS group was assessed longitudinally for changes in clinical test scores (Figure 1). Within the PMS group, Visit 5 showed significantly higher (more impaired) scores in EDSS, 25FTW and 9HPT than baseline ($p < 0.05$, $p < 0.05$, $p < 0.001$, respectively). In contrast, cognitive function test scores (PASAT and SDMT) remained stable over the five sessions ($p > 0.05$). Pairwise statistical comparisons for each test and visit are shown in supplemental Table S1. We also compared the performance on clinical tests between the RRMS and PMS group (Figure 1). RRMS compared to the baseline visit of the PMS cohort showed significant differences group-wise in EDSS, 25-foot walk, 9-hole peg test and SDMT ($p < 0.001$, $p < 0.05$, $p < 0.001$, $p < 0.001$, $p < 0.01$, respectively).

	R2t* (Cortical GM)/s ⁻¹		R2t* (Deep GM)/s ⁻¹		R2t* (NAWM)/s ⁻¹		Mean Lesion TDS		Lesion Volume/mm ³	
	rho	p	rho	p	rho	p	rho	p	rho	p
EDSS	-0.722	p<0.001	-0.153	0.223	-0.498	p<0.001	0.431	p<0.001	0.199	0.091
25FTW	-0.748	p<0.001	-0.246	0.034	-0.531	p<0.001	0.481	p<0.001	0.141	0.254
9HPT (Dominant)	0.453	p<0.001	-0.041	0.722	0.330	0.003	-0.288	0.011	-0.058	0.645
9HPT (NonDominant)	0.562	p<0.001	-0.028	0.786	0.422	p<0.001	-0.339	0.002	-0.119	0.315
PASAT	0.296	0.009	-0.119	0.315	0.335	0.003	0.087	0.481	-0.201	0.091
SDMT	0.448	p<0.001	0.051	0.702	0.390	p<0.001	-0.213	0.067	-0.128	0.302

Table 2. R2t* in the cortical GM, NAWM and lesions correlated with clinical assessments. R2t* measurements in both cortical GM and NAWM correlated with motor and cognitive tests. In addition, R2t* in cortical GM showed stronger correlation with clinical tests than NAWM. Mean lesion tissue damage score (TDS) computed using R2t* measures within lesions showed significant correlations with the EDSS and motor assessments. No significant correlation between lesion volume and clinical scores was found. Spearman rho and p values were computed in R, with age and gender as covariates. All listed p values are after multiple comparison correction using false discovery rate. Statistically significant correlations are highlighted in color.

We next investigated the relation between the R2t* of different brain tissues and clinical test performance. In the cross-sectional analysis of forty-five MS subjects (22 RRMS and 23 PMS), R2t* measurements in NAWM and cortical GM, but not deep GM, correlated significantly with the EDSS, and motor and cognitive test scores (Table 2). The association between most clinical scores and R2t* was stronger in cortical GM compared to NAWM. R2t* in WMLs, quantified relative to NAWM within each subject as the TDS, showed significant correlations with the EDSS and motor-related clinical scores (25FTW, 9HPT), but not cognitive test scores (PASAT, SDMT). Importantly, lesion volume did not show significant correlation with any clinical scores.

R2t* was measured in cortical and deep GM, NAWM and lesions (as TDS) in the RRMS and PMS groups cross-sectionally and the PMS group longitudinally (Figure 2). At baseline, cortical GM R2t* was lower in the PMS cohort compared to RRMS and control group (p<0.05). Lesion TDS showed larger (worse) values in PMS compared to RRMS. In PMS, both cortical and deep GM R2t* declined over the five imaging sessions. R2t* in NAWM did not change across visits, However, the data exhibited increasing variability at later visits, suggesting that R2t* in NAWM declines more in some subjects than others. Pairwise comparisons for each measure and visit are shown in the supporting information (Table S2).

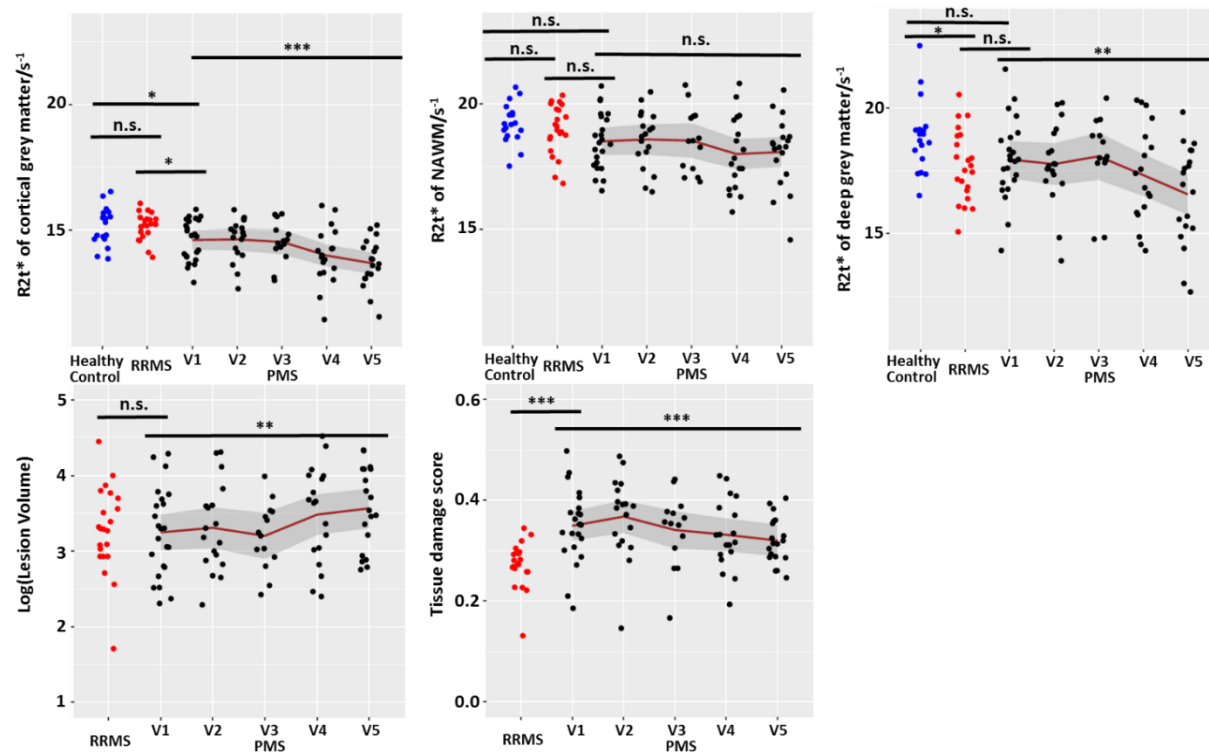


Figure 2. Longitudinal R2t*, lesion volume and tissue damage score measurements for non-relapsing PMS subjects compared with healthy control and RRMS measurements. The longitudinal PMS cohort showed significant decreasing of R2t* (greater microstructural tissue damage) in cortical GM and deep GM over five imaging sessions, performed nine months apart. R2t* of cortical GM and deep GM in visit five were significantly different from R2t* in visit one ($p < 0.001$ and $p < 0.01$, respectively). The R2t* measurements in GM of PMS subjects at baseline were significantly lower than healthy control and RRMS cohorts ($p < 0.05$). Lesion volume of PMS subjects increased over four years. Specifically, lesion volume in visit five was significantly higher than the lesion volume in visit one ($p < 0.01$). Tissue damage score declined over the last four visits, which is likely due to the decrease of R2t* in NAWM, making the differences between R2t* in NAWM and lesion volume smaller. The line represents the mean value and the gray area represents the standard error. One-way ANOVA with repetitive measurements was used to compare the difference between visit one and each subsequent visit of the PMS group. Dunnett test was used to correct for multiple comparisons. One-way ANOVA was used to compare the healthy control, RRMS and baseline visit one of the PMS group. Tukey test was used to correct for multiple comparisons. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, n.s. $p > 0.05$.

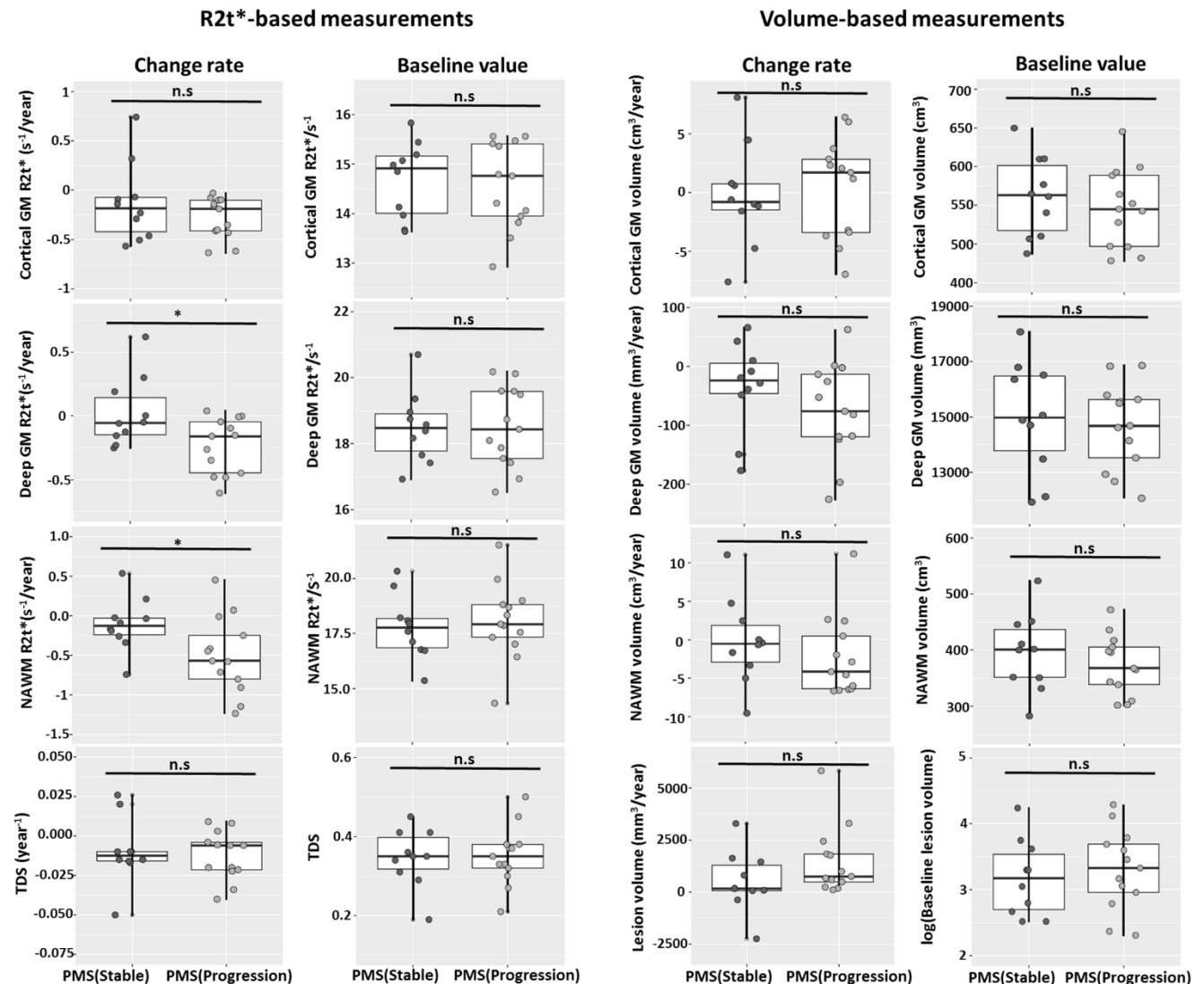


Figure 3. Rate of R2t* decline in NAWM and deep grey matter was faster in PMS subjects with clinically defined progression than in stable PMS subjects. Thirteen PMS subjects showed clinically defined progression, whereas ten subjects were stable on clinical tests. Each dot represents one subject. Each row corresponds to an imaging biomarker. The first and third columns of graphs show rate of change and the second and fourth columns show the baseline values. Rate of change of R2t* in deep grey matter and NAWM were significantly faster in PMS with clinical progression, whereas rate of change in lesion tissue damage score, lesion volume, NAWM volume and deep GM atrophy were nominally greater in the PMS group with clinical progression group than the stable group, but did not reach statistical significance. * $p < 0.05$, n.s $p > 0.05$.

To determine if imaging biomarkers distinguished stable ($n=10$) vs. clinically progressing ($n=13$) non-relapsing PMS patients, we calculated the rate of change in R2t*, lesion volume, and atrophy. Two-sample t-tests were used to compare the rates of change between the two groups. The baseline measurements of lesion volume, atrophy and R2t* were not statistically different between groups (Figure 3). PMS subjects with clinically defined progression showed faster decline of R2t* in NAWM and deep grey matter. These statistical tests were significant considered in isolation but not significant when adjusted for multiple comparisons. This result suggests that patients with

clinical progression, while not different from stable subjects at baseline, had more rapid tissue damage accumulation in NAWM and deep GM over the study period, detected by R2t* but not by volume loss.

Logistic LASSO was used to identify the unique contribution of changes in R2t* in NAWM and deep GM to the prediction of clinical progression. Predictors included in the model were initial and rate of change in R2t*, lesion volume, and atrophy (Table 3). Imaging data were standardized as Z-scores, resulting in regression coefficients (β values) that are comparable across variables. Baseline imaging values were not strongly associated with later progression. However, change in deep GM and NAWM R2t* were associated with clinical progression. The largest β values were observed in the rate of change in R2t* of deep GM, followed by NAWM and lesion volume. This result is compatible with R2t* of deep GM and NAWM contributing additional predictive information apart from atrophy and lesion burden.

<u>Predictor</u>	<u>β value</u>
Initial Value	
R2t* - CGM	0
R2t* - DGM	0
R2t* - NAWM	0
R2t* - Lesion TDS	0
Lesion Volume	0
Volume – CGM	0
Volume – DGM	0
Volume - NAWM	0
Rate of Change	
R2t* - CGM	0
R2t* - DGM	-0.4299
R2t* - NAWM	-0.0211
R2t* - Lesion TDS	0
Lesion Volume	0.0046
Volume – CGM	0
Volume – DGM	0
Volume - NAWM	0

Table 3: Imaging variables that correlate with clinical progression. LASSO logistic regression was fit with the 10 imaging parameters as predictors and the group (clinically stable PMS vs. clinical progression PMS) as outcome. Input variables were normalized to Z scores such that values exist on the same scale. $\lambda=0.1375$, the tuning parameter, was optimized by cross validation. CGM = cortical grey matter; DGM = deep grey matter; NAWM = normal appearing white matter; TDS= tissue damage score.

4. Discussion

MRI has significantly contributed to our understanding of MS pathology. WMLs, detected on gadolinium enhanced T1- and T2-weighted sequences, reliably correlate with clinical relapses in RRMS ⁴. Based on this, new or enhancing lesions are included as outcome measures in ongoing

clinical trials and represent disease activity potentially due to suboptimal treatment of RRMS²⁰. However, no currently available imaging biomarker, apart from atrophy, similarly correlates with progression in PMS. Because atrophy occurs subsequent to damage, identification of progression by atrophy is possible only in retrospect. This has, at least partially, slowed down development of treatments of PMS. The current results demonstrate that R2t* strongly correlates with clinical impairment more closely than atrophy, appearing to concurrently reflect clinical progression.

qGRE quantitatively measures R2t* ($=1/T2t^*$)¹³, which is a sub-component of total R2* GRE signal decay ($R2^*=1/T2^*$)¹¹. Previous studies showed that quantitative determination of R2* differentiates RRMS patients from healthy controls and correlates with neurological impairment^{12, 21, 22}. R2t* is theoretically more specific to tissue microstructure than total R2* because it measures a part of the GRE signal decay (transverse relaxation rate) that depends strongly on the local cellular-matrix environment of water molecules (a major source of MRI signal). R2* is affected by macroscopic field inhomogeneities, physiological fluctuations and the vascular-related blood oxygen level-dependent (BOLD) effect whereas R2t* is less affected¹³⁻¹⁵. Thus, R2t* is more closely related to intrinsic tissue integrity than R2* and is a more biologically relevant biomarker. Prior results also showed significant associations between CNS R2t* reduction and MS clinical severity^{23, 24}. Importantly, the present results expand those findings to non-relapsing clinical progression examined longitudinally and demonstrate that R2t*-based measures outperform atrophy- and lesion-based measures as clinical biomarkers.

Histopathological studies have demonstrated that MS tissue damage is complex, involving not only demyelination but also axonal injury and transections, and gliosis in WMLs, GM and NAWM²⁵. Progressive MS is thought to be due to continuing tissue injury mediated by multiple mechanisms beyond inflammatory activity²⁶. We assayed intrinsic tissue damage using R2t* as a marker of pathology in patients with RRMS and PMS. We showed that PMS patients had greater tissue injury (reduced R2t*) compared to similarly aged RRMS patients and that this injury correlated closely with clinical test performance. Consistent with ongoing damage, we found longitudinal reductions in R2t* across much of the brain in PMS patients. PMS patients were further stratified as progressing or stable over the study period. Almost all PMS patients demonstrated at least mildly worsened (but not always statistically significant) clinical test performance across the study period. Importantly, patients with confirmed progression, indicating significantly worsened impairment, had the most dramatic declines in tissue integrity in deep GM and NAWM. The association of deep GM atrophy with MS-related disability, specifically in progressive disease, is well documented^{27, 28} but tissue atrophy is likely the end result of multiple pathologies. The present results suggest that tissue integrity as measured by R2t* reflects neuropathology not yet manifesting as atrophy.

This study has a few limitations. The cohort was relatively small (23 PMS patients). Nevertheless, we detected differential changes in those who progressed versus those who did not, suggesting that R2t* might serve as a surrogate for progression in early phase clinical trials in PMS. The progressive patients were followed every nine months for five visits, a short time compared to the time-course of PMS. Some patients may have clinically progressed but remained below the detection threshold. In addition, learning effects might have masked cognitive decrements on

repeated testing. We were able to match RRMS and PMS cohorts by age but matching the two subtypes by both age and impairment was not possible due to the near-definitional requirement for PMS to have more significant clinical impairment. The histological basis of R2t* signal changes in the deep GM is uncertain. Measured R2t* reflects the combined effects of microstructural tissue damage and non-heme iron deposition. This question will be a subject of our future studies employing a method for separation of cellular and iron contributions to GRE signal recently developed by our group ²⁹.

5. Conclusion

We measured qGRE R2t* as a biomarker for tissue microstructural damage in brains of RRMS and PMS patients. We examined longitudinal changes in R2t* in non-relapsing PMS patients with and without clinically confirmed progression over the study period. We found that decreased R2t* in deep GM and NAWM correlated with clinical progression. This finding supports the potential of R2t* as a quantitative imaging biomarker of ongoing MS progression.

References

1. McDonald WI, Compston A, Edan G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Annals of neurology* 2001; 50: 121-127. Guideline

Research Support, Non-U.S. Gov't 2001/07/18.

2. Dalton CM, Brex PA, Miszkil KA, et al. New T2 lesions enable an earlier diagnosis of multiple sclerosis in clinically isolated syndromes. *Annals of neurology* 2003; 53: 673-676. Research Support, Non-U.S. Gov't 2003/05/06. DOI: 10.1002/ana.10580.

3. Brex PA, Ciccarelli O, O'Riordan JI, et al. A longitudinal study of abnormalities on MRI and disability from multiple sclerosis. *The New England journal of medicine* 2002; 346: 158-164. Research Support, Non-U.S. Gov't 2002/01/18. DOI: 10.1056/NEJMoa011341.

4. Sormani MP and Bruzzi P. MRI lesions as a surrogate for relapses in multiple sclerosis: a meta-analysis of randomised trials. *The Lancet Neurology* 2013; 12: 669-676. DOI: [https://doi.org/10.1016/S1474-4422\(13\)70103-0](https://doi.org/10.1016/S1474-4422(13)70103-0).

5. Lassmann H, van Horssen J and Mahad D. Progressive multiple sclerosis: pathology and pathogenesis. *Nat Rev Neurol* 2012; 8: 647-656. 2012/09/26. DOI: 10.1038/nrneurol.2012.168.

6. Dutta R and Trapp BD. Relapsing and progressive forms of multiple sclerosis: insights from pathology. *Current opinion in neurology* 2014; 27: 271-278. Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, Non-P.H.S.

Review 2014/04/12. DOI: 10.1097/WCO.0000000000000094.

7. Moccia M, de Stefano N and Barkhof F. Imaging outcome measures for progressive multiple sclerosis trials. *Multiple Sclerosis Journal* 2017; 23: 1614-1626. DOI: 10.1177/1352458517729456.

8. Oh J, Ontaneda D, Azevedo C, et al. Imaging outcome measures of neuroprotection and repair in MS: A consensus statement from NAIMS. *Neurology* 2019; 92: 519-533. DOI: 10.1212/wnl.00000000000007099.

9. Fischer JS, Rudick RA, Cutter GR, et al. The Multiple Sclerosis Functional Composite measure (MSFC): an integrated approach to MS clinical outcome assessment. *Multiple Sclerosis Journal* 1999; 5: 244-250. DOI: 10.1177/135245859900500409.

10. Fox RJ, Coffey CS, Conwit R, et al. Phase 2 Trial of Ibudilast in Progressive Multiple Sclerosis. *New England Journal of Medicine* 2018; 379: 846-855. DOI: 10.1056/NEJMoa1803583.

11. Wen J, Goyal MS, Astafiev SV, et al. Genetically defined cellular correlates of the baseline brain MRI signal. *Proceedings of the National Academy of Sciences* 2018; 115: E9727-E9736. DOI: 10.1073/pnas.1808121115 %J Proceedings of the National Academy of Sciences.

12. Luo J, Jagadeesan BD, Cross AH, et al. Gradient echo plural contrast imaging--signal model and derived contrasts: T2*, T1, phase, SWI, T1f, FST2*and T2*-SWI. *Neuroimage* 2012; 60: 1073-1082. 2012/02/07. DOI: 10.1016/j.neuroimage.2012.01.108.

13. Ulrich X and Yablonskiy DA. Separation of cellular and BOLD contributions to T2* signal relaxation. *Magnetic Resonance in Medicine* 2016; 75: 606-615. DOI: 10.1002/mrm.25610.

14. Yablonskiy DA, Sukstanskii AL, Luo J, et al. Voxel spread function method for correction of magnetic field inhomogeneity effects in quantitative gradient-echo-based MRI. *Magn Reson Med* 2013; 70: 1283-1292. Research Support, N.I.H., Extramural 2012/12/13. DOI: 10.1002/mrm.24585.

15. Wen J, Cross AH and Yablonskiy DA. On the role of physiological fluctuations in quantitative gradient echo MRI: implications for GEPCI, QSM, and SWI. *Magnetic resonance in medicine* 2015; 73: 195-203. 2014/01/30. DOI: 10.1002/mrm.25114.

16. Fischl B. FreeSurfer. *NeuroImage* 2012; 62: 774-781. DOI: <https://doi.org/10.1016/j.neuroimage.2012.01.021>.
17. Brier MR, Snyder AZ, Tanenbaum A, et al. Quantitative signal properties from standardized MRIs correlate with multiple sclerosis disability. *Annals of Clinical and Translational Neurology* 2021; 8: 1096-1109. DOI: <https://doi.org/10.1002/acn3.51354>.
18. Benjamini Y and Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* 1995; 57: 289-300. DOI: <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
19. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 1996; 58: 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>. DOI: <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>.
20. Brownlee WJ, Altmann DR, Prados F, et al. Early imaging predictors of long-term outcomes in relapse-onset multiple sclerosis. *Brain : a journal of neurology* 2019; 142: 2276-2287. DOI: 10.1093/brain/awz156.
21. Luo J, Yablonskiy DA, Hildebolt CF, et al. Gradient echo magnetic resonance imaging correlates with clinical measures and allows visualization of veins within multiple sclerosis lesions. *Multiple sclerosis (Houndmills, Basingstoke, England)* 2014; 20: 349-355. 2013/07/08. DOI: 10.1177/1352458513495935.
22. Sati P, Cross AH, Luo J, et al. In vivo quantitative evaluation of brain tissue damage in multiple sclerosis using gradient echo plural contrast imaging technique. *NeuroImage* 2010; 51: 1089-1097. DOI: <https://doi.org/10.1016/j.neuroimage.2010.03.045>.
23. Xiang B, Wen J, Lu HC, et al. In vivo evolution of biopsy-proven inflammatory demyelination quantified by R2t* mapping. *Ann Clin Transl Neurol* 2020; 7: 1055-1060. 2020/05/06. DOI: 10.1002/acn3.51052.
24. Xiang B, Wen J, Cross AH, et al. Single scan quantitative gradient recalled echo MRI for evaluation of tissue damage in lesions and normal appearing gray and white matter in multiple sclerosis. *Journal of Magnetic Resonance Imaging* 2019; 49: 487-498. DOI: <https://doi.org/10.1002/jmri.26218>.
25. Moll NM, Rietsch AM, Thomas S, et al. Multiple sclerosis normal-appearing white matter: Pathology–imaging correlations. *Annals of neurology* 2011; 70: 764-773. DOI: <https://doi.org/10.1002/ana.22521>.
26. Lassmann H. Multiple Sclerosis Pathology. *Cold Spring Harbor Perspectives in Medicine* 2018; 8. DOI: 10.1101/cshperspect.a028936.
27. Sotirchos ES, Gonzalez-Caldito N, Dewey BE, et al. Effect of disease-modifying therapies on subcortical gray matter atrophy in multiple sclerosis. *Multiple Sclerosis Journal* 2019; 26: 312-321. DOI: 10.1177/1352458519826364.
28. Pontillo G, Coccozza S, Lanzillo R, et al. Determinants of Deep Gray Matter Atrophy in Multiple Sclerosis: A Multimodal MRI Study. *American Journal of Neuroradiology* 2019; 40: 99. DOI: 10.3174/ajnr.A5915.
29. Yablonskiy DA, Wen J, Kothapalli SVVN, et al. In vivo evaluation of heme and non-heme iron content and neuronal density in human basal ganglia. *NeuroImage* 2021; 235: 118012. DOI: <https://doi.org/10.1016/j.neuroimage.2021.118012>.

Supporting Information:

	RRMS vs Visit1	Visit2 vs Visit1	Visit3 vs Visit1	Visit4 vs Visit1	Visit5 vs Visit1
EDSS	p<0.001	0.957	0.852	0.118	0.044
25FTW	0.011	1.000	0.777	0.145	0.023
9HPT(Dominant)	p<0.001	0.992	0.893	0.230	0.121
9HPT(NonDominant)	p<0.001	0.999	0.208	0.606	p<0.001
PASAT	0.692	0.997	1.000	0.959	0.666
SDMT	0.008	0.021	0.317	0.026	0.565

Table S1. p values of group comparisons for clinical assessments between the RRMS and PMS groups and within the PMS group. Significant difference was observed between the RRMS and PMS groups for EDSS, 25FTW, 9HPT and SDMT. EDSS, 25FTW and 9HPT of visit five of the PMS group was significantly different than visit one of the PMS group. One way ANOVA with repetitive measurements was used to compare the difference between visit one and each subsequent visit of the PMS group. Dunnett test was used to correct for multiple comparison. Two-sample t-test was used to compare the RRMS and the first visit of PMS group. Statistically significant correlations are highlighted in color.

	ANOVA	HC vs RRMS	HC vs Visit1	RRMS vs Visit1	Visit2 vs Visit1	Visit3 vs Visit1	Visit4 vs Visit1	Visit5 vs Visit1
Cortical GM	0.015	0.998	0.036	0.030	0.997	0.961	0.030	p<0.001
NAWM	0.069	0.719	0.065	0.259	0.846	0.995	0.056	0.077
Deep GM	0.046	0.045	0.134	0.844	0.577	0.965	0.394	0.003
TDS	NA	NA	NA	NA	0.438	0.868	0.253	0.166
Lesion Volume	NA	NA	NA	NA	0.909	0.969	0.180	0.010

Table S2. p values of group comparisons for MRI measurements between the healthy control, RRMS and PMS groups, and within the PMS group. One way ANOVA with repetitive measurements was used to compare the difference between visit one and each subsequent visit of the PMS group. Dunnett test was used to correct for multiple comparison. One way ANOVA was used to compare the healthy control, RRMS and visit one of the PMS group. Tukey test was used to correct for multiple comparison. Statistically significant correlations are highlighted in color.