A comparative assessment of myelin-sensitive measures in multiple sclerosis patients and healthy subjects

Reza Rahmanzadeh a,b, Matthias Weigel a,b,c, Po-Jui Lu a,b, Lester Melie-Garcia a,b, Thanh D. Nguyen d, Alessandro Cagol a,b, Francesco La Rosa e,f,g, Muhamed Barakovic a,b, Antoine Lutti h, Yi Wang d, Meritxell Bach Cuadra e,f,g, Ernst-Wilhelm Radue a, Laura Gaetano i, Ludwig Kappos b, Jens Kuhle b, Stefano Magon j, Cristina Granziera a,b,*

ABSTRACT

Introduction: Multiple Sclerosis (MS) is a common neurological disease primarily characterized by myelin damage in lesions and in normal-appearing white and gray matter (NAWM, NAGM). Several quantitative MRI (qMRI) methods are sensitive to myelin characteristics by measuring specific tissue biophysical properties. However, there are currently few studies assessing the relative reproducibility and sensitivity of qMRI measures to MS pathology in vivo in patients.

Methods: We performed two studies. The first study assessed the sensitivity of qMRI measures to MS pathology: in this work, we recruited 150 MS and 100 healthy subjects, who underwent brain MRI at 3 T including quantitative T1 mapping (qT1), quantitative susceptibility mapping (QSM), magnetization transfer saturation imaging (MTsat) and myelin water imaging for myelin water fraction (MWF). The sensitivity of qMRIs to MS focal pathology (MS lesions vs peri-plaque white/grey matter (PPWM/PPGM)) was studied lesion-wise; the sensitivity to diffuse normal appearing (NA) pathology was measured using voxel-wise threshold-free cluster enhancement (TFCE) in NAWM and vertex-wise inflated cortex analysis in NAGM. Furthermore, the sensitivity of qMRI to the identification of lesion tissue was investigated using a voxel-wise logistic regression analysis to distinguish MS lesion and PP voxels.

The second study assessed the reproducibility of myelin-sensitive qMRI measures in a single scanner. To evaluate the intra-session and inter-session reproducibility of qMRI measures, we have investigated 10 healthy subjects, who underwent two brain 3 T MRIs within the same day (without repositioning), and one after 1-week interval. Five region of interest (ROIs) in white and deep grey matter areas were segmented, and inter- and intra-session reproducibility of qMRI measures was investigated using the intra-class correlation coefficient (ICC). Further, we also investigated the voxel-wise reproducibility of qMRI measures in NAWM and NAGM.

Results: qT1 and QSM showed the highest sensitivity to distinguish MS focal WM and cortical pathology from peri-plaque WM (P < 0.0001), although QSM also showed the highest variance when applied to lesions. MWF and MTsat exhibited the highest sensitivity to NAWM pathology (P < 0.01). On the other hand, qT1 appeared to be the most sensitive measure to NAGM pathology (P < 0.01). All myelin-sensitive qMRI measures exhibited high inter/intra session ICCs in various WM and deep GM ROIs, in NAWM and in NAGM (ICC 0.82 ± 0.12).

Keywords:
Multiple sclerosis
Quantitative MRI
Myelin
Reproducibility
Sensitivity

* Corresponding author.
E-mail address: cristina.granziera@unibas.ch (C. Granziera).
1. Introduction

In the last decades, several myelin-sensitive quantitative magnetic resonance imaging (qMRI) techniques (e.g. myelin water imaging, magnetization transfer, T1 relaxometry, quantitative susceptibility imaging) have been developed, which provide more specific measures of demyelination and remyelination in multiple sclerosis (MS) patients than conventional MRI (Piredda et al., 2021; Granziera et al., 2021). To date, however, it is unclear how all these measures compare to each other, both in terms of sensitivity to MS damage and of reproducibility across different acquisition sessions (van der Weijden et al., 2021).

MS lesions contain variable amounts of inflammatory infiltrates, demyelination and axonal loss (Luchinetti et al., 2000). Also, the periplaque (PP) tissue around the MS lesion is not healthy and contains both myelin and axonal damage although to a lesser degree than the lesion itself (Lieury et al., 2014). Furthermore, the normal-appearing tissue (i.e. white and gray matter areas that do not exhibit signs of focal pathology, NAWM and NAGM, respectively) is characterized by diffuse myelin and axonal damage and microglia clusters (Multiple and Pathology, 2018; Granberg et al., 2017; Cui et al., 2017; Kutzelnigg et al., 2005). To date it is unclear which qMRI measure is the most sensitive to these different pathological hallmarks of MS pathology.

Myelin-sensitive techniques exploit different contrast mechanisms to assess specific characteristics of the myelin sheaths. Myelin water imaging (MWI) quantifies the water between myelin layers by distinguishing multiple water components in T2 relaxometry data, and provides measures (e.g. myelin water fraction, MWF) that have been validated post-mortem (Moore et al., 2000; Kozlowski et al., 2014). Quantitative susceptibility mapping (QSM) quantifies the spatial distribution of magnetic susceptibility in biological tissue (Liu et al., 2012) and has been shown to be sensitive to iron content and to myelin integrity (Hamerter et al., 2018; Wisnieff et al., 2015). Quantitative T1 mapping (qT1) quantifies T1 relaxation times that are sensitive to water content and macro/micro molecules changes within a tissue, such as the one provoked by demyelination and axonal loss in the brain (Granziera et al., 2021; Kolb et al., 2021). Although it is challenging to disentangle the contribution of these factors in acquired qT1, it has been shown that qT1 correlates well with myelin content in NAWM and MS lesions (Mottershead et al., 2003; Seewann et al., 2009). Magnetization transfer (MT) imaging measures the magnetization exchange between protons in free water and protons bound to macromolecules, which has been related to myelin integrity and content (Moccia et al., 2020). MT saturation (MTsat) was developed to improve the MT ratio (MTR) by decoupling the MTR from T1 relaxation time contributions and by incorporating excitation flip angle inhomogeneity, thereby overcoming some limitations of previous MT-based methodologies (Helms et al., 2008). Combined MRI – neuropathology studies have shown that the MT ratio correlates well with myelin content within the brain tissue but also to the presence of macrophages and astrocytes (Moccia et al., 2020; Schmierer et al., 2004).

All the above mentioned myelin-sensitive qMRI techniques might provide differential sensitivity to MS pathology hallmarks. Nevertheless, only few previous studies explored the relative sensitivity of some of these measures (e.g., qT1, quantitative T2 and MTR to MS damage in WM (O’Muircheartaigh et al., 2019; Vavasour et al., 2006), but none included QSM and MTsat as well as an assessment of cortical pathology.

Last, most previous studies evaluated myelin-sensitive qMRI reproducibility using single-contrast approaches, which renders it challenging to compare obtained results among different qMRI measures, due to the heterogeneous experimental settings.

In this work, we further expanded previous knowledge about myelin-sensitive qMRI measures by comprehensively evaluating (i) their sensitivity to MS pathology in a large cohort of patients and healthy controls; as well as (ii) their mono-centric reproducibility.

2. Materials and methods

2.1. Participants

2.1.1. Sensitivity study

To assess sensitivity to MS pathology, we have considered both sensitivity to focal damage (i.e. sensitivity of qMRI to alterations in MS lesions vs periplaque tissue) as well as to diffuse damage (i.e. sensitivity of qMRI to alterations in NAWM and NAGM vs healthy WM and GM). We enrolled 150 MS patients (92 RRMS and 58 PMS) and 100 healthy controls. The inclusion criteria were: (i) MS diagnosis according to McDonald criteria 2018 (Thompson et al., 2018) and diagnosis of active RRMS or inactive PMS as defined by Lublin et al. (Lublin et al., 2014); (ii) absence of any concomitant psychiatric or neurological disease (excluding headache); (iii) absence of contraindication to MRI. The ethical review committee of the University Hospital Basel (IRB of Northwest Switzerland) approved the study, and all participants entered the study following written consent. All subjects underwent MRI at 3 T. Clinical characteristics of patients and healthy controls are reported in Table 1.

2.1.2. Reproducibility study

We have enrolled 17 healthy subjects who underwent two scans without repositioning (time point 1 and 2, TP1 and TP2) and a third scan (TP3) was performed 1 week (±3 days) later.

The design of this reproducibility assessment allowed to investigate (i) intra-session reproducibility without subjects repositioning; as well as (ii) inter-session reproducibility after repositioning, Fig. 1.

Seven subjects were excluded from the study: motion artefacts (n = 4), which were to be mainly attributed to the applied study design (scan-rescan without repositioning, acquisition time: 1 h:25 min:12 s), technical issues of the MR system leading to either heavy artifacts (n = 1) r impossibility to acquire MRI in time at TP3 (n = 2) (Fig. 1).

2.2. MR acquisition and qMRI maps reconstruction

2.2.1. MR acquisition

MRI was performed on a 3 T whole-body MR system (Magnetom
1800 ms; R (Ganter et al., 2013); and to correct for effects of radio frequency transmit inhomogeneity on the quantitative maps (Helms and Dechent, 2008). Its sensitivity to the organization of the WM tissue and its microstructural myelin measurements (Laule et al., 2006).

By building a lookup table for these uniform intensities with protocol parameters as dependents and T1 as a free parameter, quantitative T1 mapping was performed on FLAIR and MP2RAGE images by consensus (WM: RR and CG, cortical: RR and AC) (Geurts et al., 2011). Both intracortical lesions and the GM part of leukocortical lesions were considered as “cortical lesions”. To identify the GM part of leukocortical lesions, we first manually segmented both their GM and WM part, and then masked them with the GM mask obtained with FreeSurfer (Rahmanzadeh et al., 2021). A few patients (n = 3) with dirty WM were excluded from study because for those accurate delineation of lesion border was not possible, which could have led to inaccurate estimation of qMR in lesions, PP and NA tissue (Fig. 3).

For the reproducibility study, six regions of interests (splenium of corpus callosum, genu of corpus callosum, putamen, head of caudate, thalamus and cortical grey matter) segmented automatically by imaging software package FreeSurfer (FS) (v.6.0, surfer.nmr.mgh.harvard.edu) (Fischl, 2012) were manually corrected on FLAIR and MP2RAGE. The ROIs mask were then registered to qMR spaces using a linear registration in FMRIB Software Library (FSL) (Jenkinson et al., 2012) with nearest-neighbor interpolation.

2.4. WM and cortex segmentation

For both studies, we used the imaging software package FreeSurfer (https://surfer.nmr.mgh.harvard.edu/) to segment the brain into whole WM, cortex, deep grey matter structures, and ventricles. NAWM and NAGM masks were obtained by subtracting WM and cortical lesion masks from WM and cortical masks.

An in-house developed algorithm was used to automatically produce a 2-voxel layer of NAWM surrounding the lesions; herein after denoted peri-plaque (PP).

We have opted to assess differences between lesions and peri-plaque white matter/grey matter (PPWM, PPGM) to avoid spurious changes due to the underlying anatomy. In fact, all qMRI measures applied in this study are sensitive to the organization of the WM tissue and its microscopic anatomical features (i.e. bundles orientations, myelin content, cellular distribution etc.). Comparing lesions to ROIs in more remote

---

**Fig. 1.** Reproducibility study design. Time points (TP) 1 & 2 were performed without repositioning and TP3 1 week ± 3 days later.
white matter than PPWM will suffer from the bias due to the different tissue microstructure. Furthermore, we have chosen to report absolute values to allow the comparability of the sensitivity of the different qMRI maps to MS pathology. In fact, whilst qT1, MTsat and MWF are expected to respectively increase (qT1) or decrease (MTsat and MWF) based on previous literature (summarized in Granziera et al. (Granziera et al., 2021)), QSM may either increase (due to myelin loss or iron accumulation) or decrease (due to remyelination or edema accumulation) or both (Chen et al., 2014).

The absolute delta of changes between lesions and periplaque tissue was calculated as follows: \( \Delta \text{WML-PPWM} \) \& \( \Delta \text{CL-PPGM} \) between lesion and PP tissue for WMLs and CLs as follows:

\[
\Delta \text{lesion-PP} = \frac{\text{mean qMRI measure in lesion} - \text{mean qMRI measure in PP}}{\text{mean qMRI measure in PP}} \quad \text{(formula (1))}
\]

A minority of lesions were excluded from the analysis (353 WMLs (7.6% of all WMLs) and 217 CLs (30% of all CLs)), due to the challenges in establishing an automatic procedure to assess the PP-tissue in i) lesions of very small size (due to partial-volume effects) ii) lesions close to ventricle wall and WM-GM border or iii) lesions located in/around WM bundles.

To note, QSM studies in the cortical areas are challenging due to cortical erosion and the presence of banding/streaking artifacts (Yaghmaie et al., 2021). In this work, we minimized these issues by i) identifying the lowest erosion threshold that did not impact the obtained QSM values; and ii) by carefully examining the presence of artifacts in the obtained maps.

For the reproducibility study, FS was performed using MP2RAGE acquired at baseline and FS outputs were then registered to the other time-points using the linear registration with nearest-neighbor interpolation.

2.5. Voxel-Based analyses in NAWM

NAWM maps were co-registered patient-wise to a reference brain (standard MN152 space) using an affine registration (FLIRT) followed by non-linear FNIRT registration in FSL (https://fsl.fmrib.ox.ac.uk/fsl/) (Jenkinson et al., 2012). Then, the NAWM maps were wrapped to WM templates using – applywarp commandline in FSL. As previously performed (Vrenken et al., 2006), in NAWM we excluded voxels that were not present in at least 50 percent of subjects, and (Granziera et al., 2021) filled missing data with the group mean value of those voxels present in group subjects. By using the randomize tool of FSL with Threshold-Free Cluster Enhancement (TFCE), we carried out a voxel-wise comparison of MWF, qT1, MTsat and QSM maps (i) between patients and controls (sensitivity analysis); and (ii) between TP1-TP2 & TP1-TP3 (reproducibility analysis). \( P \) values < 0.01 were considered statistically significant. To compare the extent of qMRI measure changes in areas showing significant clusters, the mean qMRIs maps were computed using FSL.

2.6. Vertex-wise analysis in NAGM

A customized volume-to-surface mapping algorithm was applied to voxels assigned to the grey matter ribbon by FreeSurfer - i.e., voxels with coordinates located between the white and pial surfaces were registered.
and projected into a standard surface. A smoothing kernel of 10-mm full-width at half-maximum was used. Then, generalized linear model (GLM) analysis was conducted to assess (i) the sensitivity of each map in NAGM to assess damage between patients and controls; and (ii) to assess the intra- and inter-scanner reproducibility of each qMRI measure in GM (TP1 vs TP2 & TP1 vs TP3, respectively). P values < 0.01 were considered statistically significant.

2.7. Intra-class correlation coefficient to assess intra- and inter-session reproducibility

In six regions of interest (ROIs: genu and splenium of the corpus callosum, putamen, head of caudate, thalamus and cortical GM), we calculated the intra-class correlation coefficient (ICC) for each qMRI measure using a two-way model with consistency of agreement within and across sessions (TP1-TP2, TP1-TP3). The ICC analysis was performed in Stata 16 statistical package. According to Koo et al. (Koo and Li, 2016), ICC < 0.5, 0.5 ≤ ICC < 0.75, 0.75 ≤ ICC < 0.90, 0.90 ≤ ICC were interpreted as indicative of poor, moderate, good and excellent reliability, respectively.

2.8. Statistical analyses

Statistical analysis was performed using GraphPad Prism version 8.0.0 for Windows, GraphPad Software, San Diego, California USA. Kolmogorov-Smirnov’s test was used to assess the normality of data. Non-parametric Mann-Whitney test and Kruskal-Wallis test with Dunn’s test for multiple comparisons correction, were used for unpaired two-group analysis and more-than-three group analysis, respectively.

To account for both differences in mean and variance in the comparison among myelin-sensitive measures, we repeated the same comparison for Z-scores of delta values, as follows:

\[
Z \text{ score} = \frac{\text{lesion } \Delta - \text{ mean lesion } \Delta \text{ across all lesions}}{\text{standard deviation of lesion } \Delta \text{ across all lesions} \ (\text{formula (2)})}
\]

We have pooled RRMS and PMS together, because there is increasing evidence showing that there’s no difference in the microstructural characteristics of focal pathology between RRMS and PMS (Lucchinetti et al., 2000; Rahmanzadeh et al., 2021; Antel et al., 2012).

For the independent two samples analysis, a Cohen’s d effect size was calculated as Cohen’s d = (M2 - M1) / SDpooled, where M1 and M2 are the group means and SDpooled = \sqrt{(SD_1^2 + SD_2^2) / 2}.

3. Results

3.1. qMRI measures sensitivity to focal and diffuse MS pathology

3.1.1. Sensitivity of myelin-sensitive qMRI measures to focal MS pathology

WM lesion-wise analysis showed that mean \( \Delta_{\text{WML-PPWM}} \) was the highest for QSM and the lowest for MTsat (Fig. 4, Table 2); \( \Delta_{\text{WML-PPWM}} \) was higher for QSM than for qT1 (P < 0.0001; Cohen’s d effect size = 0.69) and for qT1 than for MWF (P < 0.0001; Cohen’s d effect size = 0.69). Likewise, qT1 shows higher mean Z-score compared to MWF and MTsat (both P < 0.001). However, in contrast to \( \Delta \) analysis, the mean Z-score for QSM was lower than the one of MWF and qT1 (both P < 0.001; Fig. 5).

Cortical lesion-wise analysis also showed that mean \( \Delta_{\text{CL-PPGM}} \) was the highest for QSM and the lowest for MTsat (Fig. 4, Table 2); \( \Delta_{\text{CL-PPGM}} \) was higher for QSM than for qT1 (P < 0.0001; Cohen’s d effect size = 0.75) and for qT1 than for MWF (P < 0.0001; Cohen’s d effect size = 1.46). When the z-score analysis was considered, qT1 showed the highest mean Z-score compared to MWF and MTsat (both P < 0.05; Fig. 5). However, in contrast to the \( \Delta \) analysis, the mean Z-score for QSM was not higher than the one measured using the other myelin sensitive-measures.

3.1.2. Sensitivity of qMRI measures to NAWM pathology in MS

Voxel-wise analysis showed that all qMRI measures were sensitive to changes in some clusters of NAWM voxels compared to WM in healthy controls. However, MTsat, MWF, QSM and qT1 showed decreasing...
sensitivity to NAWM pathology in MS (Fig. 6, total clusters voxel number for MWF: 285820, MTsat: 340304, QSM (QSM1 + QSM2): 132548, qT1: 36757; P < 0.01).

The average qMRI measures in clusters of significant difference between patients and controls (P < 0.01) were as follows: qT1 = 799.88 vs 768.27 ms (patients vs controls); MWF = 6.35 vs 7.09 p.u. (percent unit, patients vs controls); MTsat = 1.39 vs 1.55 p.u. (patients vs controls); QSM1 = -12.15 vs -18.77 ppb (patients vs controls); QSM 2 = 26.08 vs 18.39 ppb (patients vs controls) in QSM2.

3.1.3. Sensitivity of qMRI measures to NAGM pathology in MS

Vertex-wise cortical surface analysis showed that all qMRI measures were sensitive to changes in some clusters of NAGM voxels compared to GM in healthy controls. However, qT1, QSM, MWF and MTsat appeared to show decreasing sensitivity to NAGM pathology in MS (Fig. 7, total clusters voxel number for qT1: 4223, QSM (QSM1 + QSM2): 2678, MWF: 926, MTsat: 876; P < 0.01). Except for QSM, for other qMR the surface analysis yielded no significant cluster for the comparison MS > HC.

3.2. Reproducibility of myelin-sensitive qMRI measures in a single scanner

3.2.1. ICC analysis

Intra-session (TP1-TP2) and inter-session (TP1-TP3) ICCs for qMRI measures are summarized in Table 3.

qT1 exhibited the highest intra-session and inter-session reproducibility (ICCs TP1 vs TP2: 0.83–0.98; ICC TP1 vs TP3: 0.77–0.95, respectively). MWF followed with an intra-session ICC of 0.81–0.98, and an inter-session ICC of 0.50–0.90, with lower ICCs in deep gray matter structures than CC parts (Table 3). Also, MTsat and QSM exhibited relatively high ICC in most ROIs (MTsat: intra-session ICC = 0.59–0.95 and inter-session ICC = 0.67–0.88; QSM: intra-session ICC = 0.57–0.93 and inter-session ICC = 0.62–0.90, Table 3).

Table 2

Summary of Mean and standard deviations (M, SD) for \( \Delta \text{WML-PPWM} \) and \( \Delta \text{CL-PPGM} \).

<table>
<thead>
<tr>
<th></th>
<th>MWF</th>
<th>qT1</th>
<th>MTsat</th>
<th>QSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta \text{WML-PPWM} ) (M ± SD)</td>
<td>0.19 ± 0.16</td>
<td>0.31 ± 0.17</td>
<td>0.18 ± 0.13</td>
<td>0.64 ± 0.66</td>
</tr>
<tr>
<td>( \Delta \text{CL-PPGM} ) (M ± SD)</td>
<td>0.25 ± 0.26</td>
<td>0.66 ± 0.30</td>
<td>0.18 ± 0.17</td>
<td>1.07 ± 0.71</td>
</tr>
</tbody>
</table>

Fig. 4. Comparison of the sensitivity of qMRI measures in distinguishing lesion vs peri-plaque tissue in WM and in the cortex (because of QSM changes in both direction in MS, absolute \( \Delta \text{WML-PPWM} \) & \( \Delta \text{CL-PPGM} \) are reported). * P < 0.05, ** P < 0.001, *** P < 0.0001.

Fig. 5. Comparison of Z-score of qMRI delta measures between lesion vs peri-plaque tissue in WM and in the cortex. * P < 0.05, ** P < 0.001, *** P < 0.0001.
3.2.2. Voxel-wise WM reproducibility analysis

For all qMRI measures, WM voxel-wise analysis exhibited no clusters of significant intra- and inter-session differences ($P < 0.01$).

3.2.3. Vertex-wise GM reproducibility analysis

Vertex-wise cortical surface analysis showed very small areas of intra- and inter-session differences for all qMRI measures. The total clusters numbers for each measure were as follows: $qT1 = 155$, QSM ($QSM1 + QSM2 = 182$, $MWF = 312$, $MTsat = 297$ ($P < 0.01$). There was also no substantial overlap between the significant clusters in reproducibility (TP1 vs TP2) and sensitivity (patients vs controls) as outputs of vertex-wise inflated cortex analyses.

4. Discussion

In this work, we performed a comprehensive assessment of the sensitivity of myelin-sensitive qMRI measures to MS pathology, as well as of their respective intra-scanner reproducibility.

Our results showed that qMRI measures such as $MWF$, $MTsat$, $qT1$, and QSM exhibit differential sensitivity to MS pathology in MS plaques and in WM and cortical regions outside areas of focal damage.

When a comparison of the mean changes of qMRI measures in lesion vs periplaque tissue was performed, $qT1$ and QSM showed the highest sensitivity to MS lesions pathology in both WM and cortical GM. This may be due to the broader sensitivity of $qT1$ to the tissue destruction in CNS encompassing demyelination, axonal degeneration, edema and...
tissue destruction (Lucchinetti et al., 2000; Lassmann, 2018), which are predominant within the core of the lesion and present to a smaller extent in the peri-plaque tissue (Mustafi et al., 2019; Rahmanzadeh et al., 2021; Vrenken et al., 2006; Rodriguez et al., 1993). The higher sensitivity of qT1 for cortical focal pathology may also be at least in part depend on the fact that qT1 maps were obtained using the same acquisition where cortical lesions were detected (MP2RAGE). On the other hand, the quantification of MWF, MTsat in cortical lesions may have suffered from partial volume effects due to the lower spatial resolution of the respective images compared to MP2RAGE. Interestingly, when the “sensitivity” to MS pathology was calculated using z-scores instead of mean delta measures, results were very similar with the exception of the ones obtained for QSM, which appeared the “least sensitive” among all. Assessing the sensitivity of a given measure using z-scores instead of the “delta of the mean” has the advantage to consider the variance of that measure, which is an important aspect if that variance is of technical origin (i.e. a higher variance will reduce the sensitivity to pathology). Nevertheless, we recently showed that the variance of QSM measures among MS lesions has an important pathological source (Rahmanzadeh et al., 2021). Therefore, we concluded that that the most appropriate way to show the sensitivity to MS pathology is the one that does not penalize a measure because of its variance (i.e. the mean delta): by using this measure qT1 and QSM appeared to outperform MWF and MTsat to identify focal MS lesion pathology. Interestingly, the sensitivity of qT1 to NAWM pathology was the lowest among all qMRI measures. In contrast, MTsat and MWF were the most sensitive to NAWM pathology. Indeed, MWF and MTsat exploit different contrast mechanisms allowing to detect subtle changes in myelin water and lipid/macromolecular content, which may be the consequence of diffuse microglia activation in NAMW (Cui et al., 2017; Rodriguez et al., 1993; Rodriguez and Scheithauer, 1994).

As to NAGM pathology, qT1 and QSM were the most sensitive measures to assess the consequences of neuroinflammatory and neurodegenerative processes occurring in the cortical ribbon of MS patients (Kutzelnigg et al., 2005; Magliozzi et al., 2007). To note, QSM studies in the cortical areas are challenging due to cortical erosion and the presence of banding/streaking artefacts (Yaghmaie et al., 2021). In this work, we minimized these issues by (i) identifying the lowest erosion threshold that did not impact the obtained QSM values; and (ii) carefully examining the presence of artifacts in the obtained maps.

All qMRI measures exhibited high intra- and inter-session reproducibility; qT1 obtained with MP2RAGE (Marques et al., 2010) exhibited good-to-excellent ICCs (0.77–0.98) in both WM (i.e. splenium/genu of corpus callosum) and in deep gray matter structures. Similarly, qT1 showed the highest reproducibility in voxel-wise and vertex-wise analyses. It is true that the reproducibility of qT1 measures is highly hardware-, software- and sequence-dependent (Bane et al., 2018); nevertheless, the results obtained in this study by using MP2RAGE confirm the feasibility of reproducible measurements across sessions in a single 3 T scanner, which were previously reported (Marques et al., 2010).

As to MTsat, previous reproducibility studies have shown good to excellent ICC for both inter-scanner and intra-scanner acquisitions, both in single-center and multi-center studies (Barker et al., 2005; Weiskopf et al., 2013; Ropele et al., 2005). This is in contrast with the reproducibility of MTR, which is modest across centers due to its high sensitivity to (i) sequence parameters and setup, (ii) the inhomogeneity of the transmission field, and finally to (iii) the inability of the applied reconstruction algorithm to correct qT1 that may occur in different scanners/centers (Helms et al., 2010). Our results support and extend previous findings obtained with MTsat in monocentric assessments by showing an overall good reproducibility for MT in all ROIs except in the caudate (ICC intra-session: 0.59, inter-session: 0.67). Similarly, the voxel-wise WM and vertex-wise cortical analysis showed almost no intra-/inter-session changes in MTsat, a finding that highlights the robustness of MTsat for myelin imaging in a single-scanner and single-center setting.

Regarding QSM, our results confirm previous reports of high intra-scanner and inter-scanner reproducibility for the applied MEDI reconstruction method (Deh et al., 2019). Interestingly, these results are also supported by a multi-center phantom study where the MEDI QSM reconstruction appeared to be highly reproducible (ICC > 0.99) among different clinical and preclinical scanners (Deh et al., 2019). Nevertheless, high reproducibility in QSM experiments seems not to be limited to the MEDI reconstruction method, as also other have shown high intra-scanner reproducibility (i.e. the L2 regularization QSM) (Lin et al., 2015; Santin et al., 2017; Choi et al., 2019), supporting the knowledge that QSM provide robust qMRI measures for multicentric studies.

Similarly, confirming one previous study (Nguyen et al., 2016), our work shows that the reproducibility of FAST-T2 MWF maps is high when assessed both with a region of interest and with a voxel-wise or surface-wise approach in both WM and cortical GM. Previous works also showed a high reproducibility of myelin water fraction maps obtained with other acquisition methods such as GRASE and mcDESPO (multi-component Driven Equilibrium Single Pulse Observation of T1 and T2) (Lee et al., 2018; Levy et al., 2018; Meyers et al., 2009), although in this case no voxel-wise and surface-wise comparison was performed.

Notably, the reported reproducibility performance does not only depend on the applied acquisition method, but also on (i) the relative spatial resolution/partial volume effects, which may partially explain the outperformance of qT1 maps in cortical regions; as well as on (ii) the applied reconstruction algorithm (e.g. MEDI for QSM). A comprehensive evaluation of these aspects is beyond the scope of the current study, but should be considered in future works.

Our work is unique in that it compares different qMRI measures for both focal and diffuse sensitivity to MS pathology as well as for intrasession and inter-session reproducibility at 3 T MRI. Moreover, our study applies qMR measures derived from sequences acquired in a clinically feasible time (acquisition time ranging from 4.5 min for MWF to 9.45 min for MTsat), providing therefore a reference for future clinical trials or clinical applications of myelin-sensitive qMRI. Extending previous studies (Vavasour et al., 2006), our work evaluated the reproducibility of multiple myelin-sensitive qMRI measures not only in specific regions of interest in WM but also in cortical grey matter and deep grey matter nuclei. Further, our work explored first the sensitivity of (i) a clinically-compatible T1 mapping technique such MP2RAGE; (ii) MTsat, a quantitative measure for MT effects that is less sensitive to B1 inhomogeneity and T1 effects than the previously applied MTR (Vavasour et al., 2006); and of (iii) QSM obtained from a fast 3D-EPI based T2* weighted acquisition.

Nonetheless, our study provides evidence of a good reproducibility of all qMRI measures only in a single-center and single-scanner setting: therefore, these results warrant confirmation in a future study where reproducibility among different scanners, vendors and field strengths, is assessed. Moreover, our study was not designed to assess reproducibility of time intervals that are normally applied in longitudinal studies (e.g. > 1 year), therefore future work should assess this aspect.

Another limitation of this study was the high drop-out rate (7/17): just over the half (Lucchinetti et al., 2000) due to motion artifacts derived from the length of the protocol including multiple myelin-sensitive qMRI sequences and intra-session reproducibility assessment without repositioning (cf. Materials and Methods section).

In summary, we demonstrated that the applied qMRI measures provide complementary information about the extent of MS pathology in the brain regions corresponding both WM and GM areas. Our data suggest that the combination of qT1 and MWF might be ideal for a comprehensive study of myelin-damage in the brain of MS patients in single center settings, due to their complementary and optimal sensitivity to focal and diffuse MS-damage in cortical grey and white matter as well as to their high intra-scanner reproducibility. QSM did not seem to provide advantages when average analyses were performed, although we know that it allows to qualitatively identify MS lesion subtypes (Rahmanzadeh et al., 2021; Chen et al., 2014; Absinta et al., 2018). Last,
MTsat might be a good compromise for an all-in-one approach due to its relatively high sensitivity to MS damage and reproducibility. Future studies should further assess the relative correlation of qMRI measures with quantification of tissue properties and pathology in MS brains, as well as replicate the current results in larger cohorts of MS patients in vivo.

CRedit authorship contribution statement


Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References


