Test–Retest Reproducibility of In Vivo Magnetization Transfer Ratio and Saturation Index in Mice at 9.4 Tesla

Naila Rahman, BSc,1,2* Jordan Ramnarine, BMSc(c),1 Kathy Xu, MSc,3 Arthur Brown, PhD,3,4 and Corey A. Baron, PhD1,2

Background: Magnetization transfer saturation (MTsat) imaging was developed to reduce T1 dependence and improve specificity to myelin, compared to the widely used MT ratio (MTR) approach, while maintaining a feasible scan time. As MTsat imaging is an emerging technique, the reproducibility of MTsat compared to MTR must be evaluated.

Purpose: To assess the test–retest reproducibility of MTR and MTsat in the mouse brain at 9.4 T and calculate sample sizes potentially required to detect effect sizes ranging from 6% to 14%.

Study Type: Prospective.

Subjects: Twelve healthy C57Bl/6 mice.

Field Strength/Sequence: 9.4 T; magnetization transfer imaging using FLASH-3D Gradient Echo; T2-weighted TurboRARE spin echo.

Assessment: All mice were scanned at two timepoints (5 days apart). MTR and MTsat maps were analyzed using mean region-of-interest (ROIs: corpus callosum [CC], internal capsule [IC], hippocampus [HC], cortex [CX], and thalamus [TH]), and whole brain voxel-wise analysis.

Statistical Tests: Bland–Altman plots were used to assess biases between test–retest measurements. Test–retest reproducibility was evaluated via between and within-subject coefficients of variation (bsCV and wsCV, respectively). Sample sizes required were calculated (significance level: 95%; power: 80%), given effect sizes ranging from 6% to 14%, using both between and within-subject approaches. Results were considered statistically significant at $P \leq 0.05$.

Results: Bland–Altman plots showed negligible biases between test–retest sessions (MTR: 0.0009; MTsat: 0). ROI-based and voxel-wise CVs revealed high reproducibility for both MTR (ROI-bsCV/wsCV: CC—4.5/2.8%; IC—6.1/5.2%; HC—5.7/4.6%; CX—5.1/2.3%; TH—7.4/4.9%) and MTsat (ROI-bsCV/wsCV: CC—6.3/4.8%; IC—7.3/5.1%; HC—9.5/6.4%; CX—6.7/6.5%; TH—7.2/5.3%). With a sample size of 6, changes on the order of 15% could be detected in MTR and MTsat, both between and within subjects, while smaller changes (6%–8%) required sample sizes of 10–15 for MTR, and 15–20 for MTsat.

Data Conclusion: MTsat exhibited comparable reproducibility to MTR, while providing sensitivity to myelin with less T1 dependence than MTR.

Evidence Level: 2

Technical Efficacy: Stage 1

MT is a physical process by which macromolecules and their closely associated water molecules cross-relax with protons in the free water pool. Based on this phenomenon, it is possible to quantify the protons bound to large molecules, which are not MR visible, due to their extremely short transverse relaxation time (T2). MT contrast can be generated by...
applying an off-resonance radiofrequency pre-pulse (MT pulse) to selectively saturate the spectrally broad macromolecular proton pool. This saturation then transfers to the free water proton pool via MT, resulting in a decrease in the observed free water signal. The magnitude of the MT effect can be characterized by the MTR

\[
\text{MTR} = \frac{\text{PDw} - \text{MTw}}{\text{PDw}}
\]  

where PDw is the signal without an MT pulse applied, which is proton density weighted (PDw), and MTw is the signal with the MT pulse applied, which is MT weighted (MTw). Although MTR has been shown to correlate well with histological myelin content,\textsuperscript{1,4} it is also sensitive to the choice of sequence parameters, flip angle inhomogeneities, and longitudinal relaxation time (T1).\textsuperscript{5} T1 also correlates strongly with myelin content, but is also sensitive to axon size\textsuperscript{6} and iron content,\textsuperscript{7} mitigating the power of MTR as a measure of myelin. Quantitative MT has been used in many recent works to quantify myelin,\textsuperscript{8–10} as it reduces the confounding effects of scan parameters and quantifies specific tissue characteristics, such as the macromolecular pool size.\textsuperscript{3} However, qMT relies on complex modeling of the MR signal dependence on myelin, and requires more measurements and thus longer acquisition times compared to contrast-weighted MT protocols.\textsuperscript{3}

Magnetization transfer saturation (MTsat) imaging was developed to improve MTR, by decoupling MTR from T1 effects, while maintaining a feasible scan time.\textsuperscript{5} The shorter scan time compared to qMT enables longitudinal in vivo imaging and allows the addition of other imaging techniques required to characterize microstructure. A scalar map of MTsat can be acquired using two reference scans of proton density and T1 weighting (PDw and T1w, respectively), and one MTw scan. MTsat, being more independent of system parameters and T1 weighting, and less susceptible to inhomogeneities of the receiver coil and the transmitted RF field, provides greater specificity and contrast compared to MTR.\textsuperscript{5,11} MTsat shows higher white matter contrast in the brain than MTR,\textsuperscript{5} and has been shown to correlate more with disability metrics than MTR in patients with multiple sclerosis.\textsuperscript{12} Hagiwara et al reported that MTsat may be more suited to measure myelin in the white matter, compared to the ratio of T1-weighted to T2-weighted images, which has also been proposed as a measure of myelin.\textsuperscript{13}

There is strong interest in applying MT to preclinical rodent neuroimaging studies at ultra-high field strengths (≥7 T), demonstrated by MTR\textsuperscript{14–17} and qMT studies.\textsuperscript{10,18–20} The feasibility of MTsat imaging in mice at 9.4 T has been shown previously\textsuperscript{21} and MTsat has been explored in a feline model of demyelination at 3 T.\textsuperscript{22}

While most MTsat studies have been performed at 3 T, recently, Olsson et al reported an optimized whole-brain MTsat protocol at 7 T,\textsuperscript{23} which highlights the increasing interest in this method. Although previous in vivo MTsat studies have shown high reproducibility in humans at 3 T,\textsuperscript{24,25} the comparability of MTR and MTsat reproducibility has not been fully evaluated. This also leaves open the question of MTsat reproducibility in a preclinical setting, at an ultra-high field strength. As MTsat provides a time-efficient alternative to fully quantitative techniques but with increased specificity and contrast compared to MTR, investigation of MTsat in preclinical rodent imaging will likely be of interest to other research groups. The aim of this work was to assess test–retest reproducibility of in vivo MTR and MTsat in mice at 9.4 T and provide estimates of required sample sizes, which is essential in planning preclinical neuroimaging studies involving models of disease/injury.

**Methods**

**Subjects**

All animal procedures were approved by the University of Western Ontario Animal Use Subcommittee and were consistent with guidelines established by the Canadian Council on Animal Care. Twelve adult C57Bl/6 mice (six males and six females) were scanned twice 5 days apart. The sample size was chosen to reflect similar sample sizes used in other preclinical imaging studies.\textsuperscript{26–28} Before scanning, anesthesia was induced by placing the animals in an induction chamber with 4% isoflurane and an oxygen flow rate of 1.5 L/min. Following induction, isoflurane was maintained during the imaging session at 1.8% with an oxygen flow rate of 1.5 L/min through a custom-built nose cone. The mouse head was fixed in place using ear bars and a bite bar to prevent head motion. These mice were also part of a different longitudinal study with three additional imaging sessions following the test and retest scans, at the end of which they were euthanized for histology. The mice were anesthetized with ketamine/xylazine\textsuperscript{1,2} and then underwent trans-cardiac perfusion with ice-cold saline, followed by 4% paraformaldehyde in phosphate-buffered saline.

**In Vivo MRI**

In vivo MRI experiments were conducted on a hybrid system: Agilent 9.4 Tesla, 31-cm bore magnet (Agilent, Palo Alto, CA), equipped with a 60-mm gradient coil set of 1 T/m strength ( slew rate = 4100 T/m/s) (Agilent) and Bruker Avance MRI III console with software package of Paravision-7 (Bruker BioSpin Corp, Billerica, MA). A single channel transceiver surface coil (20 × 25 mm), built in-house, was fixed in place directly above the mouse head to maximize signal-to-noise ratio (SNR). A boost in SNR in the cortex when using this surface coil, compared to a
commercially available 40-mm millipede (MP40) volume coil (Agilent), has been reported previously.29

The MT protocol required 50 minutes total scan time and comprised three FLASH-3D (fast low angle shot) scans, and one RF transmit field (B1) map scan to correct for local variations in flip angle. An MT-weighted scan, and reference T1-weighted and PD-weighted scans (MTw, T1w, and PDw, respectively) were acquired by appropriate choice of the repetition time (TR) and the flip angle (1); TR/ = 8.5 ms/20° for the T1w scan and 25 ms/90° for the PDw and the MTw scans. MT-weighting was achieved by applying an off-resonance Gaussian-shaped RF pulse (12-ms duration, 385 μm in-plane; 12 averages. The B1 map was acquired at a lower resolution of 600 × 600 × 400 μm³ and the following scan parameters: TE = 4 ms; α = 60°; short TR = 20 ms; long TR = 100 ms; 2 averages. Anatomical images were also acquired for each subject within each session using a 2D T2-weighted TurboRARE pulse sequence (150-μm in-plane resolution; 500-μm slice thickness; TE/TR = 40/5000 ms; 16 averages; total acquisition time = 22 minutes).

**Image Processing**

MTR and MTsat maps were generated using in-house MATLAB (ver. 2020b, Mathworks, Natick, MA) code. Gaussian filtering (full-width-half-maximum = 3 voxels) was first applied to the original images (MTw, PDw, and T1w images, and B1 maps) to reduce noise, while retaining image contrast. The standard MTR maps were calculated using Eq. 1. MTw, PDw, and T1w images were used to calculate MTsat maps, following the original method proposed by Helms et al,9 and outlined by Hagiwara et al.13 The following parameter estimates are influenced by local flip angle errors and are hence labeled by the subscript “app.” The apparent longitudinal relaxation rate, R1app, was calculated as follows:

\[
R_{1app} = \frac{1}{2} \frac{S_{T1} \alpha_{T1}/TR_{T1} - S_{PD} \alpha_{PD}/TR_{PD}}{S_{PD}/\alpha_{PD} - S_{T1}/\alpha_{T1}}
\]

where \( S_{T1} \) and \( S_{PD} \) denote signal intensities of T1w and PDw images, respectively; \( TR_{T1} \) and \( TR_{PD} \) denote TR of T1w and PDw images, respectively; and \( \alpha_{T1} \) and \( \alpha_{PD} \) denote excitation flip angles of T1w and PDw images, respectively. The apparent signal amplitude, Aapp, was calculated as follows:

\[
A_{app} = S_{PD} S_{T1} \frac{TR_{PD} \alpha_{T1}/\alpha_{PD} - TR_{T1} \alpha_{PD}/\alpha_{T1}}{S_{T1} TR_{PD} \alpha_{T1} - S_{PD} TR_{T1} \alpha_{PD}}
\]

Using \( R_{1app} \) and \( A_{app} \), the apparent MT saturation, MTsatapp, was calculated as follows:

\[
MT_{satapp} = \left( \frac{A_{app} \alpha_{MT}}{S_{MT}} - 1 \right) R_{1app} TR_{MT} - \alpha_{MT}^2/2
\]

where \( S_{MT} \), \( TR_{MT} \), and \( \alpha_{MT} \) denote signal intensity, TR, and excitation flip angle of the MTw image, respectively. MTsatapp is inherently robust against differences in relaxation rates and inhomogeneities of RF transmit and receive field compared with conventional MTR imaging.5,11 Furthermore, B1 maps were used to correct for small residual higher order dependencies of the MT saturation on the local RF transmit field to further improve spatial uniformity, as suggested by Weiskopf et al.24

\[
MT_{sat} = \frac{MT_{satapp} \cdot (1 - 0.4)}{1 - 0.4 \cdot RF_{local}}
\]

where \( RF_{local} \) is the relative flip angle \( \alpha \) compared to the nominal flip angle.

Brain masks were produced using the skull stripping tool from BrainSuite (ver. 19b, http://brainsuite.org/quickstart/cse/).30 Image registration was performed using affine and symmetric diffeomorphic transforms with ANTs software (https://github.com/ANTsX/ANTs).31 Region-of-interest (ROI) masks were acquired from the labeled Allen Mouse Brain Atlas.32 One T2-weighted scan was performed for each subject at each timepoint. As registration to an atlas is time-consuming, a T2-weighted scan from only one subject at the test timepoint was chosen (the “chosen T2”) to be registered to the atlas. All other T2-weighted images from other subjects, at both timepoints, were registered to the “chosen T2.” MTR parameter maps were registered to the corresponding anatomical images (from the same subject at the same timepoint). For ROI-based analysis, the inverse transformations resulting from the preceding registration steps (MTR ➝ corresponding T2 ➝ chosen T2 ➝ atlas) were then used to bring the labeled atlas to the corresponding MT space for each subject at each timepoint. The inverse transformations, computed by ANTs for each registration step, are used to perform the opposite operation (such as deforming an image in the atlas space and producing an output in the chosen T2 space), and include inverse deformation fields and inverse affine transforms. Binary masks for each ROI were generated by thresholding the labeled atlas. Each mask was eroded by one voxel, except for the corpus callosum masks, to minimize partial volume errors within a given ROI. The binary masks were visually inspected to ensure good registration quality.

Furthermore, to perform whole brain voxel-wise analysis of all subjects across both timepoints, the data were registered to a common template. MTR maps were first registered to

Rahman et al.: Reproducibility of MTR and MTsat at 9.4 Tesla
one MTR map (the “chosen MTR”). All MTsat maps were then registered to the chosen MT space using a single transform: MTR $\rightarrow$ chosen MTR.

**Data Availability**
The test–retest dataset and in-house code to compute MTR and MTsat is available online: https://osf.io/5nwae/.

---

**Figure 1**: Example axial cross sections from a single subject. An anatomical T2-weighted image, an MT weighted (MTw) image, reference T1-weighted (T1w) and proton density weighted (PDw) images, a B1 map, and corresponding MTR and MTsat maps are shown. ROIs analyzed are overlaid on an MTw image and abbreviated as follows: CC, corpus callosum; CX, cortex; IC, internal capsule; HC, hippocampus; TH, thalamus.

**Figure 2**: Violin plots showing the distribution of MTR and MTsat at the test and retest timepoints (5 days apart) for 12 subjects in several brain regions. Unregistered data (left column) and data registered to a common template (right column) are shown. The dark black line represents the median and the red lines depict the interquartile range (25th–75th percentile). The violin plots extend to the minimum and maximum values of each metric. ROIs are abbreviated as follows: CC, corpus callosum; CX, cortex; IC, internal capsule; HC, hippocampus; TH, thalamus.
Data Analysis

**ROI-BASED AND VOXEL-WISE ANALYSIS.** ROI analysis was performed using two approaches: analysis of unregistered data and analysis of data registered to a common template. For the second approach, all MTR and MTsat maps were registered to a “chosen” MTR space, as described above.

The ROI analysis focused on five different tissue regions: corpus callosum (CC), internal capsule (IC), hippocampus (HC), cortex (CX), and thalamus (TH). For both ROI-based approaches, Bland–Altman (BA) and CV analyses were performed using the mean MTR and MTsat values from each ROI. Voxel-wise CV analysis was also performed with the registered data.

**STATISTICAL ANALYSIS.** The normalized contrast between white matter (WM) and gray matter (GM) regions was calculated as \(\frac{\text{WM (MTR or MTsat)} - \text{GM (MTR or MTsat)}}{\text{GM (MTR or MTsat)}}\), where the WM MTR and MTsat values were the average of mean ROI-based CC and IC values, and the GM MTR and MTsat values were the average of mean ROI-based HC, CX, and TH values. For the test scans of the registered data, the normalized contrast was averaged across all subjects, and an unpaired two-tailed \(t\) test was performed between MTR and MTsat contrast. For both registered and unregistered data, paired two-tailed \(t\) tests were performed to test for significant differences between ROI-based test and retest mean measurements. As there were multiple ROIs, the Bonferroni–Dunn method was used to correct for multiple comparisons.

Measurement reproducibility was explored for both ROI-based analysis and whole brain voxel-wise analysis. To mitigate partial volume errors from cerebrospinal fluid (CSF) in ROI-based analysis, voxels with MTR < 0.1 were omitted in both test and retest images. In voxel-wise analysis, voxels with MTR < 0.1, as measured on the test images, were omitted. BA analysis was performed for the ROI-based analyses to identify any biases between test and retest measurements. For both analysis techniques, the scan–rescan reproducibility was characterized using the coefficient of variation (CV). The CV

![Figure 3: Bland–Altman plots depicting biases between test and retest scans for mean MTR and MTsat values (from the ROI-based analysis). Unregistered data (left column) and data registered to a common template (right column) are shown. The solid black lines represent the mean bias, and the dotted black lines represent the ±1.96 standard deviation lines. The average of the test and retest mean values is plotted along the x-axis and the difference between the test and retest mean values is plotted along the y-axis. ROIs in the legend are abbreviated as follows: CC, corpus callosum; CX, cortex; IC, internal capsule; HC, hippocampus; TH, thalamus.](image-url)
reflects both the reproducibility and variability of these metrics, as well as provides insight into necessary sample sizes and minimum detectable effect size. CVs were calculated between subjects (bsCV) and within subjects (wsCV) to quantify the between-subject and within-subject reproducibility, respectively. The between-subject CV was calculated separately for the test and retest timepoints as the standard deviation divided by the mean value across subjects 1–12. These two CV values were then averaged for the mean between-subject CV. The within-subject CV was calculated separately for each subject as the standard deviation divided by the mean of the test and retest scans. The 12 within-subject CVs were then averaged to determine the mean within-subject CV. For both registered and unregistered data, one-way analysis of variance (ANOVA) was performed to test for significant differences between ROI-based CVs, and unpaired two-tailed t tests were performed to test for significant differences between ROI-based MTR and MTsat CVs (using the Bonferroni–Dunn method to correct for multiple comparisons).

Sample size calculations were performed based on CVs from the ROI analysis of registered data. Minimum sample sizes required to detect defined biological effects (statistically significant changes of 6%, 8%, 10%, 12%, and 14%), using both between and within-subject approaches, were determined at a 95% significance level (α = 0.05) and power of 80% (1 − β = 0.80). The defined statistically significant changes were centered around 10%, as most MT studies report changes in MTR between 15% and 30%,16,33 while some studies report more subtle changes between 5% and 10%.15,34 This is explained in greater detail in Section 3. Thus, changes smaller than 10% were considered “small” changes and changes larger than 10% were considered “large” changes.

Following the procedure presented in van Belle,35 the between-subject CVs were used to determine the sample size...
required per group to detect a defined biological effect between subjects in each ROI. Assuming paired $t$ tests, the standard deviations of the differences between test–retest mean values across subjects, were used to determine the sample size required to detect a defined biological effect within subjects in each ROI, using an online sample size calculator (UCSF Clinical & Translational Science Institute, San Francisco, CA, [https://sample-size.net/sample-size-study-paired-t-test/](https://sample-size.net/sample-size-study-paired-t-test/)). BA plots, CV calculations, and sample sizes required (using a between-subjects approach) were generated using MATLAB (ver. 2020b, Mathworks, Natick, MA). All tests of statistical significance were performed using GraphPad Prism 9 (San Diego, CA). Results were considered statistically significant at $P \leq 0.05$.

Results

Parameter Maps

Representative parameter maps are shown in Fig. 1. MTsat revealed slightly greater contrast than MTR between gray matter and white matter, which was noticeable when comparing the corpus callosum and internal capsule (white matter regions) to the surrounding gray matter. The normalized contrast between gray matter and white matter regions, averaged over all subjects, in MTsat (0.376) was significantly higher ($P < 0.0001$) than in MTR (0.226).

ROI-Based Analysis

Violin plots, as shown in Fig. 2, depict the distribution of the mean values for each metric within each ROI for the 12 subjects for both registered and unregistered datasets. Across all

![Figure 5: Voxel wise average between-subject and within-subject CV maps for MTR (top row) and MTsat (bottom row). Values for the between-subject condition represent the mean CV within each voxel averaged over the test and retest timepoints. Values for the within-subject condition represent the mean CV within each voxel averaged over all 12 subjects. ROIs are abbreviated as follows: CC, corpus callosum; CX, cortex; IC, internal capsule; HC, hippocampus; TH, thalamus.](image)

![Figure 6: Distribution of whole brain voxel-wise between- and within-subject CVs for MTR and MTsat.](image)
ROIs, for both MTR and MTsat, no significant differences were found between mean test and retest values. For registered data, adjusted $P$ values for MTR were ($P = 0.864$ [CC]; 0.315 [HC]; 0.353 [CX]; 0.285 [TH]; 0.448 [IC]), while all adjusted $P$ values for MTsat were >0.999 (paired two-tailed $t$ test). For unregistered data, adjusted $P$ values for MTR and MTsat were ($P = 0.423$ [CC]; 0.273 [HC]; >0.999 [CX]; 0.260 [TH]; 0.586 [IC]) and ($P = 0.954$ [CC]; 0.902 [HC]; 0.957 [CX]; 0.954 [TH]; 0.954 [IC]), respectively. In general, the smaller ROIs (i.e., the internal capsule) showed greater distributions, while the larger ROIs (i.e., the cortex) showed much tighter distributions. For example, for registered data at the test timepoint, MTR and MTsat had a range of 0.059 and 0.0031, respectively, in the CX, and 0.075 and 0.0046, respectively, in the IC.

BA plots, as shown in Fig. 3, revealed negligible biases, with mean biases of 0.009 and 0 for MTR and MTsat, respectively. Although not significant, MTR exhibited lower between- and within-subject CVs (2.9%–8%) compared to MTsat (4.5%–10%), as shown in Fig. 4. This range represents the minimum and maximum CVs obtained across all ROIs in the registered dataset. For registered data, adjusted $P$ values for bsCV and wsCV were ($P = 0.680$ [CX]; >0.999 [all other ROIs]) and ($P = 0.080$ [CX]; >0.999 [all other ROIs]), respectively (unpaired two-tailed $t$ test between MTR and MTsat CVs). For unregistered data, adjusted $P$ values for bsCV and wsCV were ($P > 0.999$ [all ROIs]) and ($P = 0.162$ [CC]; 0.557 [CX]; >0.999 [all other ROIs]), respectively. In general, CVs were comparable across all ROIs, as no significant differences were found between ROI-based mean CVs (one-way ANOVA). For registered data, $P$ values were ($P = 0.532$ [bsCV-MTR]; 0.255 [wsCV-MTR]; 0.789 [bsCV-MTsat]; 0.817 [wsCV-MTsat]). For unregistered data, $P$ values were ($P = 0.661$ [bsCV-MTR]; 0.353 [wsCV-MTR]; 0.890 [bsCV-MTsat]; 0.666 [wsCV-MTsat]).

**Voxel-Wise Analysis**

The voxel-wise CV maps showed very high CVs in the cerebrospinal fluid (CSF), due to the low values of MTR (<0.1) and MTsat (<0.004) in the CSF (Fig. 5). In the CSF, mean bsCV/wsCV were 44.5%/36.5% and 53.2%/39.8% for MTR and MTsat, respectively, while whole brain mean bsCV/wsCV (with a mask applied to omit CSF voxels) were significantly lower at 20.5%/11.8% and 26.2%/16.5% for MTR and MTsat, respectively. Throughout the whole brain, between- and within-subject CVs showed good

![Figure 7](image.png)

**Figure 7:** Sample size estimation using a between-subjects (a) and within-subjects (b) approach on data registered to a common template. Note that the sample size range varies between plots and sample sizes exceeding the range are not shown. ROIs are abbreviated as follows: CC, corpus callosum; CX, cortex; IC, internal capsule; HC, hippocampus; TH, thalamus.
reproducibility (CV < 20%) with 67% of voxels and 87% of voxels falling within this range for MTR bsCV and wsCV, respectively. For MTsat, 54% and 80% of voxels were within this range for bsCV and wsCV, respectively. The wsCVs were significantly lower than the bsCVs for both MTR ($P < 0.0001$) and MTsat ($P < 0.0001$). For both wsCVs and bsCVs, voxelwise MTR and MTsat CVs had significantly different variances ($F$ test). As shown in Fig. 6, the MTsat histograms revealed a wider distribution (with standard deviations of 17.3 [bsCV] and 9.3 [wsCV]) compared to the MTR histograms (with standard deviations of 16.3 [bsCV] and 8.3 [wsCV]). As observed in the ROI-based CVs, MTR exhibited lower bsCVs ($P < 0.0001$) and wsCVs ($P < 0.0001$) (with peaks at 7% and 6%, respectively) compared to MTsat (with peaks at 15% and 12%, respectively), as shown in whole brain histograms (Fig. 6).

Sample Sizes and Minimum Detectable Effect

BETWEEN SUBJECTS. To detect a minimum change of 8% in all ROIs, MTR required a sample size of 15 (Fig. 7a). In comparison, MTsat required a sample size of 25 to detect an 8% change in all ROIs. The CC and CX required smaller sample sizes, with MTR requiring 12 subjects to detect a 6% change, and MTsat requiring 15 subjects to detect an 8% change.

WITHIN SUBJECTS. As shown in Fig. 7b, in the CC and CX, small changes (6%) could be detected in MTR with six subjects per group, while MTsat could detect larger changes (8%–12%) with 12 subjects per group. For MTR, small changes (6%) could be detected in the other ROIs (IC, HC, TH) with a feasible sample size of 15. MTsat could detect larger changes (8% and greater) in all ROIs with 20 subjects per group.

Discussion

This study explored the reproducibility of MTR and MTsat at 9.4 T, and will provide insight into experiment design and sample size estimation for future in vivo MTsat imaging studies. No biases were found between repeat measurements with ROI-based analysis. MTR and MTsat were shown to be reproducible in both the mean ROI analysis and the whole brain voxel-wise analysis, with MTsat CVs being slightly higher than MTR CVs (which was not significant in ROI analysis, but significant in voxel-wise analysis). Overall, within-subject CVs were lower than between-subject CVs for both ROI-based (not significant) and voxel-wise (significant) analysis, indicating less variability within subjects on a test-retest basis.

ROI-Based Reproducibility

ROI-based reproducibility was investigated using an unregistered dataset and a dataset registered to a common template, as both unregistered and registered analysis techniques have been used in neuroimaging studies, and the difference between using either analysis technique remains sparsely explored. Recently, Klingenberg et al reported that registration significantly increased the accuracy of a convolutional neural network (CNN) to detect Alzheimer’s disease, compared to no registration.\textsuperscript{36} In our study, violin plots, BA plots, and ROI-based CV analysis revealed the same trends for both registered and unregistered ROI-based analysis approaches, which indicated that either method can be used for MT analysis. However, we recommend using the registered analysis approach, as there is only one set of ROI masks to edit, making the analysis process more time efficient. The unregistered analysis approach will also introduce inter- and intra-rater variability, due to the large number of ROI masks being edited.

The MTR ROI CVs observed in this work are consistent with MTR CVs in human studies done by Welsch et al\textsuperscript{37} and Hannila et al.\textsuperscript{38} MTsat CVs reported here are comparable to MTsat CVs in human studies at 3 T.\textsuperscript{24,25} Overall, MTsat exhibits slightly higher CVs than MTR, which may arise from noise propagation through the equations used to calculate MTsat, as described by Olsson et al.\textsuperscript{23} A noticeable increase in MTsat CVs compared to MTR CVs, in the HC and CX, may be due to low MTsat values in these regions.

Voxel-Wise Reproducibility

Voxel-wise CV trends were comparable to ROI-based CV trends. Voxel-wise CV maps revealed a more noticeable increase in CVs in the superior–inferior direction of the brain in MTR, compared to MTsat. This can be related to the inherent compensation of flip angle inhomogeneities in MTsat.\textsuperscript{5}

Sample Size and Minimum Detectable Effect

The CC consistently exhibited the smallest required sample sizes, which can be related to the lower variability of myelin content in the CC, compared to the gray matter ROIs.\textsuperscript{39} Interestingly, the CC and IC (the white matter regions) required similar sample sizes to detect the same changes in MTsat (using both between- and within-subject approaches), but not in MTR, which required larger sample sizes to detect changes in the IC. This may stem from the better contrast seen between the IC and gray matter in MTsat, compared to MTR, which arises from MTsat being less susceptible to inhomogeneities of the transmitted field and more independent of T1-weighting.\textsuperscript{5,21}

Most MT studies report changes in MTR between 15% and 30%, with some studies reporting more subtle changes
between 5% and 10%. In a cuprizone demyelination model in mice, MTR decreased by 15% and 30% at 4 weeks and 6 weeks of cuprizone administration, respectively. In an ischemic injury model in mice, MTR decreased by 30% in the corpus callosum of injured mice compared to controls. In a closed head traumatic brain injury model in mice, MTR in the corpus callosum decreased by 10% from baseline at 1-day post-injury. A post-mortem study revealed a 10% decrease in MTR between normal-appearing white matter and multiple sclerosis lesions. In a recent multiple sclerosis work, MTR was able to differentiate between patients with and without cognitive impairment, showing a 7% decrease in patients with cognitive impairment.

MTR can detect changes on the order of 15%–30% (such as the changes found in the cuprizone demyelination model) with small sample sizes (N = 6). With disease and injury models resulting in less drastic changes to myelin content, our findings suggest that MTR and MTsat can detect smaller changes with feasible preclinical sample sizes. Thiessen et al showed that when there is an 80% reduction in myelinated axon density, MTR only decreases by ~30% (because it is thought that inflammation has a competing effect on MTR). So, a 2-fold difference in myelin will result in at least a 15% change in MTR. However, as MTsat provides greater specificity to myelin, a 2-fold difference in myelination should translate to a larger percent change in MTsat.

**Limitations**

Although a volume coil is more appropriate for structural imaging as it provides stable SNR throughout the brain, this study used a transceiver surface coil. The voxel-wise CV maps showed that between-subject and within-subject CVs were slightly higher toward the inferior region of the brain. However, the increase in CV was subtle and as shown in ROI-based analysis, the CVs of ROIs located in inferior regions of the brain (such as the IC) were comparable to the ROIs closer to the surface coil. Moreover, MTsat maps were comparable to MTsat maps acquired by Boretius et al in the mouse brain at 9.4 T. This shows the feasibility of acquiring MTR and MTsat data using a surface coil, which may be useful in studies in which MT imaging is combined with other methods that require a surface coil or in inherently low SNR methods that would benefit from a surface coil, such as diffusion MRI. Recent preclinical investigations have included a combination of MT imaging and diffusion MRI. Moreover, the findings in this study will complement a recent test–retest reproducibility study in advanced diffusion MRI techniques in mice at 9.4 T.

Although the sample size was chosen to reflect similar sample sizes used in other pre-clinical imaging studies, the small number of subjects is another limitation in this work. Nevertheless, we believe that these results are valuable and useful for the MT imaging community. In the statistical analyses, it should be noted that for the within-subject calculation of CV, the standard deviation was determined from only two data points (the test and retest conditions). As a result, the standard deviation may not accurately represent the spread of data within the population, leading to an unknown bias in the resulting within-subject CV.

**Conclusion**

We demonstrated that MTR and MTsat were reproducible in both ROI-based analysis, which includes both registered and unregistered analysis techniques, and voxel-wise analysis. Importantly, MTsat exhibited comparable reproducibility to MTR, and could detect small changes (<10%) with sample sizes of 15–20, while providing better contrast and maintaining a feasible scan time.

**Acknowledgments**

The authors thank the Canada First Research Excellence Fund (BrainsCAN—https://brainscan.uwo.ca/); the New Frontiers in Research Fund (NFRFE-2018-01290—https://www. sshrc-crsh.gc.ca/funding-financement/nfrf-fnfr/index-eng.aspx), awarded to Corey A. Baron; the Natural Sciences and Engineering Research Council of Canada: Canada Graduate Scholarships—Master’s Program (NSERC-CGS M), awarded to Naiia Rahman; and the Ontario Graduate Scholarship (OGS), awarded to Naiia Rahman, for their funding contributions.

**References**

1. Schmierer K, Scaravilli F, Altmann DR, Barker GJ, Miller DH. Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain. Ann Neurol 2004;56:407-415.
2. Gass A, Barker GJ, Kidd D, et al. Correlation of magnetization transfer ratio with clinical disability in multiple sclerosis. Ann Neurol 1994;36:62-67.
3. Sled JG. Modelling and interpretation of magnetization transfer imaging in the brain. Neuroimage 2018;182:128-135.
4. Mottershead JP, Schmierer K, Clemence M, et al. High field MRI correlates of myelin content and axonal density in multiple sclerosis: A post-mortem study of the spinal cord. J Neurol 2003;250:1293-1301.
5. Helms G, Dathe H, Kallenberg K, Dechent P. High-resolution maps of magnetization transfer with inherent correction for RF inhomogeneity and T1 relaxation obtained from 3D FLASH MRI. Magn Reson Med 2008;60:1396-1407.
6. Harkins KD, Xu J, Dula AN, et al. The microstructural correlates of T1 in white matter. Magn Reson Med 2016;75:1341-1345.
7. Stübner C, Morawski M, Schafer A, et al. Myelin and iron concentration in the human brain: A quantitative study of MRI contrast. Neuroimage 2014;93:95-106.
8. Sui YV, Bertisch H, Lee H-H, et al. Quantitative macromolecular proton fraction mapping reveals altered cortical myelin profile in schizophrenia spectrum disorders. Cereb Cortex Commun 2021;2:1-12.
9. Clarke MA, Lakhani DA, Wen S, et al. Perilesional neurodegenerative injury in multiple sclerosis: Relation to focal lesions and impact on disability. Mult Scler Relat Disord 2021;49:102738.
10. Khodanovich M, Pishchelko A, Glazcheva V, et al. Quantitative imaging of white and gray matter remyelination in the Cuprizone demyelination model using the macromolecular proton fraction. Cell 2019;8:1204.

11. Helms G, Dathe H, Dechent P. Modeling the influence of TR and excitation flip angle on the magnetization transfer ratio (MTR) in human brain obtained from 3D spoiled gradient echo MRI. Magn Reson Med 2010;64:177-185.

12. Lema A, Bishop C, Malik O, et al. A comparison of magnetization transfer methods to assess brain and cervical cord microstructure in multiple sclerosis. J Neuroimaging 2017;27:221-226.

13. Hagiwara A, Horii M, Kamagara K, et al. Myelin measurement: Comparison between simultaneous tissue Relaxometry, magnetization transfer saturation index, and T1 w/T2 w ratio methods. Sci Rep 2018;8:10554. https://doi.org/10.1038/s41598-018-28852-6.

14. Giorgetti E, Obrecht M, Ronco M, et al. Magnetic resonance imaging as a biomarker in rodent peripheral nerve injury models reveals an age-related impairment of nerve regeneration. Sci Rep 2019;9:1-18.

15. Tu TW, Williams RA, Lescher JD, Jikaria N, Turtzo LC, Frank JA. Radiological-pathological correlation of diffusion tensor and magnetization transfer imaging in a closed head traumatic brain injury model. Ann Neurol 2016;79:907-920.

16. Fatemi A, Wilson MA, Phillips AW, et al. In vivo magnetization transfer ratio MRI shows remyelination in an ischemic mouse model of periventricular leukomalacia. J Cereb Blood Flow Metab 2011;31:2009-2018.

17. Pohmann R, Shajan G, Balla DZ. Contrast at high field: Relaxation times, magnetization transfer and phase in the rat brain at 16.4 T. Magn Reson Med 2011;66:1572-1581.

18. Xu J, Li K, Zu Z, Li X, Gochberg DF, Gore JC. Quantitative magnetization transfer imaging of rodent glioma using selective inversion recovery. NMR Biomed 2014;27:253-260.

19. Li K, Li H, Zhang XY, et al. Influence of water compartmentation and heterogeneous relaxation on quantitative magnetization transfer imaging in rodent brain tumors. Magn Reson Med 2016;76:635-644.

20. Wu TL, Byun NE, Wang F, et al. Longitudinal assessment of recovery after spinal cord injury with behavioral measures and diffusion, quantitative magnetization transfer and functional magnetic resonance imaging. NMR Biomed 2020;33:1-14.

21. Boretius S, Dechent P, Frahm J, Helms G. Magnetization transfer mapping of myelinated fiber tracts in living mice at 9. 4 T. Proc Int Symp Magn Reson Med 2010;1252:2009.

22. Field AS, Samsonov A, Alexander AL, Mosahebi P, Duncan ID. Conventional and quantitative MRI in a novel feline model of demyelination and endogenous remyelination. J Magn Reson Imaging 2019;49:1304-1311.

23. Olsson H, Andersen M, Wirestam R, Helms G. Mapping magnetization transfer saturation (MTsat) in human brain at 7T: Protocol optimization under specific absorption rate constraints. Magn Reson Med 2021;86:2562-2576. https://doi.org/10.1002/mrm.28899.

24. Weiskopf N, Suckling J, Williams G, et al. Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: A multi-center validation. Front Neurosci 2013;7:1-11.

25. Leutritz T, Seif M, Helms G, et al. Multiparameter mapping of relaxation (R1, R2*), proton density and magnetization transfer saturation at 3 T: A multicenter dual-vendor reproducibility and repeatability study. Hum Brain Mapp 2020;41:4232-4247.

26. Xu X, Cowan M, Beraldo F, et al. Repetitive mild traumatic brain injury in mice triggers a slowly developing cascade of long-term and persistent behavioral deficits and pathological changes. Acta Neuropathol Commun 2021;9:1-27.

27. Gazzinelli LM, Mellerup M, Wang T, et al. White matter changes caused by mild traumatic brain injury in mice evaluated using neurite orientation dispersion and density imaging. J Neurotrauma 2020;37:1818-1828.

28. Wright DK, Brady RD, Kannaksh A, et al. Repeated mild traumatic brain injuries induce persistent changes in plasma protein and magnetic resonance imaging biomarkers in the rat. Sci Rep 2019;9:1-13.

29. Rahman N, Xu K, Omer M, Budde MD, Brown A, Baron CA. Test-retest reproducibility of in vivo oscillating gradient and microscopic anisotropy diffusion MRI in mice at 9.4 Tesla. PLoS One 2021;16:e0255711.

30. Shattuck David W, Leahy Richard M. BrainSuite: An automated cortical surface identification tool. Medical Image Analysis. 2002;6(2):129-142. http://doi.org/10.1016/s1361-8415(02)00054-3.

31. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 2011;54:2033-2044.

32. Lein ES, Hawrylycz MJ, Ao N, Ayres M, Benzinger T. Genome-wide atlas of gene expression in the adult mouse brain. Nat Artic 2007;445:168-176.

33. Thiessen JD, Zhang Y, Zhang H, et al. Quantitative MRI and ultrastructural examination of the cuprizone mouse model of demyelination. NMR Biomed 2013;26:1562-1581.

34. Campanhola K, Pitombeira M, Rimkus C, et al. Myelin imaging measures as predictors of cognitive impairment in MS patients: A hybrid PET-MRI study. Mult Scler Relat Disord 2021;57:1-9. https://doi.org/10.1016/j.msard.2021.103331.

35. van Belle G. Sample size. Statistical rules of thumb. Hoboken, NJ: John Wiley and Sons; 2018. p 27-51.

36. Klingenberg M, Stark D, Eitel F, Ritter K. MRI image registration considerably improves CNN-based disease classification. In: Proceedings of the 4th International Workshop, Machine Learning in Clinical Neuroimaging, Strasbourg, France, 2021. p 44-52.

37. Welsch GH, Apprich S, Zbyn S, et al. Biochemical (T2, T2* and magnetisation transfer ratio) MRI of knee cartilage. Feasibility at ultra-high field (7T) compared with high field (3T) strength. Eur Radiol 2011;21:1136-1143.

38. Hannila I, Susanna Räinä S, Tervonen O, Ojala R, Nieminen MT. Topographical variation of T2 relaxation time in the young adult knee cartilage at 1.5 T. Osteoarthr Cartil 2009;17:1570-1575.

39. Erö C, Gewaltig MO, Keller D, Markram H. A cell atlas for the mouse brain. Front Neuroinform 2018;12:1-16.