UNLOCKING THE PROGNOSTIC POTENTIAL OF QUANTITATIVE MAGNETIC RESONANCE IMAGING FOR MULTIPLE SCLEROSIS THROUGH A MULTIMODAL APPROACH

By

Kaihim Wong

A Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

of the University of Manitoba

in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Department of Physics and Astronomy

University of Manitoba

Winnipeg, Manitoba

Abstract

Magnetic resonance imaging (MRI) is the most widely used paraclinical tool for diagnosing and monitoring multiple sclerosis (MS). However, given known heterogeneity in pathophysiology, even with robust conventional MRI (cMRI) measures such as white matter lesion (WML) number and volume, the relationship with MS progression is only modest, hindering its prognostic use in MS. One major burden is the lack of characterization of the white matter (WM) space.

With recent advances in quantitative MRI (qMRI) techniques, the enhanced sensitivity and specificity of individual qMRI have shed light on its clinical potential. But none of them has yet been validated as a stand-alone biomarker for MS, and their clinical value has not been clearly demonstrated. Therefore, this thesis investigates the potential of a multi-modal qMRI approach for WM characterization and prognostic applications. The thesis work is split into four different phases. The first phase has explored the multimodal relationship between qMRI and volumetric measures across three brain regions relevant to MS – WML, normal appearing white matter (NAWM) and deep gray matter (DGM). The results have demonstrated the clinical relevance of qMRI and the feasibility of the multimodality approach. The second and third phases were then motivated to provide a deeper characterization of the WM by defining finer regions related to WML and perilesional NAWM, both cross-sectionally and longitudinally. These two phases have demonstrated that qMRI metrics not only are sensitive to WM heterogeneity, but also provide information about future WM transitions. The final phase tasked a neural network model in predicting WML evolution using separated and combined inputs from qMRI and cMRI, and assessed the performance. The positive result strongly suggests the potential of qMRI and cMRI to predict WML evolution.

These findings indicate richer levels of WM heterogeneity than the binary distinction between WML and NAWM, which can be sensitized using qMRI. While a multimodal qMRI approach helps characterize WM, cMRI also demonstrates the capability to predict WML evolution. Overall, this work sheds light on the clinical value of qMRI in characterizing the WM space, demonstrates the extended utility of cMRI, and lays a solid foundation for future studies in WML prediction, ultimately helping improve MS prognosis.

Acknowledgements

The past five years of my Ph.D. have been one of the most rewarding experiences of my life. Putting together the thesis and dissertation was a remarkable accomplishment for me, and I believe this is true for everyone else. This path offered me the capability to view and think more critically. In the meantime, I received an enormous amount of academic support that helped bring this work to the next level. This support was translated into consistent fuel for my studies (and will always be, to navigate through my lifelong journey).

Here, I would first like to acknowledge my supervisor, mentor, and friend, Dr. Chase Figley. Thank you for recognizing my potential and bringing me into this field. I deeply appreciate the guidance and support you provided throughout the M.Sc.-Ph.D. journey. The incredible duality of professionalism and approachability in you has brightened and nourished my training environment, and I believe that is why all your trainees enjoyed immersing themselves (including me) the most. Second, I would like to acknowledge my internal committee members, Dr. Melanie Martin and Dr. Stephen Pistorius. Your insights set a much wider stage for the projects and made their scope more well-rounded. At the same time, the care that you both provided to the University of Manitoba medical physics community is truly appreciated by the local student community. I would also like to thank everyone on my committee and my external committee member, Dr. Kolind Shannon, for your valuable comments and suggestions that helped elevate the quality of the dissertation.

I acknowledge the effort from all angles to launch and drive the CCOMS study initiative: all the study participants who volunteered and shared their data and all the CCOMS researchers who worked collaboratively to design the protocol and collect the data. I would like to thank all CCOMS researchers and my lab members for your feedback at different stages of this work. Especially, I greatly appreciate Ms. Teresa Figley for your patience and assistance during the

dataset walkthrough, as well as for sharing constructive comments on knowledge translation in presentations and graphic design in the conference proceedings.

I am grateful for the support and opportunities provided by the staff, faculty, and students in the Department of Physics & Astronomy at the University of Manitoba and the medical physicists at CancerCare Manitoba. There were plenty of volunteering opportunities with the student association and outreach committee that I had participated in and learnt a lot from. My colleagues in the local medical physics student groups were extremely helpful when I reached out for assistance. I would like to especially mention Ms. Susan Beshta for all your hard work in tirelessly keeping every student on the same page and up to speed throughout their graduate studies, and promptly helping us when needed.

During my studies, I was honoured to be part of several external interdisciplinary groups, including the Manitoba Multiple Sclerosis Research Centre, the endMS trainee group, and the Canadian Organization of Medical Physicists Student Council, which have prepared me to be a more well-rounded researcher and helped me thrive and succeed. In addition, I was fortunate to be awarded multiple major scholarships and funding at different stages to support my study, Canadian Graduate Scholarship – Master's (2022), MS Canada summer school supports (2022-2024), Research Manitoba (2024) and Hilary Kaufman Lerner Memorial Travel Fund (2024-2025). These scholarships and funding substantially eased my financial burden, and more importantly, motivated me to push through the limitations.

Unquestionably, this Ph.D. journey has widened my horizons and given me the courage to continue seeking the path to break the boundaries of knowledge throughout the rest of my life, while staying curious, regardless of the settings or scopes.

Dedication

Outside of research, all my surroundings had mentally supported and provided lots of unconditional love along the way. Thank you, the most important person in my life – my wife, for trusting me since the day we met, standing strong with me, easing my stress in all aspects of my life, bearing with my hectic schedule, and always putting a smile on my face. My gratitude for you extends far beyond what words can convey. Alongside, my lovely two cats – Anthony and Libby – showed an incredible patience while I was busy working. Your adorable behaviour always makes me forget about the nerves and stress at work. I would also like to shout out to my parents for your guidance and unconditional support throughout the journey. I am grateful for the decision you made to send us abroad, and it was never easy. A special mention to my sister, who is putting up an impactful work for her Ph.D. in biochemistry, for your peer support, feedback on my writings, and spending time doing dry runs of presentations. I have never had a single doubt about your professionalism and dedication in translation, and I believe they will lead to great success in the future. Interestingly, I would not have heard of the medical physics program if I had never met my academic consultant, Mr. Ron Butterfield. Thank you for your advice and guidance. I really wish you could witness this. And before I started my graduate studies, I was fortunate to work with my former summer research supervisors, Dr. Esmat Elhami, Dr. Chase Figley and Dr. Melanie Martin. You were the deciding factor of graduate study, research and an imaging career. Last but not least, friendships that I made at the local, national and international levels, both within and outside academia, including (but not limited to) music, badminton, fitness and skiing, have kept me up with the brightest mind and attitude to triumph over challenges ahead of me and to embrace being a part of this wonderful world.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AD	Axial diffusivity
AFIL	Automatic Follow-up of Individual Lesion
AI	Artificial Intelligence
AUC	Area Under the Curve
BBB	Blood Brain Barrier
BH	Benjamini Hochberg
CanProCo	Canadian prospective cohort
CAT	Computational Anatomy Toolbox
CCOMS	Comorbidity and Cognition in Multiple Sclerosis
cMRI	Conventional magnetic resonance imaging
CNS	Central Nervous System
COSMOS	Calculation Of Susceptibility using Multiple Orientation Sampling
CoV	Coefficient of Variation
CPU	Central Processing Unit
CSF	Cerebrospinal fluid
CVS	Central vein sign
DAWM	Diffusely abnormal white matter
DEM	Disease evolution map
DGM	Deep gray matter
DL	Deep learning
DMDR	Different metrics different ROIs
DMSR	Different metrics same ROIs
DMT	Disease modifying therapy
DSC	Dice similarity coefficient
DTI	Diffusion tensor imaging
EDSS	Expanded disability status scale
EvePM	Everything Parcellation Map
FA	Fractional Anisotropy
FDR	False discovery rate
FLAIR	FLuid-Attenuated Inversion Recovery
FOV	Field-of-view
FSL	Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library
GAN	Generative Adversarial Network
GE	Gradient echo
Gd	Gadolinium
GM	Gray matter
GPU	Graphics processing unit
GRASE	GRadient And Spin Echo
gT2	Geometric mean T2
IEWF	Intra- and extra-cellular water fraction
IQR	Interquartile range
MATLAB	MATrix LABoratory
MD	Mean diffusivity
ML	Machine learning

MNI	Montreal Neurological Institute
MR	Magnetic resonance
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MTR	Magnetic transfer ratio
MWF	Myelin water fraction
MWI	Myelin water imaging
NAWM	Normal appearing white matter
PRL	Paramagnetic rim lesion
pNAWM	Perilesional normal appearing white matter
pwMS	People with multiple sclerosis
qMRI	Quantitative magnetic resonance imaging
QSM	Quantitative susceptibility mapping
QSMART	Quantitative susceptibility mapping artifact reduction technique
qT1	Quantitative T1
R2*	Apparent transverse relaxation rate
RD	Radial diffusivity
ROC	Receiving Operating Characteristic
ROI	Region of Interest
SE	Spin echo
seg-gradCAM	Segmentation gradient-weighted class activation map
SEL	Slowly expanding lesion
SMDR	Same metric different ROIs
sMRI	Structural magnetic resonance imaging
SPM	Statistical Parametric Mapping
Std	Standard deviation
SWI	Susceptibility weighted imaging
T1	Longitudinal relaxation time
T1-w	T1-weighted
T1w/T2w	T1-weighted-by-T2-weighted
T2	Transverse relaxation time
T2*	Apparent transverse relaxation time
T2’	Transverse relaxation due to magnetic inhomogeneity effects
T2-w	T2-weighted
TE	Echo time
TI	Inversion time
TR	Repetition time
Vol	Volumetric measures
WM	White matter
WML	White matter lesion
XAI	Explainable AI

Chapter 1

Introduction

1.1 Purpose(s)

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system (CNS) (Compston & Coles, 2002). It affects more than 2.8 million people around the world (Walton et al., 2020a). MS exhibits substantial heterogeneity in phenotypes and symptoms, genetics, pathology and immunology (Disanto et al., 2011). MS currently lacks a cure. But fortunately, thanks to advances in research, more and more cellular pathways have been studied, therapeutic targets have been identified (Jiang et al., 2024), and a handful of these are being approved as disease modifying therapies with three ways to administer (i.e., oral, injectable and intravenous)

Retrospective studies have shown that all forms of MS have underlying progression over time, and an estimate is that 70-80% of the relapsing-remitting form of MS progresses to the progressive form of MS after 10 years (Noseworthy et al., 2000; Rejdak et al., 2010). Nonetheless, studies show that early intervention can significantly alleviate the risk of subsequent progression (Filippi et al., 2022; Freeman et al., 2022; Giovannoni et al., 2016; Singer et al., 2024). Therefore, this has underpinned the importance of MS prognosis.

MS prognosis is not straightforward. In fact, no single measure or test can accurately predict MS progression when used alone. Similar to MS diagnosis, prognosis relies on a holistic approach by examining a combination of demographic and environmental factors (e.g., age, sex and lifestyle), clinical examination results (e.g., expanded disability status scale, recovery status), and findings from paraclinical tools and tests (e.g., lab measures (e.g., level of neurofilament light chain), magnetic resonance imaging (MRI), optical coherence tomography, lumbar puncture)

(Rotstein & Montalban, 2019). Based on the consensus and neurologists' experience, having multiple red flags, including both onset and follow-up changes in features may indicate a higher risk in the person with MS who will later progress (Briggs et al., 2019; Degenhardt et al., 2009; Huang, 2025; Rotstein & Montalban, 2019). However, no clear guidelines or robust clinical thresholds are yet available to inform clinical decisions. Therefore, neurologists are required to have consistent periodic follow-ups with people with MS, with reassessments of progression and treatment responses, and personalization based on the profile and trajectory of the person with MS. Still, aside from being resource-heavy, with frequent re-visits, this approach is largely reactionary with (delayed) treatments often chasing new symptoms.

Among all paraclinical tools, MRI is an indispensable component in MS diagnosis and monitoring (Rocca et al., 2024). It is mainly used to identify and characterize lesion profiles in the brain and spinal cord, thanks to its exceptional soft-tissue contrast and flexible imaging sequence designs (Rocca et al., 2024). By comparing MRIs over time, one can observe the evolution of the brain, especially of lesions. This information is really helpful for neurologists as lesion numbers, volume, location and features have been extensively used as radiological guidelines to help assess MS severity (Montalban et al., 2025). Therefore, predicting lesion evolution in advance could be extremely helpful to neurologists in making timely decisions and even prospectively tailoring individual treatment plans. The current use of conventional MRI (cMRI) remains qualitative. But quantitative MRI (qMRI), a rapidly evolving field with significant potential recognized in MS (Granziera et al., 2021; Tranfa et al., 2022), may be in our favour by providing better characterization of brain tissues and invaluable insights into early signs that inform longitudinal changes. Therefore, this thesis project uses a series of data-driven analyses to better understand qMRI's role in tissue characterization in the brain, and to tease out qMRI’s prognostic potential

for predicting disease progression: Specifically, whether qMRI can predict the evolution of tissue changes in advance.

1.2 Thesis organization

This thesis is compiled in a manuscript-style that begins with some general background, allowing audiences to familiarize themselves with key concepts of MS, MRI, qMRI, and the Comorbidity and Cognition in Multiple Sclerosis (CCOMS) dataset (Chapter 1). This thesis has five body chapters (Chapters 2-6), each of which presents an original and unique empirical contribution (written as a standalone manuscript), and Section 1.2.2 outlines the roadmap for how the analyses are connected to achieve the overarching goal of the thesis. The thesis then concludes with a general discussion of the results and a future outlook (Chapter 7).

1.2.1 Background sections

The first section of background covers MS as a disease. This section focuses on answering the following questions: “What is MS?” and “How heterogeneous is MS?” The next section introduces MRI and how it is typically used to visualize different MS lesions before, then provides a brief introduction to how MRI works, followed by a showcase of qualitative MRI sequences routinely used in clinical settings and the qMRI sequences used in this study. After that, the current use of MRI and the gap between MS and MRI are highlighted. The last section outlines the key features of the CCOMS dataset used throughout the remainder of the thesis.

1.2.2 Preamble of the analyses and the linkages between them

The main pillar of this thesis consists of five sequentially connected chapters, each a manuscript. Chapter 2 and Chapter 3 present two closely connected manuscripts (i.e., part 1 and part 2) sharing a similar methodology and study design – "*Investigating sensitivity and correlation between microstructural MRI measures in multiple sclerosis white matter lesions, normal appearing white matter and deep gray matter regions*" and "*Volumetric MRI measures correlate with quantitative MRI measures in multiple sclerosis between white matter lesions, normal appearing white matter and deep gray matter regions*". Overall, they summarize a pair of complementary results. The overarching objective of the analysis is to explore the relationship between different qMRI metrics (including signal intensities from 10 qMRI maps, as well as volumetric measurements) in the three regions of interest (ROIs): Normal appearing white matter (NAWM), white matter lesion (WML), and deep gray matter (DGM) by using pairwise Spearman rank correlations. A number of relationships were shown to be significant after correcting for multiple comparisons. The former manuscript (Chapter 2) characterizes the relationships between qMRI metrics and between regions by decomposing the correlation matrix, and the latter manuscript (Chapter 3) focuses on the fact that many qMRI metrics correlate with volumetric measures, indicating the impact of microstructural integrity on the macroscopic characterization of tissue. Together, the results demonstrate the clinical potential of qMRI and raise awareness of information redundancy when multiple qMRI modalities are used simultaneously. However, to further characterize white matter (WM) space in MS, particularly around WML, the relatively large ROI definitions in those studies drive the next study to break the larger, whole-brain WML masks into finer, localized ROIs within and surrounding individual WMLs.

Chapter 4 presents the manuscript, "*Mapping quantitative MRI changes as a function of distance from the boundaries of multiple sclerosis white matter lesions*", which investigates qMRI

signal intensities gradient as a function of distance from WML boundaries (both within WMLs and into surrounding NAWM) using concentric shells delineated using the individual WML morphology. The findings of this study highlight qMRI's sensitivity to tissue heterogeneity in both NAWM and WML, particularly near the WML boundaries (e.g., lesion rim, lesion periphery). The findings of this study suggest that NAWM and WML are complementary in terms of qMRI information. Together, these findings motivate the next study to investigate whether this spatial sensitivity in baseline qMRI maps carries any longitudinal information about WML tissue transitions.

Chapter 5 presents the manuscript, "*Comparing baseline qMRI metrics in stable & transitioning regions in and around multiple sclerosis white matter lesions*". In this work, baseline qMRI metrics in the transitioning and stable WML ROIs, defined by the longitudinal WML masks, are compared. The positive findings of baseline qMRI metrics in these regions shed light on the potential of qMRI for predicting WML evolution. As implied, the final body chapter (Chapter 6) takes the findings of this study as a significant stepping stone toward a proof-of-concept study to test the idea of predicting WML evolution using a machine learning (ML) method.

Chapter 6 presents the manuscript, "*Can multiple sclerosis white matter lesion evolution be predicted using baseline MRI? A promising attempt using a two-year follow-up dataset and a simple U-Net architecture*". cMRI (i.e., T1-weighted and T2-FLAIR) is also recruited in this study, alongside qMRI metrics, to serve as a control for the deep learning model. The main architecture of ML is a generic 3D U-Net architecture. The result clearly indicates the ML's potential to learn the follow-up spatiotemporal changes in the WM space from both cMRI and qMRI modalities at baseline. Surprisingly, the cMRI-based model has demonstrated a comparable performance to the qMRI-based model. Even when both modalities are combined (i.e., concatenated along the feature

dimension), performance is not significantly different from that of the cMRI-based model. This further highlights the study's scalability, and demonstrates a lower barrier to clinical applicability, given the high accessibility of cMRI.

Chapter 7 is the concluding chapter of the thesis, which begins by presenting a summary of the overall findings and their significance, along with certain limitations of the study's design. This chapter concludes by imagining the outlook and outlining the key future directions for follow-up work, if given enough resources and time.

1.3 Introduction to MS

The etiology of MS remains unknown, although many hypotheses have been proposed (Haki et al., 2024). However, two vital systems, the immune system and the CNS, are compromised in MS through a sequential series of key events: Autoreactive T cells first trigger inflammatory responses in the immune system. It promotes the migration of autoreactive T cells and other proinflammatory cells from the lymphatic stream toward the interface between the blood supply and neuronal space – i.e., the blood brain barrier (BBB). In addition, the inflammatory response disrupts the BBB, allowing autoreactive lymphocytes to encounter and directly attack CNS oligodendrocytes and myelin. Consequently, this leads to demyelination, axonal loss, and secondary inflammatory responses in the CNS. Over time, these processes lead to the accumulation of sclerotic lesions throughout the brain and spinal cord: hence the name “multiple sclerosis”, which is the common translation from the original French designation “sclérose en plaques” (originally coined by Charcot in 1868 (Murray, 2005; Pearce, 2005).

Historically, MS is well recognized for its heterogeneity, and it can be regarded as an aftermath of the mixed stages of any of the aforementioned inflammatory processes, and many

other factors as previously discussed (Disanto et al., 2011). Individuals with MS experience a wide range of functional disabilities, including (but not limited to): vision loss, mobility difficulties, spasticity, and at the same time showing symptoms of fatigue, memory loss and cognitive decline, bowel dysfunction and chronic comorbidities like diabetes (Cosh & Carslaw, 2014; Marrie, 2017). Depending on the location of the lesions (e.g., cortical gray matter (GM) vs. WM vs. DGM, cerebrum vs. cerebellum, brain vs. spinal cord), the stages of lesions (e.g., chronic vs. acute, active vs. inactive), and the combination of both, various levels and types of clinical presentation that are being exhibited can be quite different from case to case. Currently, MS is mainly classified into two broad phenotypes: relapsing and remitting MS and progressive MS. However, more and more studies suggest that MS can be a continuous spectrum rather than a discrete stratification (Giovannoni, 2018; Lublin et al., 2014; T. L. Vollmer et al., 2021). Nevertheless, MS pathology, neuroplasticity, and other factors (e.g., genetics and environmental factors) contribute to MS heterogeneity and are suggested to be key indicators for stratifying MS phenotypes (Kuhlmann et al., 2023).

It then becomes clear that a better characterization of MS heterogeneity can benefit the field in many different ways. Classifying the relevant pathways and integrating them into a single framework will fundamentally bridge the gap between research and pathology, forming a complete system. And eventually, this could facilitate more efficient decision-making for MS diagnostics (a holistic approach), subsequent treatment escalation and progress monitoring to improve outcomes.

1.4 Introduction to imaging MS lesions with MRI

Adopting diagnostic imaging technologies in the health care system evidently provides us with a new set of lenses to view anatomy (and anatomical changes) to complement clinical assessments,

yielding a more comprehensive and holistic view of each patient. Ever since MRI was introduced for MS diagnosis and monitoring, it has achieved great success due to its ability to non-invasively provide 3D images with extraordinary soft tissue contrast based on regional T1 (longitudinal relaxation time) and T2 (transverse relaxation time) properties (see next section for how an MRI works in general). In this way, clinicians can identify MS lesions from routinely collected MRI scans: for example, T2-weighted (T2-w) FLuid Attenuated Inversion Recovery (FLAIR) and gadolinium (Gd)-enhanced T1 weighted (T1-w) images.

MS lesions vary in size and shape, but most appear ovoid (Filippi et al., 2019). As alluded to earlier, they are also located in different regions. Commonly, these regions are the brain and spinal cord areas, where WM and GM lesions are found (Filippi et al., 2019). Typical WMLs are seen in the juxtacortical, periventricular, infratentorial, and optical fibre regions of the brain, and in the brain stem and spinal cord (Filippi et al., 2019). Those lesions can often be revealed on T2-w MR images and exhibit hyperintensity compared to the neighbouring WM space, representing the vast majority of demyelination, neurodegeneration, edema, gliosis and general inflammation (Barkhof & Walderveen, 1999; Frohman et al., 2006; Moore et al., 2000). There are MR sequences that are being used in targeting specific lesions, and these lesions signify stages of the inflammatory process. Gd-enhanced and T1-w hypointense lesions are two classic examples. Gd-enhanced lesions are active, acute lesions that represent the current disruption of the BBB and the infiltration of contrast agent leaking into the CNS (Grossman et al., 1986; Rovira et al., 2023; Saade et al., 2018). T1-w hypointense lesions are usually chronically inactive, representing a more severely damaged site with more extensive axonal loss and destruction (Sahraian et al., 2009a). Chronic active lesions are also important and specific to MS, usually indicating disease progression (Dal-Bianco et al., 2024). One of these is the slowly expanding lesion, a type of lesions

that grow over time (Elliott et al., 2019). Susceptibility-weighted imaging (SWI), a method sensitive to field inhomogeneity, can also reveal paramagnetic rim lesions by identifying the hypointense ring around lesions, which signifies iron accumulation induced by possible microglia and other dysfunctional cells crowded around the lesion (Meaton et al., 2022). SWI is also very useful for detecting lesions with the central vein sign (where the vasculature is located at the centre of the lesion), which demonstrates high specificity for MS in diagnosis (Cagol et al., 2024; Sati et al., 2016).

1.5 An overview of MRI

Unless otherwise noted, information in this section (i.e., MRI principles) and the next section (1.6, especially about routinely used MRI modalities) have been summarized from the MRI textbooks (with greater depths and scopes), and the sources are:

1. Bernstein, M. A., King, K. F., & Zhou, X. J. (2004). *Handbook of MRI pulse sequences*. Elsevier.
2. Brown, R. W., Cheng, Y.-C. N., Haacke, E. M., Thompson, M. R., & Venkatesan, R. (2014). *Magnetic resonance imaging: physical principles and sequence design*. John Wiley & Sons.
3. Dowsett, D., Kenny, P. A., & Johnston, R. E. (2006). *The physics of diagnostic imaging second edition*. CRC Press.

Before diving into different MRI modalities to better understand the origin and physics of each MRI sequence and their contrasts, it is fair to first at least briefly discuss how MRI works, what an MRI sequence is, and how it fits into MRI image acquisition.

MRI consists of three general categories of magnetic fields: a static background magnetic field or $\boldsymbol{B_0}$, a gradient magnetic field or $\boldsymbol{G}$ and a radio frequency rotating magnetic field or $\boldsymbol{B_1}$. Each of these magnetic fields has a key function. Here, the italic-bold font indicates that the variables are in vector form. First, due to a combination of the quantum spin 1/2 property under the influence of a magnetic field (i.e., Zeeman's effect) and the relatively high abundance of hydrogen (^{1}H), $\boldsymbol{B_0}$ normalizes and aligns the ^{1}H-nuclei spins population from randomly distributed states to an ordered, equilibrium state along the direction of the field, forming a net magnetization, $\boldsymbol{M}$. The role of $\boldsymbol{B_1}$ now is to be operated at the same Larmor frequency with the field direction pointing outward in the transverse plane, rotating around the $\boldsymbol{B_0}$ field direction to "excite" the spins from the longitudinal direction toward the transverse plane. At this stage, $\boldsymbol{M}$ begins to precess about the longitudinal axis at the Larmor frequency f depending on the field strength $\boldsymbol{B}$ (or the field strength that the local $\boldsymbol{M}$ is experiencing), following the equation $f = \gamma_0 \boldsymbol{B}$ (where γ_0 is the gyromagnetic ratio of ^{1}H: 42.58 MHz/T). This precession can be analogized to a gyroscope, with the spinning phenomenon driven by angular momentum under the influence of gravity.

The physical intuition behind this excitation is much easier to understand when analyzed in the rotating frame of reference. In the rotating frame of reference at the Larmor frequency, $\boldsymbol{B_1}$ is no longer rotating and the precession of $\boldsymbol{M}$ along the longitudinal axis virtually pauses. But following the same principle, $\boldsymbol{M}$ now precesses around $\boldsymbol{B_1}$ (in the rotating frame of reference) at a rate that follows the same principle as above depending on the strength of $\boldsymbol{B_1}$. Therefore, one could imagine $\boldsymbol{B_0}$, $\boldsymbol{B_1}$ and $\boldsymbol{M}$ are orthogonal to one another during the process of excitation. The purpose of having a non-zero transverse component of $\boldsymbol{M}$ is that, in the stationary frame of reference, a signal is generated by measuring voltage changes induced by changing the magnetic field (i.e., rotating $\boldsymbol{M}$), under Faraday's law, and being recorded, using a receiving coil. This signal can be

extremely useful in some other fields, such as nuclear magnetic resonance. However, for typical use in clinic settings, one should use an additional, directionally encoded gradient ($\boldsymbol{G}$) to impart spatial information, which can eventually be decoded into the day-to-day radiological presentation (as the frequency of the entire spin population is fundamentally identical and no spatial information has yet been encoded). The $\boldsymbol{G}$s are turned on, conventionally in the order of the encoding stages (with one following another): slice-encoding (during the excitation phase, i.e., at the same time when $\boldsymbol{B_1}$ is turned on), phase-encoding, and frequency-encoding (when signal is also being recorded using the receiving coil) during the time of excitation (along the longitudinal axis where the $\boldsymbol{B_0}$ lies on), phase and frequency spatial encoding (orthogonally) in the transverse plane, respectively. In this conventional method, the precise location of interest experiences a gradient field that induces linearly varying Larmor frequencies along the orthogonal spatial axes (i.e., $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}$). The recorded signal (when the receiving coil is turned on during the frequency encoding phase) is now stored in a line in the spatial frequency domain called k-space (2D in this case), with $\boldsymbol{k_x}$ and $\boldsymbol{k_y}$, representing the principal (and orthogonal) axes under the Cartesian convention. $\boldsymbol{k_x}$ and $\boldsymbol{k_y}$ are related to $\boldsymbol{x}$ and $\boldsymbol{y}$, and therefore are also closely related to $\boldsymbol{G_x}$ and $\boldsymbol{G_y}$, both with the Fourier transformation with the equation $\boldsymbol{k}(t) = \frac{\gamma}{2\pi}\int_0^t \boldsymbol{G}(\tau)d\tau$. The k-space needs to be filled in both $\boldsymbol{k_x}$ and $\boldsymbol{k_y}$ axes in order to form an image. Moreover, that explains the naming convention of the frequency and phase encoding, where $\boldsymbol{G_x}$ and $\boldsymbol{G_y}$ are played to 'help' signal traverse k-space by embedding different frequencies and different phases to the spin population in the selected slice in the spatial domain, so they can later be effectively decoded (i.e., distinguished) using the inverse Fourier transform to output a typical MR image. However, this only presents the conventional method, which is one out of many possible approaches to encode spatial information using $\boldsymbol{G}$, or

to traverse k-space. Readers should not be limited, and therefore are encouraged to seek external resources and literature to gain a better understanding of this topic (and MRI in general).

Nonetheless, the component of $\boldsymbol{M}$ in the transverse plane shrinks or decays continuously after being excited because of a combination of natural T1 and T2 relaxation effects. And because of the frequency offset in the spin population, potentially caused by, for example, chemical environment of H^1 (i.e., chemical shift) or field inhomogeneity, the phase coherence of the precession weakens, making the transverse component of $\boldsymbol{M}$ decay even faster. Especially, this is tricky because a linearly varying magnetic field, like $\boldsymbol{G}$, can inherently cause loss of phase coherence. Luckily, these induced frequency offsets are time-independent, hence they are reversible. The usages of $\boldsymbol{G}$ and $\boldsymbol{B_{1'}}$ present two distinct methods for rephasing the spin population, producing gradient echoes (GE) and spin echoes (SE), respectively. GE and SE are useful in different settings and for different goals. $\boldsymbol{B_{1'}}$ shares the same characteristic as $\boldsymbol{B_1}$, except they are orthogonal to each other. In SE sequences, $\boldsymbol{B_{1'}}$ is applied to flip $\boldsymbol{M}$ 180° at time τ. This step is essential because it virtually inverts all accumulated phase offset in the spin population (i.e., spins precess faster/slower now gain a negative/positive phase offset relative to 0). By the time reaches 2τ, spin regains phase coherence and this is why SE sequences are insensitive to static magnetic field inhomogeneities and are therefore considered to be T2-weighted. GE sequences, on the other hand, are designed to apply dephasing gradients (e.g., $\boldsymbol{G}$ with an opposite polarity, same magnitude and half of the duration of the frequency encoding $\boldsymbol{G}$) right before the frequency encoding phase to 'artificially' prepare for the anticipated dephasing. It is important to note that this dephasing gradient is embedded in SE sequences, and other instances when and where $\boldsymbol{G}$ is applied, but not for the purpose of accumulating phase. Regardless, the dephasing from other sources that cause frequency offsets is not rephased by GE. So, GE sequences carry information about field

inhomogeneities induced by, for example, chemical shift, field imperfection, air/bone-tissue interfaces, iron accumulation, and therefore are considered T2*-weighted (where $\frac{1}{T2^*} = \frac{1}{T2} + \frac{1}{T2\prime}$ and T2' is the relaxation time induced by the local field inhomogeneity as discussed).

A pulse sequence is a set of power instructions for MRI hardware that specify when the components of the system should be switched on and off (and that govern the signal traversing patterns and signal decoding methods). Furthermore, because the longitudinal relaxation time (T1, an exponent coefficient indicating how quickly the longitudinal component of $\boldsymbol{M}$ restores after excitation) and the transverse relaxation time (T2, an exponent coefficient indicating how quickly the transverse component of $\boldsymbol{M}$ decays after excitation) vary in tissues, the timing of each component in the pulse sequence matters. More importantly, they are the key pillars on which different pulse sequences were designed in the first place. Having this flexibility to produce different contrasts makes MRI a well-suited imaging modality for MS.

T1 and T2 can be affected by many factors at both microscopic and macroscopic scales. Cytoarchitecture, magnetic exchange rate, susceptibility, temperature and diffusion all play important roles. Since ^{1}H is the resonance agent of interest in typical MRI, the tissue environment around water is one of the most common first-level factors that determines the relaxation times.

1.6 MRI modalities in MS

Clinical uses of MRI in MS lie within the structural domain. T1-w and T2-w images are the most common ones. The contrast of these images is determined by the corresponding relaxation time, T1 or T2, and is achieved by adjusting the repetition time (TR) and echo time (TE) in the pulse sequence. TR (ranging around 400-1200 ms) differentiates $\boldsymbol{M}$ from tissue relaxation in the longitudinal axis as a result of T1 decay, whereas TE (ranging around 20-80 ms) differentiates $\boldsymbol{M}$

in the transverse plane during the T2 decay process. From the perspective of T1, using the three main tissue types from the CNS, WM, GM and cerebrospinal fluid (CSF), $T1_{CSF}>T1_{GM}>T1_{WM}$. Therefore, in T1-w images, WM is the brightest, followed by GM and CSF. In contrast, because $T2_{CSF}>T2_{GM}>T2_{WM}$, we can see an almost inverted picture using a T2w sequence where CSF is the brightest, followed by GM and WM. Therefore, because WML in MS loses its microstructural integrity (e.g., demyelination, axonal degeneration), the cellular environment is left with fewer structures, allowing water molecules to diffuse and move more freely. WMLs are usually hypointense (i.e., having a relatively lower intensity) on T1 images and hyperintense (i.e., having a relatively higher intensity) on T2 images relative to the neighbouring WM space.

These common MR sequences can be modified to provide better contrast and more specific features related to MS. Gd-enhanced T1 is the most commonly used contrast-enhanced MR sequence in MS. The use of Gd is because of its paramagnetic property. When it is injected, it enters the bloodstream and will infiltrate and accumulate in the active WMLs, where BBB is acutely damaged. The paramagnetic property of Gd shortens the T1 relaxation of the local lattice or tissue space. Therefore, the intensity of the rims of these WMLs with acute inflammation and active BBB disruption is elevated in the Gd-T1-w images, as opposed to the often hypointense core (Bou Fakhredin et al., 2016). On the other hand, because of the long T2 in CSF, some WMLs are not visually recognized and difficult to delineate, especially those that are close to the ventricles (i.e., periventricular WML). But T2-FLAIR, as mentioned, nullifies the signal from CSF by cleverly utilizing the longitudinal relaxation of completely free “water”. The mechanism is to first excite $\boldsymbol{M}$ beyond the transverse plane (i.e., > 90°), toward the opposite polarity of the longitudinal axis. In case of FLAIR, where $\boldsymbol{B_1}$ is designed to excite $\boldsymbol{M}$ to the opposite side of the longitudinal axis (i.e., 180°), this is called inversion recovery, one of the specialized magnetization preparation

techniques. During inversion recovery, $\boldsymbol{M}$ does not have any component in the transverse plane (orthogonal to the main magnetic field $\boldsymbol{B_0}$ direction); therefore, $\boldsymbol{M}$ effectively undergoes only T1 relaxation, beginning from the negative end of the longitudinal axis opposite to its equilibrium state. Following this concept, CSF's signal (or $\boldsymbol{M_{CSF}}$) can be virtually suppressed simply by applying the conventional excitation process (see Section 1.5), for this context of the T2-FLAIR sequence, a conventional T2 sequence at the inversion time (TI) for CSF or TI_{CSF} (where TI_{CSF} is when $\boldsymbol{M_{CSF}}$ reaches at theoretically '0' in the longitudinal axis during the T1 relaxation process after the magnetization preparation process). This method produces much better contrast for WML, allowing clinicians to identify the crucial WML locations much faster and thereby improving the diagnostic quality. Strategically, TI can be adjusted to suppress other tissue types of interest to meet specific clinical needs.

SWI also shows an interesting contrast that is becoming increasingly useful in the MS domain (Clarke et al., 2020). It is a GE-based sequence as one normally does to produce a T2*-weighted image. The key feature of SWI sequence acquisition/processing is that it uses both magnitude and phase images, whereas most other imaging sequences discard phase information. Phase is a natural parameter output by the Fourier transformation, and it has a meaningful value in extracting information about the distribution of the spin population in the transverse plane. Because the dephasing process in the transverse plane is sensitive to field inhomogeneities induced by particles with significant magnetic susceptibility, such as iron, deoxyhemoglobin, calcium, and air (Halefoglu & Yousem, 2018; Haller et al., 2021), local phase changes indicate a higher local concentration of these elements within the tissue. As an end product, the phase image is overlaid on the magnitude image using pointwise multiplication to highlight regions of high susceptibility. As briefly mentioned, SWI is now a modality recommended in practice for detecting central vein

sign and paramagnetic rim lesions, thereby further enhancing diagnostic accuracy and guiding clinical decisions (Montalban et al., 2025).

The flexibility in MRI allows us to view MS from multiple aspects. Beyond structural MRI (sMRI), these include, but are not limited to, perfusion MRI for blood flow analysis (Jahng et al., 2014), functional MRI for brain network analysis (van den Heuvel & Hulshoff Pol, 2010), and MR spectroscopy for analysis of the underlying chemical composition (Tognarelli et al., 2015). For the scope of the thesis, qMRI methods and metrics are briefly discussed, focusing only on those included in the current study. Over the past two decades, qMRI has received much more attention in the MS research landscape. This is because qMRI images are acquired and processed so that signal intensities are standardized and have physical units, meaning that voxels are not only spatially relevant but also individually interpretable within a biophysical framework. One related example to ensure a smooth transition in understanding from sMRI to qMRI is to quantitatively calculate the relaxation time, such as quantitative T1 (qT1). In contrast to T1-w images, qT1 in theory, requires at least two T1-w images with different TRs to be acquired. So, an exponential decay curve under the MR relaxation model (i.e., the Bloch equation) can be fitted to estimate T1 for voxels in time units.

With harmonized units and more delicate biophysical frameworks, qMRI is more sensitive and specific in measuring underlying microstructural integrity (Granziera et al., 2021; Tranfa et al., 2022). The ones that are included in this thesis work, also as one of the major parts of neuroimaging protocol of the master CCOMS dataset (see later section), includes five primary frameworks: diffusion tensor imaging (DTI), myelin water imaging (MWI), T1-weighted-by-T2-weighted (T1w/T2w) ratio, apparent transverse relaxation rate, ($R2^* = \frac{1}{T2^*}$), and quantitative susceptibility mapping (QSM).

DTI is a specific imaging technique within the broader field of diffusion imaging. Using diffusion imaging, one can infer the underlying microstructural environment and organization. Particles and molecules in our body are constantly undergoing diffusion, driven at the microscopic level by random Brownian motion. Several factors affect diffusion, including temperature, the cellular structural environment, and the size and spatial distribution of the elements of interest (Afzali et al., 2021; Winston, 2012). For example, given a set amount of time, water molecules can diffuse farther in a region with fewer obstacles. And that phenomenon can be described by a relatively higher diffusion coefficient (with a lower diffusion coefficient where diffusion is more constrained). In MRI, any use of $\boldsymbol{G}$, including spatial encoding steps, automatically implies dephasing of the spin population, therefore diminishing the measurable signal. However, it is easy to compensate for that by applying a gradient with the same effective area (e.g., identical duration and strength), and an opposite polarity to recover most of the initial signal (minus any T2 or T2* weighting due to the overall time of the forward and inverse gradients). This set of gradient pairs is essentially the feature diffusion gradient module added to the diffusion imaging sequence to sensitize the mobile water pool (i.e., decreasing the signal relative to the amount of diffusion along the direction of the encoding gradient) while leaving the signal in the static pool untouched. In other words, the further a water molecule travels between being dephased and then rephased (in the direction of the gradients), the less its spin population can gain back phase coherence, ultimately, leaving little to no signal to be detected. Since $\boldsymbol{G}$ can be applied in any direction using the vector combination, diffusion can be measured in multiple directions (at least six for DTI, although ideally more to yield accurate estimation (Zhan et al., 2010)) and with different amounts of diffusion-encoding (at least two values, including a reference measurement with no diffusion-encoding). Diffusion for an individual voxel can be estimated by fitting the probed signals (i.e.,

attenuated signals) to a rank-2, 3-by-3 tensor using linear algebra (i.e., eigendecomposition or diagonalization). The diffusion tensor of a voxel (after diagonalization) is often visualized as an ellipsoid, with the principal axes (i.e., primary, secondary and tertiary) representing the three orthogonal diffusivities or diffusion coefficients. Using these three diffusion coefficients, four classic DTI metrics are produced, axial diffusivity or AD (i.e., diffusion coefficients of the primary axes), radial diffusivity or RD (i.e., the mean of two diffusion coefficient of the secondary and tertiary axes), mean diffusivity or MD (i.e., the mean of all three diffusion coefficients from all axes) and fractional anisotropy or FA (i.e., an aggregated metric of all three diffusion coefficients from all axes that describes how isotropic the diffusion tensor is). In MS, DTI is useful for sensitizing areas with water molecules that diffuse abnormally due to the compromised microstructural integrity, such as neurodegeneration and edema. Some of these early changes and subtle heterogeneity in WM are not detectable in routinely collected MRI (Abdo et al., 2026; Gloor et al., 2023).

MWI is an imaging framework that discards the traditional assumption that each voxel of the spin population undergoes the same mono-exponential decay in the transverse plane, and instead adopts a more biologically plausible model with a multi-exponential decay resulting from a spectrum of spin populations. This is useful simply because of the scale discrepancy between MRI resolution and cellular anatomy and the fact that water molecules from different pools within each voxel have different T2 decays. Myelin is an extremely delicate structure that wraps around axons, facilitating the transmission of neural signals with minimal energy expenditure (Stadelmann et al., 2019). At the same time, some water molecules are trapped in the myelin sheath, and studies have shown that myelin water generally has a T2 relaxation time between 10-50 ms (A. MacKay et al., 2006). This concept extends to other components in the T2 spectrum, including intra- and

extra-cellular water, free water and CSF. Therefore, fractions of different components, such as the myelin water fraction (MWF) and the intra- and extra-cellular water fraction (IEWF), can be estimated by dividing the amount of signal originating from a certain T2 range by the amount of signal from the entire T2 spectrum. In the framework used in this study, the T2 scale is chosen to be logarithmic for more stable spectrum fitting with limited data. Therefore, geometric mean T2 (gT2) is a more suitable descriptive mean. The MWI sequence can be time-consuming. That is because of two interplaying factors. One is that the T2 decay needs to be sampled sufficiently so that a spectrum can be properly estimated with minimal numerical error, and this is difficult because of the limitation of how many spin echoes and gradient echoes can be acquired in a whole brain sequence. Second, due to the potential magnetization effect at the adjacent slices introduced by imperfect slice profile during excitation, which can severely impact myelin water pool (Vavasour et al., 2000), and lead to inaccuracies in T2 spectrum estimation, a 3D sequence is needed (Prasloski et al., 2012a). That is, instead of exciting slice by slice, 3D slabs with greater depth can now be used to enhance the signal. However, the trade-off of a 3D sequence is that those kinds of sequence have longer idle time because TR is a hard limit before the sequences run again, and there are not enough waitlisted volumes left to be excited during the downtime. Nevertheless, substantial efforts have been made in the field to accelerate 3D myelin water imaging (Alonso-Ortiz et al., 2015a; Prasloski et al., 2012a).

T1w/T2w ratio, originally proposed by Glasser et al. (Glasser & van Essen, 2011), is one of the most cost-effective quantitative measures, simply because both T1-w and T2-w images are routinely collected in the clinic, thereby enhancing retrospective analyses. Although T1w/T2w ratio comes with inherent challenges due to the qualitative nature of the two images (Filippi et al., 1997; Han et al., 2006; Kruggel et al., 2010), standardization and calibration can help alleviate the

concerns (Ganzetti et al., 2014; Misaki et al., 2015) and enhance results in studies with disease like MS (Cooper et al., 2019). Myelin is a bipolar lipid layer. Because the water molecules are in a bound water pool and the lipid's susceptibility impact on the local field, T1w/T2w ratio can theoretically enhance myelin contrast by using their relative T1 and (especially short) T2 relaxation characteristics (Glasser & van Essen, 2011). Nevertheless, studies have found that T1w/T2w ratio is not specific to myelin but also sensitive to other microstructural characteristics (Baum et al., 2022; Sandrone et al., 2023a; Uddin et al., 2019; Uddin, Figley, Marrie, et al., 2018a).

R2* is an enhanced measure inversely related to T2*. Similar to how qT1 is estimated, R2* is derived by fitting a mono-exponential equation using multiple sampled points of the signal decay curve. Because it measures apparent transverse relaxation using gradient echoes, R2* is sensitive to susceptibility-induced field inhomogeneity, especially those that induce a short T2* decay, such as iron (paramagnetic), and calcium and myelin (diamagnetic). In MS, iron accumulation can occur in both WM and GM regions, and it plays roles in both repair and degeneration mechanisms (Williams, Buchheit, Berman, & LeVine, 2012). However, because myelin is diamagnetic, studies have also shown that R2* is sensitive to myelin content (Paling et al., 2012; Zhang et al., 2016a).

QSM has been studied extensively in the past decade. Because susceptibility encompasses both paramagnetic (positive) and diamagnetic (negative) components, it is a unique metric with strong potential for characterizing the susceptibility distribution in the brain. To compute the QSM, two stages are normally considered: 1) background field removal and 2) local field inversion. First, the background field induced by the air-tissue interface is removed from the total field estimated by the phase image. Then the resultant local field is inverted by treating voxels as individual dipoles with the electromagnetic equation. However, this is problematic if the field perturbation is acquired only at one orientation, because the "magic angle" derived by the angular part of the

equation, $3\cos^2(\theta) - 1$, nullifies the numerical outcome. In other words, when the angle of the perturbed voxel relative to the voxel that causes local field inhomogeneity due to its magnetic property is approximately 54.7°, the dipole effect is virtually none. This then becomes an ill-posed problem with uncertainties introduced into the estimation. Although a solution is available through repeated imaging with two additional orthogonal orientations using the spatial encoding method, this is not feasible in clinical settings (Gkotsoulias et al., 2024). Currently, many different methods are being developed to tackle the problem. The studies in this thesis use the Quantitative Susceptibility Mapping Artifact Reduction Technique featuring a composite map that combines two parallel stages of local field inversion (with and without vasculature estimation) to suppress spurious artifacts, including streak patterns near vascular structures (Yaghmaie et al., 2021).

1.7 Current gap of MRI in MS

The current MRI findings in WML have been an invaluable feature in MS, indicating disease activity for diagnosis and monitoring. Even though MRI can provide multiple types of image contrast with a single machine, its use in clinical settings remains largely qualitative. Various studies have tried to use lesion count and total lesion size for characterizing disease severity and progression; however, these assessments have shown relatively limited correlations with specific clinical deficits, and poor predictive value in terms of progression, leading to the so-called 'clinico-radiological paradox' (Barkhof, 2002), and suggesting that more advantageous imaging-based methods should be sought to characterize MS better.

In addition, although studies have shown that earlier treatment is generally more effective at halting progression and preventing relapses in MS, a one-size-fits-all approach is not ideal, as high efficacy DMTs are sensitive to disease phenotypes and stages. Therefore, in parallel with the

process of identifying important therapeutic targets to develop new DMTs, better characterization of the disease should be accompanied to create a greater impact.

As outlined in the previous section, advances in qMRIs improve disease characterization because they are more specific and sensitive to different aspects of MS pathophysiology. To date, none of these metrics has yet been individually validated as an MS biomarker. It is also recognized that each metric faces its unique challenges for large-scale clinical implementation (e.g., harmonization of imaging sequences and processing pipelines). Nevertheless, this poses many exciting opportunities for research spanning from basic science to clinical research. WM is a large space with an intertwined network of axons that relay neural signals into interpretable actions, and WM pathophysiology in MS is both spatially heterogeneous and longitudinally dynamic. This thesis, therefore, aims at gaining a better understanding of a set of qMRIs in their ability to characterize the WM space in MS, with a series of analyses focusing on their 1) compatibility, sensitivity in both 2) coarse and 3) fine region-of-interest (ROI) definition oriented around WML and 4) application in predicting WML spatial evolution. This is a translational work that involves both cross-sectional and longitudinal data from the MS cohort of the CCOMS study (briefly described in the next section).

1.8 The features of the MS cohort of CCOMS

CCOMS is a dataset collected for the intended use of studying comorbidity and cognition in MS (Uddin et al., 2022). It is an information-rich, multimodal and multidimensional dataset that captures comprehensive demographic information, batteries of neurological tests, and a comprehensive list of MRI modalities (Uddin et al., 2022). The inclusion criteria for the MS cohort focus on MS in general rather than on collecting data equally from the different MS phenotypes

(i.e., relapsing-remitting, primary progressive, and secondary progressive). Instead, the goal was to obtain a dataset that represents the general MS population in the public. Within the scope of the quantitative MRI in the list of MRI scans of the MS cohort, a total of 10 metrics (as discussed above in Section 1.6) are selected to conduct the studies in the thesis. The intention in using the multimodal approach is to leverage the strengths of each metric to collectively capture richer information about white matter microstructure in MS. This MS cohort has a longitudinal component with two time points (i.e., baseline and 2-year follow-up) to facilitate both cross-sectional and longitudinal analyses. The follow-up dataset is used for the cross-sectional analyses (Chapters 2 – 4) with a maximized sample size (N=79), and the longitudinal dataset (both baseline and 2-year follow-up) is used for the longitudinal analyses (Chapters 5 and 6). Due to the reported long scanning time for the participants during the initial phase of data acquisition, the multi-echo gradient-echo sequence (i.e., the sequence that produces R2* and QSM) was temporarily placed offline for further optimization before being put back online. Therefore, a significant portion of participants with MS at baseline did not have the multi-echo gradient-echo sequence included in their image acquisition, therefore reducing the sample size (N=35) for the longitudinal analyses, when all possible quantitative MRI metrics are included.

To streamline the analyses in later chapters, the data-preprocessing steps were optimized to minimize artifacts, based on the research group's best judgment and the literature and research advances at the time of data collection. They were also standardized and fully scripted for each participant. Automation and standardization are instrumental for two main reasons. First, both approaches minimize human interaction and intervention, therefore suppressing human-induced bias, errors and noise introduced into the data analysis, which can affect the interpretation of results. Second, they streamline the implementation process and decrease both cost and time compared to

recruiting human raters. The entire pre-processing pipelines for each sequence, including the modality-specific pre-processing pipelines to convert raw images into subject-space qMRI maps, and the standardized co-registration and normalization steps to bring qMRI maps from the subject space to the template space, are documented in our previously published protocol manuscript (Uddin et al., 2022). However, since WMLs are a central focus of the dissertation and the WML masks will be used throughout the subsequent analyses (Chapters 2-6), it is perhaps worth reiterating how lesion masks for each MS participant were produced. Briefly, T2-hyperintense white matter lesion masks were segmented using two independent algorithms: 1) the Lesion Segmentation Tool (LST; http://www.applied-statistics.de/lst.html) in SPM12, which employs a lesion growth algorithm (Schmidt et al., 2012); and 2) FSL's automated Brain Intensity AbNormality Classification Algorithm (BIANCA; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BIANCA), which employs a k-nearest neighbor algorithm (Griffanti et al., 2016). Then, the final lesion mask for each participant in the analyses was taken as the overlap between the LST and BIANCA lesion maps. This two-step process is more conservative than either method alone (especially for edges/boarders and small/isolated lesions), and increases the probability that voxels included in the final WML mask were correctly classified as T2 hyperintensities. Regardless, it is also important to note that the primary input for the lesion segmentation algorithms in this study were participants' T2-FLAIR images, so the results of subsequent analyses should be interpreted with this in mind, since T2-hyperintense lesions are not the only type of WMLs found in MS, and may only have partial overlap with other lesion types, such as T1-hypointense lesions and paramagnetic rim lesions (PRLs).

1.9 References

Abdo, A. I., Ramadan, M. R., elfeshawy, mohamed salah, & Ahmed, M. G. (2026). Role of Magnetic Resonance Diffusion Tensor Imaging in Assessment of Normal Appearing White Matter in Patients with Multiple Sclerosis. *Al-Azhar International Medical Journal*, *0*(0), 184–189. https://doi.org/10.21608/aimj.2026.451416.3097

Afzali, M., Pieciak, T., Newman, S., Garyfallidis, E., Özarslan, E., Cheng, H., & Jones, D. K. (2021). The sensitivity of diffusion MRI to microstructural properties and experimental factors. *Journal of Neuroscience Methods*, *347*, 108951. https://doi.org/10.1016/j.jneumeth.2020.108951

Alonso-Ortiz, E., Levesque, I. R., & Pike, G. B. (2015). MRI-based myelin water imaging: A technical review. *Magnetic Resonance in Medicine*, *73*(1), 70–81. https://doi.org/10.1002/mrm.25198

Barkhof, F. (2002). The clinico-radiological paradox in multiple sclerosis revisited. *Current Opinion in Neurology*, *15*(3), 239–245. https://doi.org/10.1097/00019052-200206000-00003

Barkhof, F., & Walderveen, M. van. (1999). Characterization of tissue damage in multiple sclerosis by nuclear magnetic resonance. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *354*(1390), 1675–1686. https://doi.org/10.1098/rstb.1999.0511

Baum, G. L., Flournoy, J. C., Glasser, M. F., Harms, M. P., Mair, P., Sanders, A. F. P., Barch, D. M., Buckner, R. L., Bookheimer, S., Dapretto, M., Smith, S., Thomas, K. M., Yacoub, E., Van Essen, D. C., & Somerville, L. H. (2022). Graded Variation in T1w/T2w Ratio during Adolescence: Measurement, Caveats, and Implications for Development of Cortical Myelin. *The Journal of Neuroscience*, *42*(29), 5681–5694. https://doi.org/10.1523/JNEUROSCI.2380-21.2022

Bou Fakhredin, R., Saade, C., Kerek, R., El-Jamal, L., Khoury, S. J., & El-Merhi, F. (2016). Imaging in multiple sclerosis: A new spin on lesions. *Journal of Medical Imaging and Radiation Oncology*, *60*(5), 577–586. https://doi.org/10.1111/1754-9485.12498

Briggs, F. B. S., Thompson, N. R., & Conway, D. S. (2019). Prognostic factors of disability in relapsing remitting multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *30*, 9–16. https://doi.org/10.1016/j.msard.2019.01.045

Cagol, A., Cortese, R., Barakovic, M., Schaedelin, S., Ruberte, E., Absinta, M., Barkhof, F., Calabrese, M., Castellaro, M., Ciccarelli, O., Cocozza, S., De Stefano, N., Enzinger, C., Filippi, M., Jurynczyk, M., Maggi, P., Mahmoudi, N., Messina, S., Montalban, X., … Yousry, T. (2024). Diagnostic Performance of Cortical Lesions and the Central Vein Sign in Multiple Sclerosis. *JAMA Neurology*, *81*(2), 143. https://doi.org/10.1001/jamaneurol.2023.4737

Clarke, M. A., Pareto, D., Pessini-Ferreira, L., Arrambide, G., Alberich, M., Crescenzo, F., Cappelle, S., Tintoré, M., Sastre-Garriga, J., Auger, C., Montalban, X., Evangelou, N., & Rovira, À. (2020). Value of 3T Susceptibility-Weighted Imaging in the Diagnosis of Multiple Sclerosis. *American Journal of Neuroradiology*, *41*(6), 1001–1008. https://doi.org/10.3174/ajnr.A6547

Compston, A., & Coles, A. (2002). Multiple sclerosis. *The Lancet*, *359*(9313), 1221–1231. https://doi.org/10.1016/S0140-6736(02)08220-X

Cooper, G., Finke, C., Chien, C., Brandt, A. U., Asseyer, S., Ruprecht, K., Bellmann-Strobl, J., Paul, F., & Scheel, M. (2019). Standardization of T1w/T2w Ratio Improves Detection of Tissue Damage in Multiple Sclerosis. *Frontiers in Neurology*, *10*, 334. https://doi.org/10.3389/fneur.2019.00334

Cosh, A., & Carslaw, H. (2014). Multiple sclerosis: symptoms and diagnosis. *InnovAiT: Education and Inspiration for General Practice*, *7*(11), 651–657. https://doi.org/10.1177/1755738014551618

Dal-Bianco, A., Oh, J., Sati, P., & Absinta, M. (2024). Chronic active lesions in multiple sclerosis: classification, terminology, and clinical significance. *Therapeutic Advances in Neurological Disorders*, *17*. https://doi.org/10.1177/17562864241306684

Degenhardt, A., Ramagopalan, S. V., Scalfari, A., & Ebers, G. C. (2009). Clinical prognostic factors in multiple sclerosis: a natural history review. *Nature Reviews Neurology*, *5*(12), 672–682. https://doi.org/10.1038/nrneurol.2009.178

Disanto, G., Berlanga, A. J., Handel, A. E., Para, A. E., Burrell, A. M., Fries, A., Handunnetthi, L., De Luca, G. C., & Morahan, J. M. (2011). Heterogeneity in Multiple Sclerosis: Scratching the Surface of a Complex Disease. *Autoimmune Diseases*, *2011*, 1–12. https://doi.org/10.4061/2011/932351

Elliott, C., Wolinsky, J. S., Hauser, S. L., Kappos, L., Barkhof, F., Bernasconi, C., Wei, W., Belachew, S., & Arnold, D. L. (2019). Slowly expanding/evolving lesions as a magnetic resonance imaging marker of chronic active multiple sclerosis lesions. *Multiple Sclerosis Journal*, *25*(14), 1915–1925. https://doi.org/10.1177/1352458518814117

Filippi, M., Amato, M. P., Centonze, D., Gallo, P., Gasperini, C., Inglese, M., Patti, F., Pozzilli, C., Preziosa, P., & Trojano, M. (2022). Correction to: Early use of high-efficacy disease-modifying therapies makes the difference in people with multiple sclerosis: an expert opinion. *Journal of Neurology*, *269*(12), 6690–6691. https://doi.org/10.1007/s00415-022-11385-4

Filippi, M., Preziosa, P., Banwell, B. L., Barkhof, F., Ciccarelli, O., De Stefano, N., Geurts, J. J. G., Paul, F., Reich, D. S., Toosy, A. T., Traboulsee, A., Wattjes, M. P., Yousry, T. A., Gass, A.,

Lubetzki, C., Weinshenker, B. G., & Rocca, M. A. (2019). Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. In *Brain* (Vol. 142, Number 7, pp. 1858–1875). Oxford University Press. https://doi.org/10.1093/brain/awz144

Filippi, M., van Waesberghe, J. H., Horsfield, M. A., Bressi, S., Gasperini, C., Yousry, T. A., Gawne-Cain, M. L., Morrissey, S. P., Rocca, M. A., Barkhof, F., Lycklama à Nijeholt, G. J., Bastianello, S., & Miller, D. H. (1997). Interscanner variation in brain MRI lesion load measurements in MS: Implications for clinical trials. *Neurology*, *49*(2), 371–377. https://doi.org/10.1212/WNL.49.2.371

Freeman, L., Longbrake, E. E., Coyle, P. K., Hendin, B., & Vollmer, T. (2022). High-Efficacy Therapies for Treatment-Naïve Individuals with Relapsing–Remitting Multiple Sclerosis. *CNS Drugs*, *36*(12), 1285–1299. https://doi.org/10.1007/s40263-022-00965-7

Frohman, E. M., Racke, M. K., & Raine, C. S. (2006). Multiple Sclerosis — The Plaque and Its Pathogenesis. *New England Journal of Medicine*, *354*(9), 942–955. https://doi.org/10.1056/NEJMra052130

Ganzetti, M., Wenderoth, N., & Mantini, D. (2014). Whole brain myelin mapping using T1- and T2-weighted MR imaging data. *Frontiers in Human Neuroscience*, *8*(SEP). https://doi.org/10.3389/fnhum.2014.00671

Giovannoni, G. (2018). Disease-modifying treatments for early and advanced multiple sclerosis: a new treatment paradigm. *Current Opinion in Neurology*, *31*(3), 233–243. https://doi.org/10.1097/WCO.0000000000000561

Giovannoni, G., Butzkueven, H., Dhib-Jalbut, S., Hobart, J., Kobelt, G., Pepper, G., Sormani, M. P., Thalheim, C., Traboulsee, A., & Vollmer, T. (2016). Brain health: time matters in

multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *9*, S5–S48. https://doi.org/10.1016/j.msard.2016.07.003

Gkotsoulias, D. G., Jäger, C., Müller, R., Gräßle, T., Olofsson, K. M., Møller, T., Unwin, S., Crockford, C., Wittig, R. M., Bilgic, B., & Möller, H. E. (2024). Chaos and COSMOS—Considerations on QSM methods with multiple and single orientations and effects from local anisotropy. *Magnetic Resonance Imaging*, *110*, 104–111. https://doi.org/10.1016/j.mri.2024.04.020

Glasser, M. F., & van Essen, D. C. (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *Journal of Neuroscience*, *31*(32), 11597–11616. https://doi.org/10.1523/JNEUROSCI.2180-11.2011

Gloor, M., Andelova, M., Gaetano, L., Papadopoulou, A., Burguet Villena, F., Sprenger, T., Radue, E.-W., Kappos, L., Bieri, O., & Garcia, M. (2023). Longitudinal analysis of new multiple sclerosis lesions with magnetization transfer and diffusion tensor imaging. *European Radiology*, *34*(3), 1680–1691. https://doi.org/10.1007/s00330-023-10173-6

Granziera, C., Wuerfel, J., Barkhof, F., Calabrese, M., De Stefano, N., Enzinger, C., Evangelou, N., Filippi, M., Geurts, J. J. G., Reich, D. S., Rocca, M. A., Ropele, S., Rovira, À., Sati, P., Toosy, A. T., Vrenken, H., Gandini Wheeler-Kingshott, C. A. M., Kappos, L., Barkhof, F., … Yousry, T. (2021). Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. *Brain*, *144*(5), 1296–1311. https://doi.org/10.1093/brain/awab029

Griffanti, L., Zamboni, G., Khan, A., Li, L., Bonifacio, G., Sundaresan, V., Schulz, U. G., Kuker, W., Battaglini, M., Rothwell, P. M., & Jenkinson, M. (2016). BIANCA (Brain Intensity AbNormality Classification Algorithm): A new tool for automated segmentation of white matter hyperintensities. *NeuroImage*, *141*, 191–205. https://doi.org/10.1016/j.neuroimage.2016.07.018

Grossman, R. I., Gonzalez-Scarano, F., Atlas, S. W., Galetta, S., & Silberberg, D. H. (1986). Multiple sclerosis: gadolinium enhancement in MR imaging. *Radiology*, *161*(3), 721–725. https://doi.org/10.1148/radiology.161.3.3786722

Haki, M., AL-Biati, H. A., Al-Tameemi, Z. S., Ali, I. S., & Al-hussaniy, H. A. (2024). Review of multiple sclerosis: Epidemiology, etiology, pathophysiology, and treatment. *Medicine*, *103*(8), e37297. https://doi.org/10.1097/MD.0000000000037297

Halefoglu, A. M., & Yousem, D. M. (2018). Susceptibility weighted imaging: Clinical applications and future directions. *World Journal of Radiology*, *10*(4), 30–45. https://doi.org/10.4329/wjr.v10.i4.30

Haller, S., Haacke, E. M., Thurnher, M. M., & Barkhof, F. (2021). Susceptibility-weighted imaging: Technical essentials and clinical neurologic applications. In *Radiology* (Vol. 299, Number 1, pp. 3–26). Radiological Society of North America Inc. https://doi.org/10.1148/RADIOL.2021203071

Han, X., Jovicich, J., Salat, D., van der Kouwe, A., Quinn, B., Czanner, S., Busa, E., Pacheco, J., Albert, M., Killiany, R., Maguire, P., Rosas, D., Makris, N., Dale, A., Dickerson, B., & Fischl, B. (2006). Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, *32*(1), 180–194. https://doi.org/10.1016/j.neuroimage.2006.02.051

Huang, Y. (2025). Predictive factors for the progression of multiple sclerosis: a meta-analysis. *Neurological Sciences*, *46*(8), 3765–3773. https://doi.org/10.1007/s10072-025-08142-6

Jahng, G.-H., Li, K.-L., Ostergaard, L., & Calamante, F. (2014). Perfusion Magnetic Resonance Imaging: A Comprehensive Update on Principles and Techniques. *Korean Journal of Radiology*, *15*(5), 554. https://doi.org/10.3348/kjr.2014.15.5.554

Jiang, F., Zhang, F., Su, Y., Zhang, C., & Chang, T. (2024). Knowledge mapping of disease-modifying therapy (DMT) in multiple sclerosis (MS): A bibliometrics analysis. *Heliyon*, *10*(11), e31744. https://doi.org/10.1016/j.heliyon.2024.e31744

Kruggel, F., Turner, J., & Muftuler, L. T. (2010). Impact of scanner hardware and imaging protocol on image quality and compartment volume precision in the ADNI cohort. *NeuroImage*, *49*(3), 2123–2133. https://doi.org/10.1016/j.neuroimage.2009.11.006

Kuhlmann, T., Moccia, M., Coetzee, T., Cohen, J. A., Correale, J., Graves, J., Marrie, R. A., Montalban, X., Yong, V. W., Thompson, A. J., Reich, D. S., Amato, M. P., Banwell, B., Barkhof, F., Chataway, J., Chitnis, T., Comi, G., Derfuss, T., Finlayson, M., … Waubant, E. (2023). Multiple sclerosis progression: time for a new mechanism-driven framework. *The Lancet Neurology*, *22*(1), 78–88. https://doi.org/10.1016/S1474-4422(22)00289-7

Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sørensen, P. S., Thompson, A. J., Wolinsky, J. S., Balcer, L. J., Banwell, B., Barkhof, F., Bebo, B., Calabresi, P. A., Clanet, M., Comi, G., Fox, R. J., Freedman, M. S., Goodman, A. D., Inglese, M., Kappos, L., … Polman, C. H. (2014). Defining the clinical course of multiple sclerosis. *Neurology*, *83*(3), 278–286. https://doi.org/10.1212/WNL.0000000000000560

MacKay, A., Laule, C., Vavasour, I., Bjarnason, T., Kolind, S., & Mädler, B. (2006). Insights into brain microstructure from the T2 distribution. *Magnetic Resonance Imaging*, *24*(4), 515–525. https://doi.org/10.1016/j.mri.2005.12.037

Marrie, R. A. (2017). Comorbidity in multiple sclerosis: implications for patient care. *Nature Reviews Neurology*, *13*(6), 375–382. https://doi.org/10.1038/nrneurol.2017.33

Meaton, I., Altokhis, A., Allen, C. M., Clarke, M. A., Sinnecker, T., Meier, D., Enzinger, C., Calabrese, M., De Stefano, N., Pitiot, A., Giorgio, A., Schoonheim, M. M., Paul, F., Pawlak,

M. A., Schmidt, R., Granziera, C., Kappos, L., Montalban, X., Rovira, À., … Evangelou, N. (2022). Paramagnetic rims are a promising diagnostic imaging biomarker in multiple sclerosis. *Multiple Sclerosis Journal*, *28*(14), 2212–2220. https://doi.org/10.1177/13524585221118677

Misaki, M., Savitz, J., Zotev, V., Phillips, R., Yuan, H., Young, K. D., Drevets, W. C., & Bodurka, J. (2015). Contrast enhancement by combining T1- and T2-weighted structural brain MR Images. *Magnetic Resonance in Medicine*, *74*(6), 1609–1620. https://doi.org/10.1002/mrm.25560

Montalban, X., Lebrun-Frénay, C., Oh, J., Arrambide, G., Moccia, M., Pia Amato, M., Amezcua, L., Banwell, B., Bar-Or, A., Barkhof, F., Butzkueven, H., Ciccarelli, O., Chataway, J., Cohen, J. A., Comi, G., Correale, J., Deisenhammer, F., Filippi, M., Fiol, J., … Thompson, A. J. (2025). Diagnosis of multiple sclerosis: 2024 revisions of the McDonald criteria. *The Lancet Neurology*, *24*(10), 850–865. https://doi.org/10.1016/S1474-4422(25)00270-4

Moore, G. R. W., Leung, E., MacKay, A. L., Vavasour, I. M., Whittall, K. P., Cover, K. S., Li, D. K. B., Hashimoto, S. A., Oger, J., Sprinkle, T. J., & Paty, D. W. (2000). A pathology-MRI study of the short-T2 component in formalin-fixed multiple sclerosis brain. *Neurology*, *55*(10), 1506–1510. https://doi.org/10.1212/WNL.55.10.1506

Murray, T. J. (2005). Multiple sclerosis: the history of a disease. In *Journal of the Royal Society of Medicine* (Vol. 98, Number 6).

Noseworthy, J. H., Lucchinetti, C., Rodriguez, M., & Weinshenker, B. G. (2000). Multiple Sclerosis. *New England Journal of Medicine*, *343*(13), 938–952. https://doi.org/10.1056/NEJM200009283431307

Paling, D., Tozer, D., Wheeler-Kingshott, C., Kapoor, R., Miller, D. H., & Golay, X. (2012). Reduced R_2' in multiple sclerosis normal appearing white matter and lesions may reflect

decreased myelin and iron content. *Journal of Neurology, Neurosurgery & Psychiatry*, *83*(8), 785–792. https://doi.org/10.1136/jnnp-2012-302541

Pearce, J. M. S. (2005). Historical Descriptions of Multiple Sclerosis. *European Neurology*, *54*(1), 49–53. https://doi.org/10.1159/000087387

Prasloski, T., Rauscher, A., MacKay, A. L., Hodgson, M., Vavasour, I. M., Laule, C., & Mädler, B. (2012). Rapid whole cerebrum myelin water imaging using a 3D GRASE sequence. *Neuroimage*, *63*(1), 533–539. https://doi.org/10.1016/j.neuroimage.2012.06.064

Rejdak, K., Jackson, S., & Giovannoni, G. (2010). Multiple sclerosis: a practical overview for clinicians. *British Medical Bulletin*, *95*(1), 79–104. https://doi.org/10.1093/bmb/ldq017

Rocca, M. A., Preziosa, P., Barkhof, F., Brownlee, W., Calabrese, M., De Stefano, N., Granziera, C., Ropele, S., Toosy, A. T., Vidal-Jordana, À., Di Filippo, M., & Filippi, M. (2024). Current and future role of MRI in the diagnosis and prognosis of multiple sclerosis. *The Lancet Regional Health - Europe*, *44*, 100978. https://doi.org/10.1016/j.lanepe.2024.100978

Rotstein, D., & Montalban, X. (2019). Reaching an evidence-based prognosis for personalized treatment of multiple sclerosis. *Nature Reviews Neurology*, *15*(5), 287–300. https://doi.org/10.1038/s41582-019-0170-8

Rovira, À., Doniselli, F. M., Auger, C., Haider, L., Hodel, J., Severino, M., Wattjes, M. P., van der Molen, A. J., Jasperse, B., Mallio, C. A., Yousry, T., & Quattrocchi, C. C. (2023). Use of gadolinium-based contrast agents in multiple sclerosis: a review by the ESMRMB-GREC and ESNR Multiple Sclerosis Working Group. *European Radiology*, *34*(3), 1726–1735. https://doi.org/10.1007/s00330-023-10151-y

Saade, C., Bou-Fakhredin, R., Yousem, D. M., Asmar, K., Naffaa, L., & El-Merhi, F. (2018). Gadolinium and Multiple Sclerosis: Vessels, Barriers of the Brain, and Glymphatics. *American Journal of Neuroradiology*, *39*(12), 2168–2176. https://doi.org/10.3174/ajnr.A5773

Sahraian, M. A., Radue, E.-W., Haller, S., & Kappos, L. (2009). Black holes in multiple sclerosis: definition, evolution, and clinical correlations. *Acta Neurologica Scandinavica*, *122*(1), 1–8. https://doi.org/10.1111/j.1600-0404.2009.01221.x

Sandrone, S., Aiello, M., Cavaliere, C., Thiebaut de Schotten, M., Reimann, K., Troakes, C., Bodi, I., Lacerda, L., Monti, S., Murphy, D., Geyer, S., Catani, M., & Dell'Acqua, F. (2023). Mapping myelin in white matter with T1-weighted/T2-weighted maps: discrepancy with histology and other myelin MRI measures. *Brain Structure and Function*, *228*(2), 525–535. https://doi.org/10.1007/s00429-022-02600-z

Sati, P., Oh, J., Constable, R. T., Evangelou, N., Guttmann, C. R. G., Henry, R. G., Klawiter, E. C., Mainero, C., Massacesi, L., McFarland, H., Nelson, F., Ontaneda, D., Rauscher, A., Rooney, W. D., Samaraweera, A. P. R., Shinohara, R. T., Sobel, R. A., Solomon, A. J., Treaba, C. A., … Reich, D. S. (2016). The central vein sign and its clinical evaluation for the diagnosis of multiple sclerosis: a consensus statement from the North American Imaging in Multiple Sclerosis Cooperative. *Nature Reviews Neurology*, *12*(12), 714–722. https://doi.org/10.1038/nrneurol.2016.166

Schmidt, P., Gaser, C., Arsic, M., Buck, D., Förschler, A., Berthele, A., Hoshi, M., Ilg, R., Schmid, V. J., Zimmer, C., Hemmer, B., & Mühlau, M. (2012). An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. *NeuroImage*, *59*(4), 3774–3783. https://doi.org/10.1016/j.neuroimage.2011.11.032

Singer, B. A., Feng, J., & Chiong-Rivero, H. (2024). Early use of high-efficacy therapies in multiple sclerosis in the United States: benefits, barriers, and strategies for encouraging adoption. *Journal of Neurology*, *271*(6), 3116–3130. https://doi.org/10.1007/s00415-024-12305-4

Stadelmann, C., Timmler, S., Barrantes-Freer, A., & Simons, M. (2019). Myelin in the Central Nervous System: Structure, Function, and Pathology. *Physiol Rev*, *99*, 1381–1431. https://doi.org/10.1152/physrev.00031.2018.-Oligodendro

Tognarelli, J. M., Dawood, M., Shariff, M. I. F., Grover, V. P. B., Crossey, M. M. E., Cox, I. J., Taylor-Robinson, S. D., & McPhail, M. J. W. (2015). Magnetic Resonance Spectroscopy: Principles and Techniques: Lessons for Clinicians. *Journal of Clinical and Experimental Hepatology*, *5*(4), 320–328. https://doi.org/10.1016/j.jceh.2015.10.006

Tranfa, M., Pontillo, G., Petracca, M., Brunetti, A., Tedeschi, E., Palma, G., & Cocozza, S. (2022). Quantitative MRI in Multiple Sclerosis: From Theory to Application. *American Journal of Neuroradiology*, *43*(12), 1688–1695. https://doi.org/10.3174/ajnr.A7536

Uddin, M. N., Figley, T. D., Kornelsen, J., Mazerolle, E. L., Helmick, C. A., O'Grady, C. B., Pirzada, S., Patel, R., Carter, S., Wong, K., Essig, M. R., Graff, L. A., Bolton, J. M., Marriott, J. J., Bernstein, C. N., Fisk, J. D., Marrie, R. A., & Figley, C. R. (2022). The comorbidity and cognition in multiple sclerosis (CCOMS) neuroimaging protocol: Study rationale, MRI acquisition, and minimal image processing pipelines. *Frontiers in Neuroimaging*, *1*. https://doi.org/10.3389/fnimg.2022.970385

Uddin, M. N., Figley, T. D., Marrie, R. A., & Figley, C. R. (2018). Can T_1 w/ T_2 w ratio be used as a myelin-specific measure in subcortical structures? Comparisons between FSE-based T_1 w/ T_2 w ratios, GRASE-based T_1 w/ T_2 w ratios and multi-echo GRASE-based myelin water fractions. *NMR in Biomedicine*, *31*(3). https://doi.org/10.1002/nbm.3868

Uddin, M. N., Figley, T. D., Solar, K. G., Shatil, A. S., & Figley, C. R. (2019). Comparisons between multi-component myelin water fraction, T1w/T2w ratio, and diffusion tensor imaging measures in healthy human brain structures. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-39199-x

van den Heuvel, M. P., & Hulshoff Pol, H. E. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacology*, *20*(8), 519–534. https://doi.org/10.1016/j.euroneuro.2010.03.008

Vavasour, I. M., Whittall, K. P., Li, D. K. B., & MacKay, A. L. (2000). Different magnetization transfer effects exhibited by the short and longT2 components in human brain. *Magnetic Resonance in Medicine*, *44*(6), 860–866. https://doi.org/10.1002/1522-2594(200012)44:6<860::AID-MRM6>3.0.CO;2-C

Vollmer, T. L., Nair, K. V., Williams, I. M., & Alvarez, E. (2021). Multiple Sclerosis Phenotypes as a Continuum. *Neurology Clinical Practice*, *11*(4), 342–351. https://doi.org/10.1212/CPJ.0000000000001045

Walton, C., King, R., Rechtman, L., Kaye, W., Leray, E., Marrie, R. A., Robertson, N., La Rocca, N., Uitdehaag, B., van der Mei, I., Wallin, M., Helme, A., Angood Napier, C., Rijke, N., & Baneke, P. (2020). Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Multiple Sclerosis Journal*, *26*(14), 1816–1821. https://doi.org/10.1177/1352458520970841

Williams, R., Buchheit, C. L., Berman, N. E. J., & LeVine, S. M. (2012). Pathogenic implications of iron accumulation in multiple sclerosis. *Journal of Neurochemistry*, *120*(1), 7–25. https://doi.org/10.1111/j.1471-4159.2011.07536.x

Winston, G. P. (2012). The physical and biological basis of quantitative parameters derived from diffusion MRI. *Quantitative Imaging in Medicine and Surgery*, *2*(4), 254. https://doi.org/10.3978/j.issn.2223-4292.2012.12.05

Yaghmaie, N., Syeda, W. T., Wu, C., Zhang, Y., Zhang, T. D., Burrows, E. L., Brodtmann, A., Moffat, B. A., Wright, D. K., Glarin, R., Kolbe, S., & Johnston, L. A. (2021). QSMART: Quantitative susceptibility mapping artifact reduction technique. *NeuroImage*, *231*, 117701. https://doi.org/10.1016/j.neuroimage.2020.117701

Zhan, L., Leow, A. D., Jahanshad, N., Chiang, M.-C., Barysheva, M., Lee, A. D., Toga, A. W., McMahon, K. L., de Zubicaray, G. I., Wright, M. J., & Thompson, P. M. (2010). How does angular resolution affect diffusion imaging measures? *NeuroImage*, *49*(2), 1357–1371. https://doi.org/10.1016/j.neuroimage.2009.09.057

Zhang, Y., Gauthier, S. A., Gupta, A., Chen, W., Comunale, J., Chiang, G. C.-Y., Zhou, D., Askin, G., Zhu, W., Pitt, D., & Wang, Y. (2016). Quantitative Susceptibility Mapping and R2* Measured Changes during White Matter Lesion Development in Multiple Sclerosis: Myelin Breakdown, Myelin Debris Degradation and Removal, and Iron Accumulation. *American Journal of Neuroradiology*, *37*(9), 1629–1635. https://doi.org/10.3174/ajnr.A4825

Chapter 2

Investigating sensitivity and correlation between microstructural MRI measures in multiple sclerosis white matter lesions, normal appearing white matter and deep gray matter regions

Publication status

This chapter (and Chapter 3) are part of the abstract entitled '*Correlation Between Microstructural and Volumetric MRI Measures in Multiple Sclerosis White Matter Lesions, Normal Appearing White Matter, and Deep Gray Matter Regions'* accepted for presentation at the **Society for Neuroscience conference** in 2022. The manuscript is currently under internal review before submission to a peer-reviewed journal.

2.1 Abstract

Due to the complex pathophysiology in multiple sclerosis (MS), different imaging techniques can be used to leverage their unique sensitivities to particular tissue characteristics. And by combining multiple quantitative MRI (qMRI) methods, more information is likely to be obtained. However, the likelihood of encountering redundancy when recruiting a large number of qMRI metrics also increases. Using a cohort of 79 persons with MS, this cross-sectional study aimed to investigate 10 qMRI metrics (axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD), fractional anisotropy (FA), myelin water fraction (MWF), intra- and extra- cellular water fraction (IEWF), geometric mean T2 (gT2), T1-weighted-by-T2-weighted (T1w/T2w) ratio, R2* and QSM)) in their sensitivity and relationships across three tissue classes: normal appearing white matter (NAWM), white matter lesion (WML) and deep gray matter (DGM). These tissue classes were used to extract signals from the 10 qMRI maps, yielding a total of 30 variables. Individual qMRI metric in

different tissue classes were compared pairwise using the Wilcoxon Signed-Rank test. The qMRI metrics were further evaluated for their ability to distinguish WML from NAWM using the Receiver Operating Characteristic (ROC) Curve. Lastly, the correlation between qMRI metrics across tissue classes were examined using Spearman rank correlation analysis. Even after adjusting for multiple comparisons using the false discovery rate (FDR) correction, 28/30 comparisons remain significant ($p_{BH\text{-}FDR} < 0.05$). All qMRI metrics yield a better performance than a random classifier on the binary classification task (i.e., area under the curve > 0.5), with some even achieving a perfect score of 1. The Spearman rank correlation after adjusting the significance level using the FDR correction, yielded 179 significant (out of 435 total) correlations ($p_{BH\text{-}FDR} < 0.05$). Despite the moderate number of correlations, the overall pattern with a more careful inspection shows the complementarity between these qMRI metrics in the three tissue classes for MS. With the high sensitivity of qMRI metrics and the potentially unique information captured for each qMRI metric, these empirical findings suggest a multi-modal approach forward to better characterize MS as a disease.

2.2 Introduction

Multiple sclerosis (MS) is widely recognized as an autoimmune disease characterized by the invasion of the central nervous system by autoreactive immune cells, primarily targeting oligodendrocytes and myelin, leading to inflammation, edema, demyelination and axonal injury (Compston & Coles, 2002). At the same time, iron is dysregulated and accumulated in the subcortical regions (Stankiewicz et al., 2014), toxifying the environment, and causing further damage. Magnetic resonance imaging (MRI) is a noninvasive imaging technique that excels at soft tissue contrast. In diagnostics in neurodegenerative diseases, MRI often plays a role mainly in detecting brain atrophy and lesions (Young et al., 2020). Specifically in MS, the information provided by conventional MRI images (e.g., T1-weighted images, T2-weighted images) is extremely valuable. The specificity and sensitivity of MS diagnosis are significantly improved with the aid of MRI. It visualizes lesion morphology, patterns and spatial distributions to help classify diseases with similar lesion profiles. (Filippi et al., 2019). Since the first inclusion of MRI in the MS diagnostic criteria (McDonald et al., 2001) to the most recent revision in 2025 (Montalban et al., 2025), MRI has played an increasing role, and has started to include certain advanced imaging methods (e.g., quantitative MRI (qMRI)) alongside routinely collected MRI.

Despite conventional MRI demonstrating the potential for more accurate diagnosis with emerging imaging markers and advances in artificial intelligence (Rocca et al., 2024), it relies heavily on spatial features rather than directly quantifying the degree of neurodegenerative damage (Granziera et al., 2021). qMRI, on the other hand, demonstrates its value in enhancing image contrast with respect to various microstructural changes in a repeatable and quantifiable way, and therefore holds promise for making MS diagnosis and prognosis more comprehensive (Granziera et al., 2021). Besides, it provides a reasonable resolution for structural visualization of the brain,

and qMRI has an exceptional advantage: voxels are intrinsically normalized and physically interpretable (Tranfa et al., 2022), meaning that the outputs are not just images but spatially encoded data maps of radiomic features, which can also be theoretically examined and analyzed without the spatial context. Over the last few decades, emerging qMRI methods have been developed. Diffusion tensor imaging (DTI), in the field of diffusion imaging, is one of the earlier developed qMRI techniques and remains widely used in brain research (O'Donnell & Westin, 2011) to produce maps of: axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD) and fractional anisotropy (FA). The diffusion of water molecules depends on many factors, such as size, tissue and extracellular structures. (Beaulieu, 2002; le Bihan et al., 2001; Syková & Nicholson, 2008). In MS, it can be used to measure the degree of microstructural damage and assess structural connectivity (Filippi et al., 2016; Sbardella et al., 2013).

Myelin water imaging (MWI), as the name suggests, is an imaging technique designed to sensitize myelin content in the brain. This method is based on the relationship between the MR relaxation curves of the water molecules and the surrounding tissue structures. Water molecules that are trapped in the bounding pool (e.g., myelin) have a relatively short T2 (below 50 ms) in the entire T2 spectrum (15 ms to 2000 ms) (Whittall et al., 1997). Therefore, the myelin water fraction (MWF) can be calculated as the ratio of short T2 content to the entire T2 distribution (Prasloski et al., 2012b). MWF has been shown to be quite myelin-specific (Laule et al., 2006, 2008; Laule & Moore, 2018; A. L. MacKay & Laule, 2016; Sandrone et al., 2023), and has been found to correlate with the clinical presentation of MS (Edwards et al., 2022; Kolind et al., 2012), and has been used to monitor lesion activities throughout the course of MS (Vargas et al., 2015; Vavasour et al., 2009). In addition, the intra- and extra- cellular water fraction (IEWF) and geometric mean T2 (gT2) can be extracted from the same imaging sequence. IEWF is calculated by taking the contribution of

the moderate T2 component (80 to 200 ms) from the overall spectrum. In theory, the water within the intra- and extra-cellular space makes up the majority of the total water content in white matter (WM) (Tait et al., 2008; Zhou et al., 2023). Although the number of studies on MS and IEWF is limited, the negative association between IEWF and age was found in both healthy (Zhou et al., 2023) and MS cases (implicitly shown by the positive relationship between cerebrospinal fluid fraction and age) (Zhou et al., 2021). On the other hand, gT2 summarizes the entire spectrum with one single value. In general, gT2 is more reproducible than MWF (Meyers et al., 2013).

Following the same concept of a short relaxation time according to myelin characteristics, T1-weighted-by-T2-weighted (T1w/T2w) ratio was originally proposed to enhance cortical myelin signals (Glasser & van Essen, 2011). However, a potentially sensitive and useful tool to characterize WM abnormalities, the myelin specificity of T1w/T2w ratio signals in WM regions has been debated for some time (Arshad et al., 2017; Uddin, Figley, & Figley, 2018; Uddin, Figley, Marrie, et al., 2018b), and a more recent study using histochemical comparison confirmed that this technique is not specific to myelin content (Sandrone et al., 2023). Nonetheless, T1w/T2w ratio has been recruited in many MS studies (Boaventura et al., 2022; Cooper et al., 2021; Hannoun et al., 2022; Righart et al., 2017). Most importantly, the cost of the metric is negligible because the two imaging sequences are typically standardized and implemented across all MR vendors and are already adopted in most MRI clinical protocols.

Demyelination is the main feature of MS, but not the full picture. As mentioned, elevated iron concentrations in the parenchyma and deep gray matter regions have become another important feature of MS pathology. R2* ($1/T2^*$) and quantitative susceptibility mapping (QSM) are two metrics used to quantify iron levels in the brain (Zhang et al., 2016b). R2* mapping is a passive method to probe local field inhomogeneity because effective T2 (T2*) is shortened when

the local susceptibility of a voxel is not homogeneous. Although studies have reported a strong relationship between the R2* signal and iron content (Langkammer et al., 2012; Wood et al., 2005), R2* can also be sensitive to myelin (Kor et al., 2019), making R2* in pathophysiologically complex environments, like in the case of MS (especially with the co-existence of iron and myelin), difficult to interpret. QSM is an imaging method that can quantify voxel-level susceptibility by modelling each voxel as a sum of dipoles (Schweser et al., 2016; Wang & Liu, 2015). In certain deep gray matter regions, it is quite successful in detecting iron content in MS (Stüber et al., 2016) and longitudinally assessing abnormal changes in MS (Hagemeier et al., 2018).

The aforementioned qMRI metrics were shown to be sensitive for detecting brain WM abnormalities that are highly relevant to MS. Unfortunately, the majority of them have not yet been widely clinically adopted for various reasons, including the scan time required to add even a subset (never mind all) of them to the routine MS MRI protocol (Granziera et al., 2021). Therefore, determining the uniqueness of the information provided by the techniques is important yet easily overlooked. Especially, these technologies are advantageous for opening the door to study designs with higher-dimensional analyses. Redundant data should be omitted to promote efficiency in data collection and enhance the clarity of the results. Previously, Uddin et al. reported that several qMRI metrics (i.e., MWF, T1w/T2w ratio, AD, RD, MD and FA) across 25 brain structures (WM and subcortical gray matter) with healthy subjects have shown that those metrics all have comparable contribution in sensitizing underlying microstructures, yielding complementary information to help characterize heterogeneous sub-compartments of the brain (Uddin et al., 2019). Furthermore, in a similar scope but in the context of MS, Lipp et al. compared a smaller group of metrics (i.e., FA, RD, MTR and MWF) and found that, again, this group of qMRI metrics were complementary to each other by offering slightly different insights into the microstructural damages (i.e.,

sensitivity to two different types of lesions) (Lipp et al., 2019). However, previous studies examined only a subset of qMRI metrics across modalities and did not explore much on inter-region-of-interest (ROI) associations between metrics. Therefore, building on the momentum these studies have generated, this paper aims to broaden the scope by analyzing a larger set of qMRI metrics across three tissue classes in an MS cohort.

2.3 Methods

Dataset summary

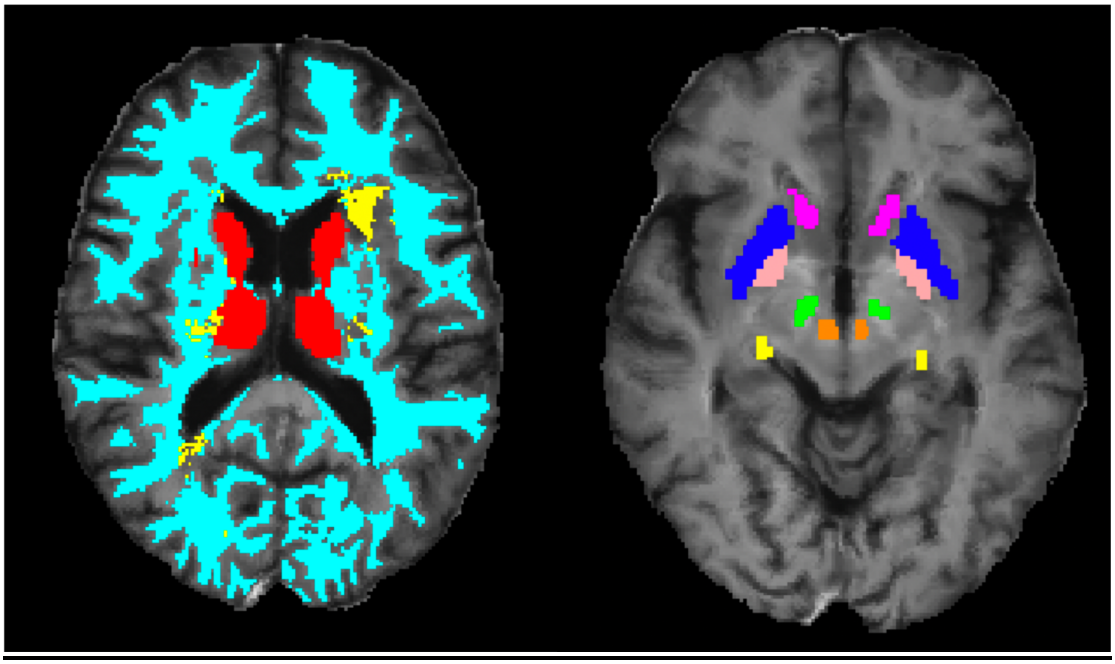
This study was approved by the University of Manitoba Health Research Ethics Board and consent from the participants was collected before their enrollment. The data used for the subsequent analyses were obtained from 79 pwMS (67 females) for whom a full battery of qMRI maps (outlined below) was available within the follow-up visit of The Comorbidity and Cognition in Multiple Sclerosis (CCOMS) Study (Uddin et al., 2022). All pwMS were previously diagnosed with multiple sclerosis using the current McDonald criteria at the time of diagnosis. Participants' median expanded disability status scale is 3.5 (IQR: [3,5.5]), and mean age is 50.6 (range: 22.8 – 79.0). For more details on the inclusion criteria, one may refer to our previously published imaging protocol (Uddin et al., 2022).

qMRI map pre-processing and ROI mask computation

This study included a total of 10 metrics (i.e., AD, RD, MD, FA, MWF, IEWF, gT2, T1w/T2w ratio, R2* and QSM) and three brain regions of interest (Figure 2-1): 1) normal appearing white matter (NAWM) 2) white matter lesion (WML) 3) Deep gray matter (DGM). Global WM and WML masks were previously calculated for each participant in the MS cohort of the CCOMS Study (Uddin et al., 2022), so the NAWM masks were created by subtracting each person's WML

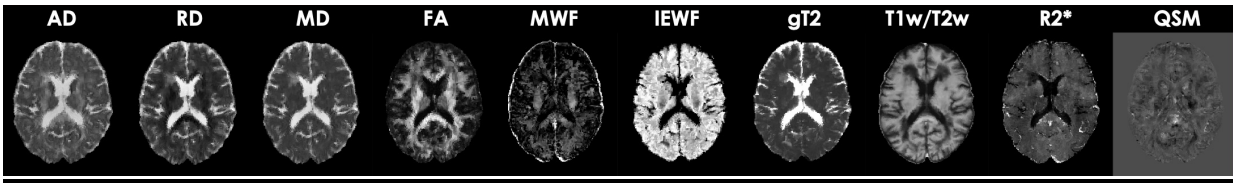
mask from their global WM mask; and the DGM mask (Figure 2-1, right) was generated by combining six subcortical gray matter structures from the MNI-space "EvePM" Atlas (Lim et al., 2013), that have been shown with excessive iron accumulation in MS: Thalamus, Globus pallidus, Putamen, Caudate, Substantia nigra and Red nucleus (Blazejewska et al., 2015; Uddin et al., 2016; Williams, Buchheit, Berman, & Levine, 2012). In addition, the trade-off between the reduced field-of-view near the inferior cerebellar region and imaging time during the MWI acquisition stage imposed an apparent brain mask on each subject, to ensure only valid brain regions were considered for ROI mean calculation (i.e., virtually excluding '0' entries from the background) for MWI-derived metrics (MWF, IEWF, gT2). This apparent brain mask was computed by applying the 'imclose' function in MATLAB (version R2020a; The MathWorks Inc., Natick, MA, USA) with a spherical element of a radius of 10 voxels on the MWF map. The ten qMRI metrics of one representative participant are displayed in the axial views in Figure 2-2. For more in-depth details of the imaging sequence parameters and pre-processing pipelines, one can visit the imaging protocol paper for the CCOMS study (Uddin et al., 2022). WM and DGM ROI masks were inspected so that these ROIs do not overlap. Any overlaps between WM and DGM were excluded from WM. Then the final NAWM and WML masks were computed by simply excluding and overlapping the adjusted (when necessary) WM mask with the lesion mask.

Figure 2-1: Illustration of NAWM, WML and DGM for one random MS participant.



(Left) color coded tissue ROIs overlaid on the participant's T1-weighted image, cyan: NAWM, yellow: WML, red: DGM. (Right) DGM mask is made up of six DGM regions, magenta: Caudate nucleus, blue: Putamen, pink: Globus pallidus, green: Substantia nigra, orange: Red nucleus and yellow, thalamus.

Figure 2-2: Sample axial slices of 10 quantitative maps of one random MS participant of the study.



Statistical analyses

Box-and-whisker plots and the associated descriptive statistics (i.e., mean, standard deviation (Std) and coefficient of variation (CoV)) for 30 variables were computed to provide an overview of the data distribution across all qMRI metrics in all three tissue classes. Two-tailed Wilcoxon Signed-Rank tests were performed on all possible ROI pairs within each metric, corrected using the step-up Benjamini-Hochberg false discovery rate (BH-FDR) procedure (Benjamini & Hochberg, 1995). Then, to assess the sensitivity of each qMRI metric with respect to WML and NAWM classification, a ROC/AUC analysis was performed. Finally, Spearman rank correlations were performed for all qMRI comparisons across all these tissue classes, controlling for age and sex, and the results were corrected for multiple comparisons using the BH-FDR method (Benjamini & Hochberg, 1995). These Spearman rank correlations were calculated alongside another work simultaneously, using the same set of ROIs but with a different scope and goal – exploring the relationship between brain volumetric measures and qMRI (Chapter 3). Therefore, BH-FDR correction was instead applied to the entire correlation matrix (with 93 more correlations by incorporating 3 additional volumetric variables). The significance level cutoff was set to an alpha value to 0.05 after the BH-FDR correction. The number of associated significant correlations of each metric, and their shared percentage, calculated as the number of the associated significant correlations of each metric divided by the total unique significant correlations from the correlation map among qMRI correlations only, were computed to gain an entry-level understanding of how the significant correlations was distributed. Later, the corrected and controlled Spearman rank correlation matrix can be viewed from three perspectives by decomposing the master matrix into the following categories: same metric, different ROIs (SMDR); different metrics, same ROI (DMSR); and different metrics, different ROIs (DMDR). In addition, DMSR was decomposed into individual

tissue classes to narrow the perspective. The above analyses were done in MATLAB (version R2020a; The MathWorks Inc., Natick, MA, USA). Note that the nonparametric comparison and correlation method is used because not all distributions pass the Kolmogorov–Smirnov test (Massey, 1951). Though rank correlation holds a relatively lower statistical power (if parametric assumptions are met), the nonparametric correlation: 1) does not require a normal distribution, 2) is less susceptible to outliers, and 3) will better detect (i.e., not penalize) nonlinear relationships (which are also of interest in this case).

2.4 Results

Box-and-whisker plots and Wilcoxon Signed-Rank tests

In Figure 2-3, box-and-whisker plots of 10 qMRI metrics are shown together. Descriptive statistics (i.e., mean, Std and CoV) of all distributions are summarized in Table 2-1. Compared to NAWM and DGM box-and-whisker plots, WML shows a much wider range, as demonstrated in the space between the first and third quartiles for most metrics, and WML plots tended to have more outliers and skewed distributions. According to the two-tailed Wilcoxon Signed-Rank test results summarized in Table 2-1, except for MWF (between NAWM and DGM) and IEWF (between WML and DGM), the remaining 28 (out of 30) pairs of comparisons between tissue classes (within each qMRI metric) were statistically significant.

Figure 2-3: Box-and-whisker plots of 10 qMRI metrics in NAWM, WML and DGM.

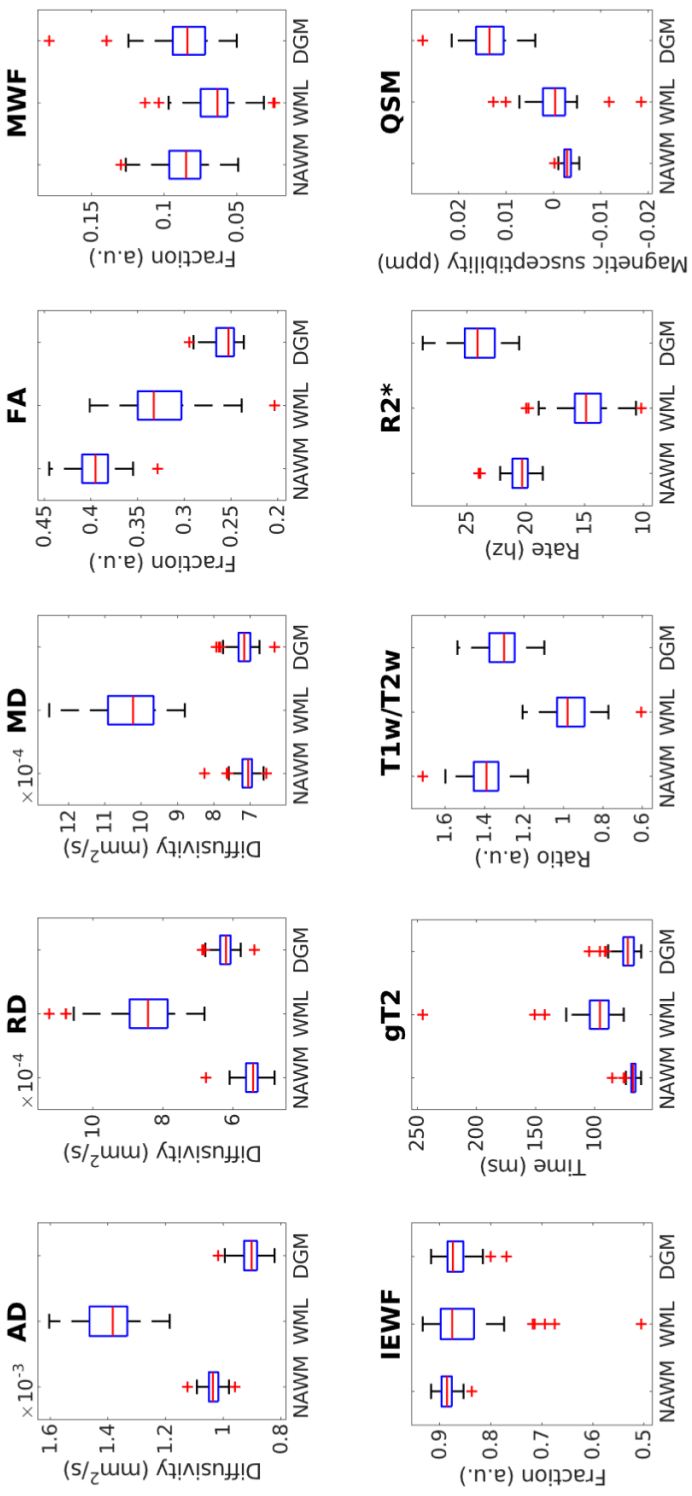


Table 2-1: Mean, Std and CoV of all 30 qMRI-ROI pairings across participants.

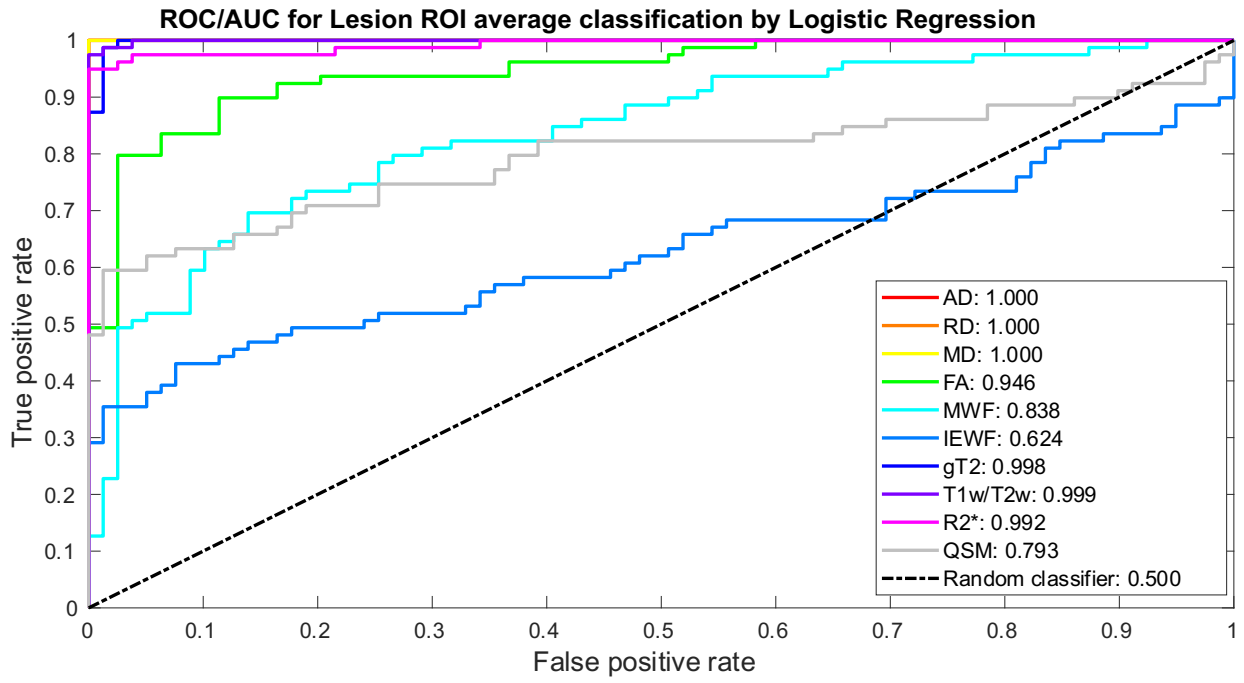
Quantitative metrics (units)	Tissue classes	mean	Sample standard deviation	Coefficient of variation	Wilcoxon Signed Rank Test	Corrected p-value
AD (mm^2/s)	*NAWM*	0.0010	0.00003	0.026	*NAWM-WML*	**2.24E-14**
	WML	0.0014	0.00009	0.062	*WML-DGM*	**2.24E-14**
	DGM	0.0009	0.00004	0.040	*DGM-NAWM*	**2.24E-14**
RD (mm^2/s)	*NAWM*	0.0005	0.00003	0.055	*NAWM-WML*	**2.24E-14**
	WML	0.0009	0.00009	0.103	*WML-DGM*	**2.24E-14**
	DGM	0.0006	0.00003	0.045	*DGM-NAWM*	**2.24E-14**
MD (mm^2/s)	*NAWM*	0.0007	0.00003	0.038	*NAWM-WML*	**2.24E-14**
	WML	0.0010	0.00008	0.079	*WML-DGM*	**2.24E-14**
	DGM	0.0007	0.00003	0.042	*DGM-NAWM*	**1.23E-02**
FA (a.u.)	*NAWM*	0.3939	0.02138	0.054	*NAWM-WML*	**5.07E-14**
	WML	0.3258	0.03950	0.121	*WML-DGM*	**1.01E-13**
	DGM	0.2564	0.01330	0.052	*DGM-NAWM*	**2.24E-14**
MWF (a.u.)	*NAWM*	0.0865	0.01587	0.183	*NAWM-WML*	**2.01E-13**
	WML	0.0650	0.01628	0.250	*WML-DGM*	**1.25E-11**
	DGM	0.0868	0.02058	0.237	*DGM-NAWM*	4.83E-01
IEWF (a.u.)	*NAWM*	0.8851	0.01489	0.017	*NAWM-WML*	**1.69E-04**
	WML	0.8559	0.06669	0.078	*WML-DGM*	8.41E-01
	DGM	0.8671	0.02642	0.030	*DGM-NAWM*	**7.69E-09**
gT2 (ms)	*NAWM*	67.6399	3.64860	0.054	*NAWM-WML*	**2.24E-14**
	WML	98.7070	21.02011	0.213	*WML-DGM*	**2.50E-14**
	DGM	72.9367	8.13183	0.111	*DGM-NAWM*	**2.88E-11**
T1w/T2w ratio (a.u.)	*NAWM*	1.3908	0.10274	0.074	*NAWM-WML*	**2.24E-14**
	WML	0.9641	0.10224	0.106	*WML-DGM*	**2.24E-14**
	DGM	1.3107	0.10000	0.076	*DGM-NAWM*	**5.19E-14**
R2* (Hz)	*NAWM*	20.4878	1.01227	0.049	*NAWM-WML*	**2.24E-14**
	WML	14.7693	1.93013	0.131	*WML-DGM*	**2.24E-14**
	DGM	24.0461	1.75999	0.073	*DGM-NAWM*	**2.36E-14**
QSM (ppm)	*NAWM*	-0.0030	0.00102	*	*NAWM-WML*	**6.34E-08**
	WML	-0.0002	0.00420	*	*WML-DGM*	**2.24E-14**
	DGM	0.0137	0.00466	*	*DGM-NAWM*	**2.24E-14**

*BH-FDR corrected two-tailed Wilcoxon Signed-Rank tests' p-values are shown on the right of the Table, where significant results were shown in bold fonts *Note: QSM's mean ranges from negative to positive, therefore the coefficient of variation was not calculated.*

ROC/AUC

In Figure 2-4, an ROC/AUC analysis was performed to further assess how well the 10 qMRI metrics distinguish WML from NAWM using the mean measure from each ROI. AD, RD and MD achieved perfect classification scores. T1w/T2w ratio, gT2 and R2* have almost perfect AUC scores of 0.999, 0.998 and 0.992, respectively. FA and MWF have relatively lower AUCs (0.946 and 0.838, respectively). Compared with the metrics mentioned, QSM has a moderate AUC (0.793) and IEWF performs only slightly better than chance at distinguishing mean values between WML and NAWM (AUC = 0.624).

Figure 2-4: ROC curves of 10 qMRI metrics in classifying white matter lesions from normal appearing white matter.

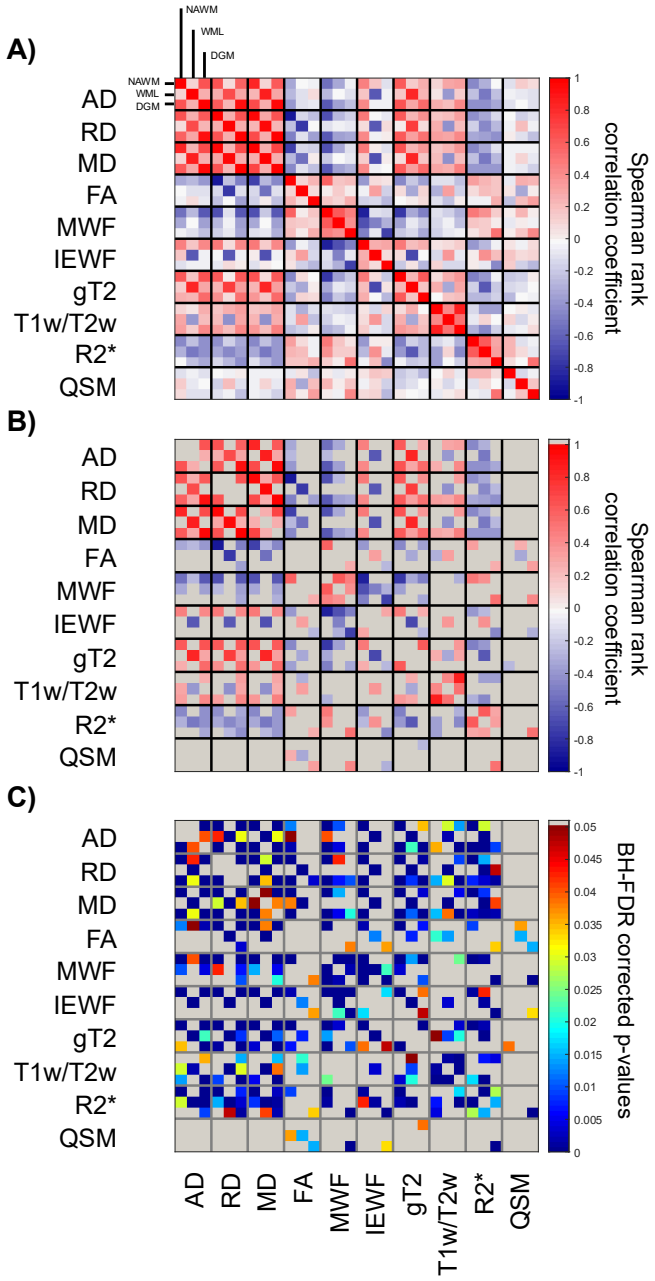


AUC score for each qMRI metric is calculated in the legend (bottom-right corner). AD (red), RD (orange) sensitivity and accuracy curves are completely hidden behind MD (yellow) because they all achieved perfect scores. A dotted random classifier line is plotted as a reference.

Spearman rank correlation

The pair-wise Spearman rank correlations, significant p-values after BH-FDR correction and significant Spearman rank correlations were visualized in Figure 2-5. Table 2-2 summarizes the Spearman rank correlation coefficients and associated significant levels. After controlling for the covariates (i.e., age and sex) and correcting for multiple comparisons using the BH-FDR method, 179 (out of 435) unique pairwise correlations (41%) remain significant ($p < 0.05$). The distribution of the number of associated significant correlations and shared percentage (over 179), of each metric is visualized in Figure 2-6. Briefly, MD has the most shared significant correlations among all qMRI metrics with (46, partially shared 25.7% of significant pairs), followed by RD (45, 25.1%) and AD (43, 24.0%), and QSM has the least (7, 3.9%). The corrected Spearman rank correlation matrix and its decomposed categories (i.e., SMDR, DMSR and DMDR) are shown in Figure 2-7. Since the correlation is diagonally symmetric, only the upper triangle portion of the matrices was shown. As mentioned, DMSR is further decomposed into tissue classes in Figure 2-8. By inspecting the correlation between different metrics after correction across the three tissue classes, NAWM, WML and DGM hold 28, 27 and 27 significant relationships, respectively. For SMDR, most metrics show significant positive correlations with themselves across ROIs except FA (i.e., none is significant) and QSM (i.e., none is significant). DMSR and DMDR, on the other hand, span most of the upper triangle, and include negative correlations with other metrics/ROIs. Compared to DMDR, DMSR's elements on average have a higher magnitude (0.53 vs 0.39). The significant correlations between every pair of tissue classes are shown in Figure 2-9 by separating the combined DMDR and SMDR matrices: NAWM-WML, WML-DGM and DGM-NAWM. The significant correlations in tissue pair skew toward DGM and NAWM (WML-NAWM (26) : WML-DGM (24) : DGM-NAWM (47)).

Figure 2-5: Age- and sex-corrected Spearman rank correlation matrices of the 10 qMRI metrics across 3 ROIs.



(A) Sex and age adjusted Spearman rank correlation matrix, (B) significant Spearman rank correlation matrix and associated (C) significant BH-FDR corrected p-values. Gray boxes indicate the entry is no longer significant after the BH-FDR correction.

Table 2-2: A hybrid display of the Spearman rank rho and the associated p-values.

Spearman Rho / p values		AD			RD			MD			FA			MWF			IEWF			gT2			T1w/T2w ratio			R2*			QSM		
		NAWM	WML	DGM	NAWM	WML	DGM	NAWM	WML	DGM	NAWM	WML	DGM	NAWM	WML	DGM	NAWM	WML	DGM	NAWM	WML	DGM	NAWM	WML	DGM	NAWM	WML	DGM	NAWM	WML	DGM
AD	NAWM			0.64	0.65		0.57	0.83		0.61	-0.32			-0.55	-0.33		0.44			0.60		0.28		0.29	0.32	-0.45	-0.29				
	WML	0.48		0.27	0.27	0.68	0.28		0.85	0.28	-0.26			-0.27				-0.64			0.84			-0.38			-0.52				
	DGM	0.00	0.04		0.60		0.91	0.68		0.97	-0.39			-0.64	-0.36		0.44			0.65	0.30	0.56	0.28		0.44	-0.42	-0.44	-0.33			
RD	NAWM	0.00	0.04	0.00			0.64	0.95	0.29	0.63	-0.90			-0.66	-0.27		0.48			0.61		0.40			0.33	-0.42	-0.32				
	WML	0.61	0.00	0.09	0.05				0.95			-0.77						-0.64			0.74			-0.44			-0.46	-0.26			
	DGM	0.00	0.03	0.00	0.00	0.06		0.68	0.28	0.98	-0.49		-0.36	-0.64	-0.35	-0.33	0.43			0.60	0.34	0.51	0.32	0.29	0.43	-0.34	-0.37	-0.41			
MD	NAWM	0.00	0.08	0.00	0.00	0.08	0.00		0.26	0.69	-0.75			-0.67	-0.31		0.50			0.65		0.37			0.35	-0.46	-0.34				
	WML	0.49	0.00	0.05	0.03	0.00	0.03	0.05		0.27	-0.27	-0.56						-0.69			0.84			-0.43			-0.53	-0.27			
	DGM	0.00	0.03	0.00	0.00	0.07	0.00	0.00	0.04		-0.45			-0.64	-0.36	-0.30	0.43			0.62	0.33	0.54	0.31	0.27	0.45	-0.38	-0.40	-0.38			
FA	NAWM	0.01	0.05	0.00	0.00	0.07	0.00	0.00	0.04	0.00				0.56			-0.39			-0.47		-0.35				0.34				0.28	
	WML	0.98	0.44	0.53	0.35	0.00	0.45	0.37	0.00	0.51	0.41							0.32			-0.34		-0.30	0.32						-0.32	
	DGM	0.81	0.59	0.61	0.10	0.91	0.00	0.21	0.87	0.06	0.07	0.59				0.27			-0.28									0.28			0.32
MWF	NAWM	0.00	0.04	0.00	0.00	0.41	0.00	0.00	0.15	0.00	0.00	0.99	0.18		0.60	0.46	-0.87			-0.76	-0.32	-0.49			-0.29	0.47	0.40				
	WML	0.01	0.98	0.00	0.04	0.75	0.01	0.01	0.91	0.00	0.20	0.30	0.42	0.00		0.38	-0.69	-0.50	-0.30	-0.36											
	DGM	0.75	0.51	0.06	0.21	0.69	0.01	0.41	0.63	0.02	0.12	0.87	0.04	0.00	0.00		-0.48		-0.74			-0.33						0.52			0.42
IEWF	NAWM	0.00	0.26	0.00	0.00	0.58	0.00	0.00	0.37	0.00	0.00	0.81	0.22	0.00	0.00	0.00			0.31	0.48		0.27				-0.40	-0.27				
	WML	0.21	0.00	0.85	0.86	0.00	0.71	0.89	0.00	0.75	0.63	0.01	0.33	0.45	0.00	0.63	0.06				-0.71			0.37			0.40				
	DGM	0.49	0.85	0.76	0.58	0.63	0.95	0.51	0.91	0.91	0.74	0.23	0.04	0.39	0.02	0.00	0.02	0.60				-0.26						-0.44			-0.28
gT2	NAWM	0.00	0.19	0.00	0.00	0.99	0.00	0.00	0.65	0.00	0.00	0.55	0.87	0.00	0.00	0.14	0.00	0.62	0.17			0.62			0.37	-0.48	-0.38				
	WML	0.39	0.00	0.02	0.06	0.00	0.01	0.08	0.00	0.01	0.11	0.01	0.88	0.01	0.56	0.30	0.11	0.00	0.71	0.07			0.26	-0.34	0.30		-0.63				
	DGM	0.03	0.08	0.00	0.00	0.59	0.00	0.00	0.26	0.00	0.01	0.49	0.65	0.00	0.12	0.01	0.04	0.91	0.05	0.00	0.18								-0.27		
T2w r	NAWM	0.41	0.73	0.04	0.26	0.06	0.01	0.24	0.13	0.02	0.49	0.02	0.44	0.49	0.30	0.23	0.83	0.35	0.50	0.62	0.05	0.99		0.41	0.86		-0.35	-0.36			
	WML	0.03	0.00	0.13	0.17	0.00	0.03	0.09	0.00	0.04	0.51	0.01	0.06	0.19	0.77	0.63	0.63	0.00	0.39	0.06	0.01	0.20	0.00		0.50						
	DGM	0.01	0.31	0.00	0.01	0.15	0.00	0.01	0.13	0.00	0.07	0.19	0.97	0.02	0.08	0.75	0.37	0.68	0.29	0.00	0.02	0.11	0.00	0.00		-0.31	-0.45				
R2*	NAWM	0.00	0.10	0.00	0.00	0.21	0.01	0.00	0.10	0.00	0.01	0.65	0.74	0.00	0.20	0.72	0.00	0.99	0.86	0.00	0.19	0.06	0.39	0.55	0.02		0.60	0.29			
	WML	0.03	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.08	0.11	0.76	0.00	0.30	0.36	0.04	0.00	0.65	0.00	0.00	0.06	0.01	0.50	0.00	0.00		0.32			
	DGM	0.91	0.09	0.01	0.33	0.05	0.00	0.41	0.04	0.00	0.26	0.26	0.03	0.30	0.86	0.00	0.77	0.13	0.00	0.78	0.15	0.22	0.00	0.99	0.24	0.03	0.01				0.51
QSM	NAWM	0.47	0.75	0.51	0.92	0.84	0.98	0.78	0.81	0.83	0.63	0.85	0.07	0.86	0.66	1.00	0.81	0.26	0.39	0.17	0.38	0.04	0.91	0.49	0.63	0.10	0.10	0.64			
	WML	0.35	0.63	0.89	0.27	0.15	0.43	0.80	0.42	0.63	0.04	0.01	0.14	0.09	0.98	0.19	0.70	0.67	0.25	0.24	0.46	0.99	0.97	0.39	0.78	0.49	0.57	0.96	0.30		
	DGM	0.57	0.49	0.50	0.46	0.74	0.09	0.76	0.63	0.21	0.08	0.96	0.01	0.86	0.73	0.00	0.66	0.52	0.03	0.98	0.87	0.71	0.63	0.91	0.60	0.72	0.55	0.00	0.69	0.19	

Only significant Spearman rank rho (Upper triangle, green) and associated p-values after BH-FDR correction (Lower triangle, blue) are displayed. Correlations that failed to reach significance are grayed out and diagonal values (i.e., the self-correlations) were left blank.

Figure 2-6: Bar graph of 10 qMRI metrics correlation and their shared percentages within 179 (out of 435) total unique significant correlations.

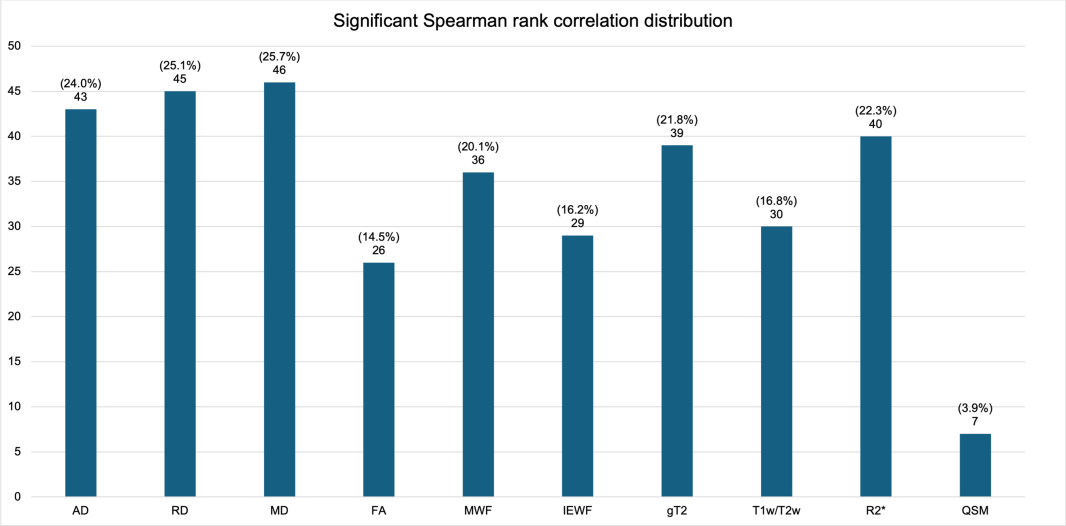
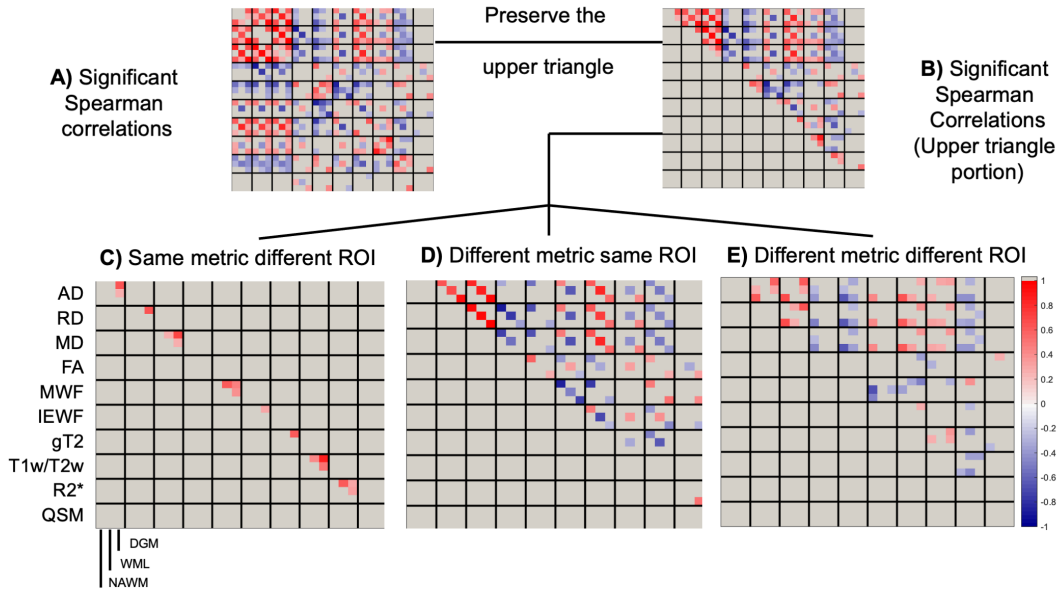
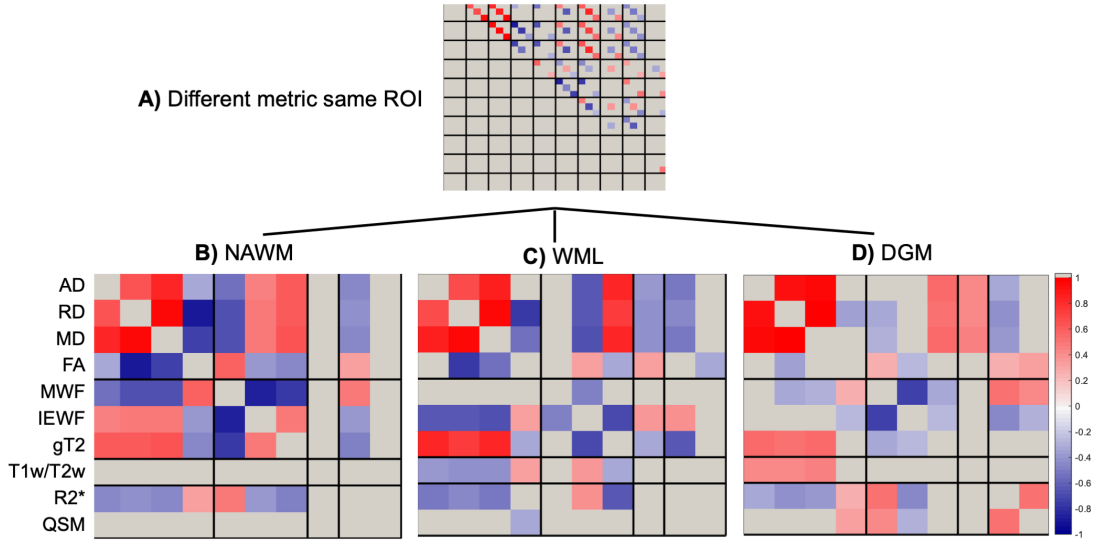


Figure 2-7: Decomposition of the unique components of the entire Spearman rank correlation matrix.



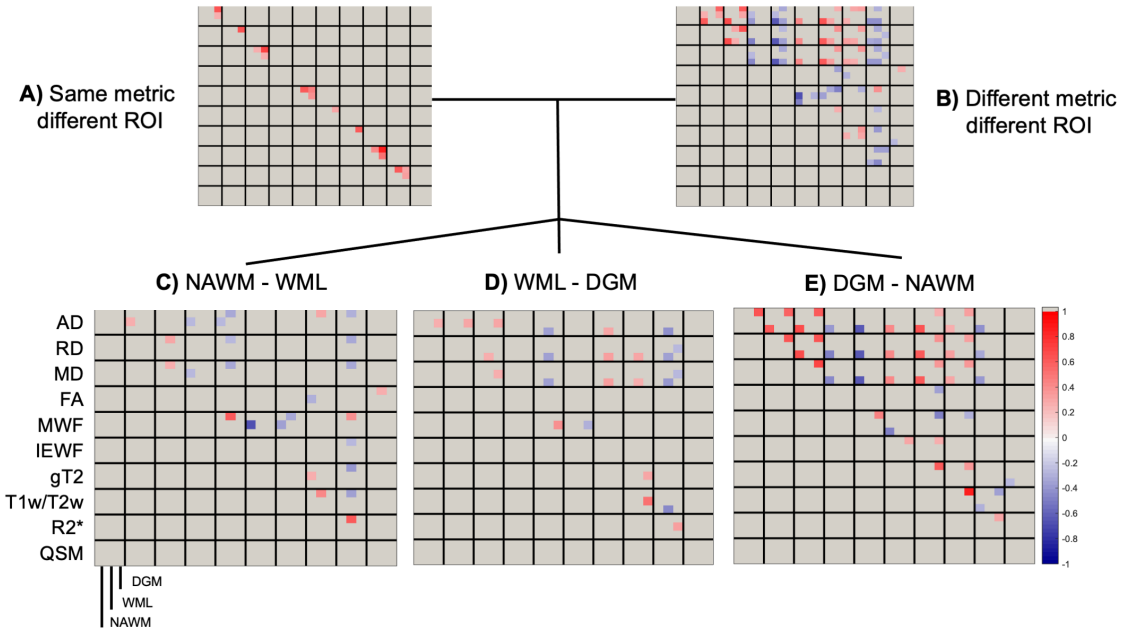
Breakdown of the (A) significant Spearman rank correlation into (B) symmetric upper triangular view and further split into (C) SMDR, (D) DMSR and (E) DMDR.

Figure 2-8: Decomposition of the DMSR components.



Breakdown of the (A) different metric same ROI matrix into (B) NAWM, (C) WML and (D) DGM.

Figure 2-9: Decomposition of the SMDR & DMDR components.



Breakdown of the [(A) same metric different ROI matrix + (B) different metric different ROI] into (C) NAWM-WML pair, (D) WML-DGM pair, and (E) DGM-NAWM pair.

2.5 Discussion

qMRI metrics show distinct sensitivity across ROIs

Before diving into the associations among metrics and their relationships within and across ROIs, these qMRI metrics exhibit unique patterns of signal hierarchy (signal distribution being ordered from highest to lowest) (Figure 2-3, and from the Wilcoxon Signed-Rank tests, Table 2-1), except for MWF and IEWF. The widespread significant results (from the Wilcoxon Signed-Rank tests, Table 2-1) have indicated that those qMRI metrics are quite sensitive not only to the differences within WM space between NAWM and WML, but also to the differences between DGM and WM in general, highlighting the findings independently shown previously with both healthy (Uddin et al., 2019) and MS (Lipp et al., 2019) cohorts. Most qMRI metrics in NAWM and DGM have similar CoVs and they are smaller than those in WML, except for R2* and QSM in DGM. This implies an inherent signal consistency in the NAWM and DGM ROIs. In addition, the relatively higher CoV in WML across qMRIs strongly reflects the heterogeneity in MS lesions known in the field (Lucchinetti et al., 2000).

Most qMRI metrics are capable of classifying mean values of WML from NAWM's

qMRI metrics has demonstrated their sensitivity in detecting changes in the brain that can benefit early MS diagnosis (Granziera et al., 2021; Lavrova et al., 2021; Rocca et al., 2024). However, how well qMRI can delineate the "boundary" between lesions and surrounding tissues remains an open question and an active area of research. Lipp et al. investigated the potential of a subset of qMRI in radiological lesions detection using ROC analysis (T1 hypointense and T2 hyperintense) and found MWF to be the more suitable candidate among other metrics (i.e., RD, FA and magnetic transfer ratio) in the classification task (Lipp et al., 2019). In our work, all qMRI metrics have an

AUC value greater than '0.5', and many qMRI metrics have demonstrated a superior ability (i.e., AUC > 0.95) to classify WML means from NAWM means in AD, RD, MD, FA (to a similar extent), gT2, T1w/T2w ratio and R2*. It is interesting to note how well the diffusion metrics do in this analysis. It is known that the MS autoimmune response at lesion sites damages tissues and cells. Diffusivity of water molecules and the fluid in those areas becomes less obstructed and has greater freedom to flow around compared to the healthy environment (Filippi et al., 2001; Rovaris et al., 2005; Werring et al., 1999). Diffusion MR was also probed with the feasibility of lesion detection in multiple sclerosis (Tadayon et al., 2016). Similarly, gT2 and T1w/T2w ratio are heavily affected as the structure surrounding the water molecules is compromised. R2*, on the other hand, can also classify WML out of NAWM very well. Due to abnormally high iron levels in active lesions resulting from the death of iron-rich oligodendrocytes and the unregulated release of ferritin into the environment (Zierfuss et al., 2023), WML can exhibit high R2* (Zhang et al., 2016b). Nonetheless, R2* can also be affected by myelin level (Bagnato et al., 2018; Zhang et al., 2016b). Since the distribution of MS phenotypes is weighted toward relapsing and remitting in our MS cohort (Uddin et al., 2022) and pwMS were not scanned within 30 days of relapse, one might expect that the majority of WML measures from participants may come from inactive lesions. That is a possible reason why R2* is lower in WML than in NAWM. On the other hand, contrary to the results shown previously by Lipp et al. (Lipp et al., 2019), MWF is not the best candidate in our case to distinguish between WML and NAWM. Though it is worth mentioning that the differences in the exact study design (e.g., sample sizes, ROI definitions, qMRI metric selection). However, although the MWF's AUC score (0.838) was still quite high in our study (and it potentially offers more unique information in WML regions, see next discussion), demyelination can occur in NAWM and MWF would still be measuring myelin (or even myelin debris (A. L. MacKay & Laule,

2016)) within the WML mask, leading to imperfect classification. Lastly, though IEWF and QSM in WML can be useful for inferring structural changes in WML relative to NAWM as indicated by the Wilcoxon Signed-Rank test results, their ROC curves went below the random classifier curve in higher false positive rate regions, hindering their ability to accurately classify WML from NAWM.

qMRI metrics share information in MS, but not completely redundant

Besides addressing the conventional problems in preparing for clinical implementation, such as comprehensive biological validation, and scan time reduction, sequences and pipelines harmonization (Granziera et al., 2021), one should also be aware of the level of redundancy among metrics when recruiting a large number of qMRI metrics. The latter issues become more serious when resources are limited (e.g., patient wait time and MRI workload in a clinic) so identifying potentially redundant qMRI metrics within the context of MS was one of the main motivations for the current study. 41% of unique pairwise correlations between metrics, ranging from moderate to high in magnitude, within and across ROIs, are significant after correction. Some hot spots (i.e., high correlation in magnitude) are clustered in the matrix regions with qMRI metrics derived from the same sequence and pre-processing pipeline. Especially in diffusion tensor imaging, AD, RD and MD have the highest numbers of shared correlation with other qMRI metrics (Figure 2-6), consistent with the previous findings with MS subjects (Lipp et al., 2019) and healthy human subjects with even finer ROI definition in the brain (Uddin et al., 2019), highlighting their lack of specificity. However, contradicting the study (possibly because of the difference between MS vs. healthy subjects) (Uddin et al., 2019), the current work has shown that correlations between pairs of qMRI metrics are found to be much stronger when they are in the same ROI (more so in NAWM

and WML) as DMSR has the highest mean correlation magnitude (0.53) among the three decomposed matrices. On the opposite, T1w/T2w ratio (with the fewest significant correlation) in NAWM (Figure 2-8B), and MWF (with the fewest significant correlation) in WML (Figure 2-8C) and QSM (with the fewest significant correlations) in all regions (Figure 2-5C) individually provides a unique perspective compared to the rest and potentially offers the most complementary information to the other qMRI metrics, in their respective context of regions. In addition, qMRI metrics dynamically share information with each other between different ROIs (Figure 2-7E). While many correlations are statistically significant, their relationships are not necessarily symmetrical (i.e., in terms of significance, magnitude and direction). That is, A in WML significantly correlates with B in NAWM does not imply that A in NAWM correlates significantly with B in WML. Even if some of the relationships do, the magnitude and the direction may not be identical. This finding further reinforces the idea that these qMRI metrics may hold rich information that complements one another, because it suggests the sensitivity and specificity of these qMRI are region-dependent, aligning with the study (Uddin et al., 2019).

The global association in microstructural measures across ROIs as measured by the same qMRI metric

As shown in the SMDR portion (Figure 2-7C) of the master Spearman rank correlation matrix, all significant relationships are unidirectional across ROIs (while FA and QSM do not have any significant correlation between ROIs). The positive correlations captured by the same qMRI metrics indicate that all of these ROIs are experiencing microstructural changes at least in the same direction, highlighting a global trend of microstructural changes in MS as an aftermath of the subtle interplay between neurodegeneration and repair mechanisms.

Interpreting qMRI metrics can be challenging

However, the underlying activities in different tissue classes (with fundamentally different underlying microstructural properties (e.g., myelin concentration, axon/cell densities)) may share similar MR characteristics. Therefore, a single qMRI metric can sensitize multiple types of microstructural changes simultaneously (i.e., limited specificity) (Granziera et al., 2021; Gulani & Seiberlich, 2020; Saltarelli et al., 2025; Tranfa et al., 2022). This outlines one of the major challenges in qMRI metrics – the interpretation. Briefly, even with a relatively specific qMRI metric like MWF as discussed earlier, which shows robust results in histological staining with myelin content, it has also shown sensitivity with iron, as a result of very similar relaxometry properties exhibited by two biological constituents (Birkl et al., 2019; Laule et al., 2006, 2008; Laule & Moore, 2018; A. L. MacKay et al., 2009; Sandrone et al., 2023). Similarly, R2*, a qMRI metric commonly used to "measure" iron concentration in DGM, can also be affected by myelin concentration (Kor et al., 2019). Particularly, the significant correlations between MWF and R2* in both NAWM and DGM (Figure 2-8) reflect this, especially the gray matter (GM), which is generally poorly myelinated (Mercadante & Tadi, 2026). Consistent with previous findings from same lab investigating the validity of T1w/T1w ratio in measuring myelin using healthy subjects (Uddin et al., 2019; Uddin, Figley, Marrie, et al., 2018b), and rigorous result shown by a recent histology study (Sandrone et al., 2023), MWF and T1w/T2w ratio were not at all significantly correlated with each other across all three ROIs in current work (Figure 2-8). Nonetheless, the colocalization of myelin and iron and their interplay in MS is just an intuitive example given their commonality to help readers understand. But in reality, the pathophysiology of MS is complex.

Inter-ROI correlations between qMRI metrics

Last but not least, the authors hope to draw readers' attention to the fact that information between ROIs can also be dynamically shared between qMRI metrics, as shown by the SMDR and DMDR portions of the Spearman rank correlation matrix (Figure 2-7C, E). Although these correlations are relatively weaker than DMSR (i.e., with a lower mean magnitude), DGM and NAWM ROI pair has the most significant correlations, and these correlations are consistently stronger than the correlations from the two other possible pairs of ROIs (Figure 2-9). In MS, besides more severe neurodegeneration seen in WML, mild abnormalities and inflammation have been observed in both DGM (Ontaneda et al., 2021; Vercellino et al., 2009) and NAWM (Maillard et al., 2011; Zhong & Lou, 2016). One possible explanation can be that, due to anterograde degeneration, the lesions originated in GM (e.g., dying cell bodies) extend/expand along the axons toward white matter, and vice versa with retrograde degeneration to reverse the direction of degeneration therefore relating DGM and NAWM (Kuceyeski et al., 2015). However, a longitudinal study with volumetric measures modeling the interaction between lesions (in the white matter) and subcortical gray matter regions has suggested that the progression of tissue damage in the two classes happens quite independently, questioning the contribution of anterograde/retrograde degeneration linkage between the two tissue classes (Fuchs et al., 2018). And interestingly, the NAWM-WML pair, are only weakly correlated with each other, despite the fact that those two ROIs are both in the WM space and they are more proximal to each other. Nevertheless, further investigation is needed to elucidate the underlying mechanisms of these correlations in particular.

Study limitations

This study offers a distinct and unique perspective on the relationships among existing qMRI metrics in the MS realm, but one must be aware of the caveats in interpreting the results.

ROI computational steps and choices (Combining DGM regions, WM/Lesion segmentation)

Combining the six small DGM regions into a single larger ROI was done to facilitate interpretation, reduce the number of comparisons, reduce variance and improve confidence in volume estimates, since DGM regions are relatively smaller than the other two tissue classes. However, some DGM regions (e.g., thalamus) are larger than others, so the combined DGM measurements (i.e., qMRI and Vol) in this study are systematically biased toward larger sub-structures. In MS, DGM regions have attracted increased interest in neuropathology such as neuroinflammation, iron deposition and regional atrophy, and not all the regions respond to MS the same way (Haider et al., 2014). Therefore, more in-depth work is needed to study each DGM region and elucidate the associations; however, the same argument goes toward NAWM and WML with known heterogeneity in those two regions. However, since the primary goal of the study is to explore empirical trends across three macroscopically defined brain regions without *a priori* hypotheses, the authors chose to constrain the number of ROIs. Distinguishing WML and NAWM mean values in ROC/AUC analyses is heavily dependent on the quality of automated WM and lesion tissue segmentation. Tissue segmentation can be performed in various ways, yielding different results (Agrawal & Sharma, 2014). Moreover, this plays a role in the definition of ROI.

Choices of imaging sequence parameters and metric processing pipelines

It is also worth noting that almost all the qMRI metrics included in this study can be computed from more than one imaging sequence and can have multiple different processing pipelines. In other words, depending on the choice of acquisition and most notably the fitting model, qMRI

maps of the same type can have different values (Weingärtner et al., 2022). For example, diffusion MR imaging has been shown to be highly reproducible when the same acquisition and post-processing pipeline is used (Heiervang et al., 2006), but scanner differences contribute significantly to DTI computation (Takao et al., 2011). The myelin water imaging method, and MWF in particular, is highly reproducible (Chan et al., 2023), and it correlates strongly with histological staining quantification of myelin (Laule et al., 2008), but can also be computed in several different ways (Alonso-Ortiz et al., 2015b; Deoni et al., 2008; Lee et al., 2021; Lee, Nam, et al., 2017; S.-H. Oh et al., 2013). On the other hand, T1w/T2w ratio maps have not yet gained a consensus on how to normalize the two images (Ganzetti et al., 2014; Scalco et al., 2020). A study has shown that the reproducibility of T1w/T2w ratio images is quite high, but there is still regional variability, partly due to bias field differences (Glasser et al., 2022; Nerland et al., 2021). QSM is a relatively younger technique and different labs have proposed their methods and pipelines for computation. QSM computed by calculation of susceptibility using multiple orientation sampling (COSMOS) with 3 orthogonal probing directions was once considered the gold standard (Wang & Liu, 2015), but other techniques using background field removal and field-to-source inversion have more recently been developed for single orientation acquisitions. We chose the QSMART pipeline (Yaghmaie et al., 2021) mainly because it generates vascular masks and uses the iLSQR (iterative least square) fitting algorithm to remove major streak artifacts in the occipital lobe, which we encountered at the early stage in data preprocessing. However, fibre-orientation-induced susceptibility error remains to be corrected in QSM measures derived from single-orientation acquisition (Schweser et al., 2016). In addition, though R2* seems straightforward to compute, the final result may not accurately reflect the underlying biology. MR relaxometry fitting has long considered the monoexponentially decaying model the gold standard. However, a study of T2*

mapping shows substantial nonlinear deviations between the signal and model (van Gelderen et al., 2012), and the same concept of fibre orientation mentioned above can be extended toward R2* mapping (Lee, Shin, et al., 2017; van Gelderen et al., 2012). Another T2 mapping simulation study shows that the signal noise level, number of echoes, echo time and repetition time all play crucial roles in the estimation (Yang, 2024). Nevertheless, the choices of imaging sequence and post-processing pipeline for the current study (Uddin et al., 2022) were all based on widely used, previously established and validated methods, and the choice to use large ROI-based analyses should mitigate potential effects of localized image artifacts, imperfections in spatial normalization, etc.

Decomposition of the Spearman rank correlation matrix

In this study, the Spearman rank correlation matrix is broken down into smaller components to organize the discussion. The rationale for the breakdown is systematic and forms a complete basis of the master matrix (i.e., mutual exclusivity). However, it was only "an" attempt to turn the results and empirical findings into something more easily digestible by the readers, with the realization that these perspectives are not exhausted by any means. Most importantly, individual significant relationships do not imply causality and should be regarded and interpreted with caution and known limitations.

2.6 Conclusion

This study has investigated the sensitivity and relationships of the ten qMRI metrics across three different ROIs in MS. The great sensitivity of the qMRI metrics across ROIs demonstrated by the Wilcoxon Signed-Rank tests, and especially the ROC/AUC result from the classification task

between WML and NAWM, highlighted their early clinical potentials in MS. And although Spearman rank correlation revealed a wide range of relationships between qMRI metrics both within and across ROIs, the information shared among qMRI metrics are not completely redundant, suggesting that using a multimodal approach, and specifically gathering unique metrics like QSM, may yield a much richer representation of the underlying MS pathophysiology to help better characterize the disease. However, the nontrivial interpretation of qMRI maps, specifically with a complex disease like MS, poses an issue that significantly hinders the clinical applicability, underpinning the need for more work to advance our understanding of qMRI in MS to improve the translatability.

2.7 References

Agrawal, R., & Sharma, M. (2014). Review of Segmentation Methods for Brain Tissue with Magnetic Resonance Images. *International Journal of Computer Network and Information Security*, *6*(4), 55–62. https://doi.org/10.5815/ijcnis.2014.04.07

Alonso-Ortiz, E., Levesque, I. R., & Pike, G. B. (2015). MRI-based myelin water imaging: A technical review. *Magnetic Resonance in Medicine*, *73*(1), 70–81. https://doi.org/10.1002/mrm.25198

Arshad, M., Stanley, J. A., & Raz, N. (2017). Test–retest reliability and concurrent validity of in vivo myelin content indices: Myelin water fraction and calibrated T_1 w/T_2 w image ratio. *Human Brain Mapping*, *38*(4), 1780–1790. https://doi.org/10.1002/hbm.23481

Bagnato, F., Hametner, S., Boyd, E., Endmayr, V., Shi, Y., Ikonomidou, V., Chen, G., Pawate, S., Lassmann, H., Smith, S., & Welch, E. B. (2018). Untangling the R2* contrast in

multiple sclerosis: A combined MRI-histology study at 7.0 Tesla. *PLOS ONE*, *13*(3), e0193839. https://doi.org/10.1371/journal.pone.0193839

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR in Biomedicine*, *15*(7–8), 435–455. https://doi.org/10.1002/nbm.782

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Birkl, C., Birkl-Toeglhofer, A. M., Endmayr, V., Höftberger, R., Kasprian, G., Krebs, C., Haybaeck, J., & Rauscher, A. (2019). The influence of brain iron on myelin water imaging. *NeuroImage*, *199*, 545–552. https://doi.org/10.1016/j.neuroimage.2019.05.042

Blazejewska, A. I., Al-Radaideh, A. M., Wharton, S., Lim, S. Y., Bowtell, R. W., Constantinescu, C. S., & Gowland, P. A. (2015). Increase in the iron content of the substantia nigra and red nucleus in multiple sclerosis and clinically isolated syndrome: A 7 Tesla MRI study. *Journal of Magnetic Resonance Imaging*, *41*(4), 1065–1070. https://doi.org/10.1002/jmri.24644

Boaventura, M., Sastre-Garriga, J., Garcia-Vidal, A., Vidal-Jordana, A., Quartana, D., Carvajal, R., Auger, C., Alberich, M., Tintoré, M., Rovira, À., Montalban, X., & Pareto, D. (2022). T1/T2-weighted ratio in multiple sclerosis: A longitudinal study with clinical associations. *NeuroImage: Clinical*, *34*, 102967. https://doi.org/10.1016/j.nicl.2022.102967

Chan, K.-S., Chamberland, M., & Marques, J. P. (2023). On the performance of multi-compartment relaxometry for myelin water imaging (MCR-MWI) – test-retest repeatability and inter-protocol reproducibility. *NeuroImage*, *266*, 119824. https://doi.org/10.1016/j.neuroimage.2022.119824

Compston, A., & Coles, A. (2002). Multiple sclerosis. *The Lancet*, *359*(9313), 1221–1231. https://doi.org/10.1016/S0140-6736(02)08220-X

Cooper, G., Chien, C., Zimmermann, H., Bellmann-Strobl, J., Ruprecht, K., Kuchling, J., Asseyer, S., Brandt, A. U., Scheel, M., Finke, C., & Paul, F. (2021). Longitudinal analysis of T1w/T2w ratio in patients with multiple sclerosis from first clinical presentation. *Multiple Sclerosis Journal*, *27*(14), 2180–2190. https://doi.org/10.1177/13524585211003479

Deoni, S. C. L., Rutt, B. K., Arun, T., Pierpaoli, C., & Jones, D. K. (2008). Gleaning multicomponent T_1 and T_2 information from steady-state imaging data. *Magnetic Resonance in Medicine*, *60*(6), 1372–1387. https://doi.org/10.1002/mrm.21704

Edwards, E. M., Wu, W., & Fritz, N. E. (2022). Using myelin water imaging to link underlying pathology to clinical function in multiple sclerosis: A scoping review. *Multiple Sclerosis and Related Disorders*, *59*, 103646. https://doi.org/10.1016/j.msard.2022.103646

Filippi, M., Cercignani, M., Inglese, M., Horsfield, M. A., & Comi, G. (2001). Diffusion tensor magnetic resonance imaging in multiple sclerosis. *Neurology*, *56*(3), 304–311. https://doi.org/10.1212/WNL.56.3.304

Filippi, M., Pagani, E., Preziosa, P., & Rocca, M. A. (2016). The Role of DTI in Multiple Sclerosis and Other Demyelinating Conditions. In *Diffusion Tensor Imaging* (pp. 331–341). Springer New York. https://doi.org/10.1007/978-1-4939-3118-7_16

Filippi, M., Preziosa, P., Banwell, B. L., Barkhof, F., Ciccarelli, O., De Stefano, N., Geurts, J. J. G., Paul, F., Reich, D. S., Toosy, A. T., Traboulsee, A., Wattjes, M. P., Yousry, T. A., Gass, A., Lubetzki, C., Weinshenker, B. G., & Rocca, M. A. (2019). Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. In *Brain* (Vol. 142, Number 7, pp. 1858–1875). Oxford University Press. https://doi.org/10.1093/brain/awz144

Fuchs, T. A., Carolus, K., Benedict, R. H. B., Bergsland, N., Ramasamy, D., Jakimovski, D., Weinstock-Guttman, B., Kuceyeski, A., Zivadinov, R., & Dwyer, M. G. (2018). Impact of Focal White Matter Damage on Localized Subcortical Gray Matter Atrophy in Multiple Sclerosis: A 5-Year Study. *American Journal of Neuroradiology*, *39*(8), 1480–1486. https://doi.org/10.3174/ajnr.A5720

Ganzetti, M., Wenderoth, N., & Mantini, D. (2014). Whole brain myelin mapping using T1- and T2-weighted MR imaging data. *Frontiers in Human Neuroscience*, *8*(SEP). https://doi.org/10.3389/fnhum.2014.00671

Glasser, M. F., Coalson, T. S., Harms, M. P., Xu, J., Baum, G. L., Autio, J. A., Auerbach, E. J., Greve, D. N., Yacoub, E., Van Essen, D. C., Bock, N. A., & Hayashi, T. (2022). Empirical transmit field bias correction of T1w/T2w myelin maps. *NeuroImage*, *258*, 119360. https://doi.org/10.1016/j.neuroimage.2022.119360

Glasser, M. F., & van Essen, D. C. (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *Journal of Neuroscience*, *31*(32), 11597–11616. https://doi.org/10.1523/JNEUROSCI.2180-11.2011

Granziera, C., Wuerfel, J., Barkhof, F., Calabrese, M., De Stefano, N., Enzinger, C., Evangelou, N., Filippi, M., Geurts, J. J. G., Reich, D. S., Rocca, M. A., Ropele, S., Rovira, À., Sati, P., Toosy, A. T., Vrenken, H., Gandini Wheeler-Kingshott, C. A. M., Kappos, L., Barkhof, F., … Yousry, T. (2021). Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. *Brain*, *144*(5), 1296–1311. https://doi.org/10.1093/brain/awab029

Gulani, V., & Seiberlich, N. (2020). Quantitative MRI: Rationale and Challenges. In *Advances in magnetic resonance technology and applications* (Vol. 1, pp. xxxvii–li). Elsevier. https://doi.org/10.1016/B978-0-12-817057-1.00001-9

Hagemeier, J., Zivadinov, R., Dwyer, M. G., Polak, P., Bergsland, N., Weinstock-Guttman, B., Zalis, J., Deistung, A., Reichenbach, J. R., & Schweser, F. (2018). Changes of deep gray matter magnetic susceptibility over 2 years in multiple sclerosis and healthy control brain. *NeuroImage: Clinical*, *18*, 1007–1016. https://doi.org/10.1016/j.nicl.2017.04.008

Haider, L., Simeonidou, C., Steinberger, G., Hametner, S., Grigoriadis, N., Deretzi, G., Kovacs, G. G., Kutzelnigg, A., Lassmann, H., & Frischer, J. M. (2014). Multiple sclerosis deep grey matter: the relation between demyelination, neurodegeneration, inflammation and iron. *Journal of Neurology, Neurosurgery & Psychiatry*, *85*(12), 1386–1395. https://doi.org/10.1136/jnnp-2014-307712

Hannoun, S., Kocevar, G., Codjia, P., Barile, B., Cotton, F., Durand-Dubief, F., & Sappey-Marinier, D. (2022). T1/T2 ratio: A quantitative sensitive marker of brain tissue integrity in multiple sclerosis. *Journal of Neuroimaging*, *32*(2), 328–336. https://doi.org/10.1111/jon.12943

Heiervang, E., Behrens, T. E. J., Mackay, C. E., Robson, M. D., & Johansen-Berg, H. (2006). Between session reproducibility and between subject variability of diffusion MR and tractography measures. *NeuroImage*, *33*(3), 867–877. https://doi.org/10.1016/j.neuroimage.2006.07.037

Kolind, S., Matthews, L., Johansen-Berg, H., Leite, M. I., Williams, S. C. R., Deoni, S., & Palace, J. (2012). Myelin water imaging reflects clinical variability in multiple sclerosis. *NeuroImage*, *60*(1), 263–270. https://doi.org/10.1016/j.neuroimage.2011.11.070

Kor, D., Birkl, C., Ropele, S., Doucette, J., Xu, T., Wiggermann, V., Hernández-Torres, E., Hametner, S., & Rauscher, A. (2019). The role of iron and myelin in orientation dependent R_2^* of white matter. *NMR in Biomedicine*, e4092. https://doi.org/10.1002/nbm.4092

Kuceyeski, A. F., Vargas, W., Dayan, M., Monohan, E., Blackwell, C., Raj, A., Fujimoto, K., & Gauthier, S. A. (2015). Modeling the Relationship among Gray Matter Atrophy, Abnormalities in Connecting White Matter, and Cognitive Performance in Early Multiple Sclerosis. *American Journal of Neuroradiology*, *36*(4), 702–709. https://doi.org/10.3174/ajnr.A4165

Langkammer, C., Schweser, F., Krebs, N., Deistung, A., Goessler, W., Scheurer, E., Sommer, K., Reishofer, G., Yen, K., Fazekas, F., Ropele, S., & Reichenbach, J. R. (2012). Quantitative susceptibility mapping (QSM) as a means to measure brain iron? A post mortem validation study. *NeuroImage*, *62*(3), 1593–1599. https://doi.org/10.1016/j.neuroimage.2012.05.049

Laule, C., Kozlowski, P., Leung, E., Li, D. K. B., MacKay, A. L., & Moore, G. R. W. (2008). Myelin water imaging of multiple sclerosis at 7 T: Correlations with histopathology. *NeuroImage*, *40*(4), 1575–1580. https://doi.org/10.1016/j.neuroimage.2007.12.008

Laule, C., Leung, E., Li, D. K. B., Traboulsee, A. L., Paty, D. W., MacKay, A. L., & Moore, G. R. W. (2006). Myelin water imaging in multiple sclerosis: quantitative correlations with histopathology. *Multiple Sclerosis Journal*, *12*(6), 747–753. https://doi.org/10.1177/1352458506070928

Laule, C., & Moore, G. R. W. (2018). Myelin water imaging to detect demyelination and remyelination and its validation in pathology. *Brain Pathology*, *28*(5), 750–764. https://doi.org/10.1111/bpa.12645

Lavrova, E., Lommers, E., Woodruff, H. C., Chatterjee, A., Maquet, P., Salmon, E., Lambin, P., & Phillips, C. (2021). Exploratory Radiomic Analysis of Conventional vs. Quantitative Brain MRI: Toward Automatic Diagnosis of Early Multiple Sclerosis. *Frontiers in Neuroscience*, *15*. https://doi.org/10.3389/fnins.2021.679941

le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., & Chabriat, H. (2001). Diffusion tensor imaging: Concepts and applications. *Journal of Magnetic Resonance Imaging*, *13*(4), 534–546. https://doi.org/10.1002/jmri.1076

Lee, J., Hyun, J., Lee, J., Choi, E., Shin, H., Min, K., Nam, Y., Kim, H. J., & Oh, S. (2021). So You Want to Image Myelin Using MRI: An Overview and Practical Guide for Myelin Water Imaging. *Journal of Magnetic Resonance Imaging*, *53*(2), 360–373. https://doi.org/10.1002/jmri.27059

Lee, J., Nam, Y., Choi, J. Y., Kim, E. Y., Oh, S., & Kim, D. (2017). Mechanisms of T_2* anisotropy and gradient echo myelin water imaging. *NMR in Biomedicine*, *30*(4). https://doi.org/10.1002/nbm.3513

Lee, J., Shin, H.-G., Jung, W., Nam, Y., Oh, S.-H., & Lee, J. (2017). An R2* model of white matter for fiber orientation and myelin concentration. *NeuroImage*, *162*, 269–275. https://doi.org/10.1016/j.neuroimage.2017.08.050

Lim, I. A. L., Faria, A. V., Li, X., Hsu, J. T. C., Airan, R. D., Mori, S., & van Zijl, P. C. M. (2013). Human brain atlas for automated region of interest selection in quantitative susceptibility mapping: Application to determine iron content in deep gray matter structures. *NeuroImage*, *82*, 449–469. https://doi.org/10.1016/j.neuroimage.2013.05.127

Lipp, I., Jones, D. K., Bells, S., Sgarlata, E., Foster, C., Stickland, R., Davidson, A. E., Tallantyre, E. C., Robertson, N. P., Wise, R. G., & Tomassini, V. (2019). Comparing MRI metrics to quantify white matter microstructural damage in multiple sclerosis. *Human Brain Mapping*, hbm.24568. https://doi.org/10.1002/hbm.24568

Lucchinetti, C., Brück, W., Parisi, J., Scheithauer, B., Rodriguez, M., & Lassmann, H. (2000). Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of

demyelination. *Annals of Neurology*, *47*(6), 707–717. https://doi.org/10.1002/1531-8249(200006)47:6<707::AID-ANA3>3.0.CO;2-Q

MacKay, A. L., & Laule, C. (2016). Magnetic Resonance of Myelin Water: An in vivo Marker for Myelin. *Brain Plasticity*, *2*(1), 71–91. https://doi.org/10.3233/bpl-160033

MacKay, A. L., Vavasour, I. M., Rauscher, A., Kolind, S. H., Mädler, B., Moore, G. R. W., Traboulsee, A. L., Li, D. K. B., & Laule, C. (2009). MR Relaxation in Multiple Sclerosis. *Neuroimaging Clinics of North America*, *19*(1), 1–26. https://doi.org/10.1016/j.nic.2008.09.007

Maillard, P., Fletcher, E., Harvey, D., Carmichael, O., Reed, B., Mungas, D., & DeCarli, C. (2011). White Matter Hyperintensity Penumbra. *Stroke*, *42*(7), 1917–1922. https://doi.org/10.1161/STROKEAHA.110.609768

Massey, F. J. (1951). The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, *46*(253), 68. https://doi.org/10.2307/2280095

McDonald, W. I., Compston, A., Edan, G., Goodkin, D., Hartung, H., Lublin, F. D., McFarland, H. F., Paty, D. W., Polman, C. H., Reingold, S. C., Sandberg-Wollheim, M., Sibley, W., Thompson, A., Van Den Noort, S., Weinshenker, B. Y., & Wolinsky, J. S. (2001). Recommended diagnostic criteria for multiple sclerosis: Guidelines from the international panel on the diagnosis of multiple sclerosis. *Annals of Neurology*, *50*(1), 121–127. https://doi.org/10.1002/ana.1032

Mercadante, A. A., & Tadi, P. (2026). Neuroanatomy, Gray Matter. In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/pubmed/31785402

Meyers, S. M., Vavasour, I. M., Mädler, B., Harris, T., Fu, E., Li, D. K. B., Traboulsee, A. L., MacKay, A. L., & Laule, C. (2013). Multicenter measurements of myelin water fraction and

geometric mean T_2 : Intra- and intersite reproducibility. *Journal of Magnetic Resonance Imaging*, *38*(6), 1445–1453. https://doi.org/10.1002/jmri.24106

Montalban, X., Lebrun-Frénay, C., Oh, J., Arrambide, G., Moccia, M., Pia Amato, M., Amezcua, L., Banwell, B., Bar-Or, A., Barkhof, F., Butzkueven, H., Ciccarelli, O., Chataway, J., Cohen, J. A., Comi, G., Correale, J., Deisenhammer, F., Filippi, M., Fiol, J., … Thompson, A. J. (2025). Diagnosis of multiple sclerosis: 2024 revisions of the McDonald criteria. *The Lancet Neurology*, *24*(10), 850–865. https://doi.org/10.1016/S1474-4422(25)00270-4

Nerland, S., Jørgensen, K. N., Nordhøy, W., Maximov, I. I., Bugge, R. A. B., Westlye, L. T., Andreassen, O. A., Geier, O. M., & Agartz, I. (2021). Multisite reproducibility and test-retest reliability of the T1w/T2w-ratio: A comparison of processing methods. *NeuroImage*, *245*, 118709. https://doi.org/10.1016/j.neuroimage.2021.118709

O'Donnell, L. J., & Westin, C.-F. (2011). An Introduction to Diffusion Tensor Image Analysis. *Neurosurgery Clinics of North America*, *22*(2), 185–196. https://doi.org/10.1016/j.nec.2010.12.004

Oh, S.-H., Bilello, M., Schindler, M., Markowitz, C. E., Detre, J. A., & Lee, J. (2013). Direct visualization of short transverse relaxation time component (ViSTa). *NeuroImage*, *83*, 485–492. https://doi.org/10.1016/j.neuroimage.2013.06.047

Ontaneda, D., Raza, P. C., Mahajan, K. R., Arnold, D. L., Dwyer, M. G., Gauthier, S. A., Greve, D. N., Harrison, D. M., Henry, R. G., Li, D. K. B., Mainero, C., Moore, W., Narayanan, S., Oh, J., Patel, R., Pelletier, D., Rauscher, A., Rooney, W. D., Sicotte, N. L., … Azevedo, C. J. (2021). Deep grey matter injury in multiple sclerosis: a NAIMS consensus statement. *Brain*, *144*(7), 1974–1984. https://doi.org/10.1093/brain/awab132

Prasloski, T., Rauscher, A., MacKay, A. L., Hodgson, M., Vavasour, I. M., Laule, C., & Mädler, B. (2012). Rapid whole cerebrum myelin water imaging using a 3D GRASE sequence. *NeuroImage*, *63*(1), 533–539. https://doi.org/10.1016/j.neuroimage.2012.06.064

Righart, R., Biberacher, V., Jonkman, L. E., Klaver, R., Schmidt, P., Buck, D., Berthele, A., Kirschke, J. S., Zimmer, C., Hemmer, B., Geurts, J. J. G., & Mühlau, M. (2017). Cortical pathology in multiple sclerosis detected by the T1/T2-weighted ratio from routine magnetic resonance imaging. *Annals of Neurology*, *82*(4), 519–529. https://doi.org/10.1002/ana.25020

Rocca, M. A., Preziosa, P., Barkhof, F., Brownlee, W., Calabrese, M., De Stefano, N., Granziera, C., Ropele, S., Toosy, A. T., Vidal-Jordana, À., Di Filippo, M., & Filippi, M. (2024). Current and future role of MRI in the diagnosis and prognosis of multiple sclerosis. *The Lancet Regional Health - Europe*, *44*, 100978. https://doi.org/10.1016/j.lanepe.2024.100978

Rovaris, M., Gass, A., Bammer, R., Hickman, S. J., Ciccarelli, O., Miller, D. H., & Filippi, M. (2005). Diffusion MRI in multiple sclerosis. *Neurology*, *65*(10), 1526–1532. https://doi.org/10.1212/01.wnl.0000184471.83948.e0

Saltarelli, G., Di Cerbo, G., Innocenzi, A., De Felici, C., Splendiani, A., & Di Cesare, E. (2025). Quantitative MRI in Neuroimaging: A Review of Techniques, Biomarkers, and Emerging Clinical Applications. *Brain Sciences*, *15*(10), 1088. https://doi.org/10.3390/brainsci15101088

Sandrone, S., Aiello, M., Cavaliere, C., Thiebaut de Schotten, M., Reimann, K., Troakes, C., Bodi, I., Lacerda, L., Monti, S., Murphy, D., Geyer, S., Catani, M., & Dell'Acqua, F. (2023). Mapping myelin in white matter with T1-weighted/T2-weighted maps: discrepancy with histology and other myelin MRI measures. *Brain Structure and Function*, *228*(2), 525–535. https://doi.org/10.1007/s00429-022-02600-z

Sbardella, E., Tona, F., Petsas, N., & Pantano, P. (2013). DTI Measurements in Multiple Sclerosis: Evaluation of Brain Damage and Clinical Implications. *Multiple Sclerosis International*, *2013*, 1–11. https://doi.org/10.1155/2013/671730

Scalco, E., Belfatto, A., Mastropietro, A., Rancati, T., Avuzzi, B., Messina, A., Valdagni, R., & Rizzo, G. (2020). T2w-MRI signal normalization affects radiomics features reproducibility. *Medical Physics*, *47*(4), 1680–1691. https://doi.org/10.1002/mp.14038

Schweser, F., Deistung, A., & Reichenbach, J. R. (2016). Foundations of MRI phase imaging and processing for Quantitative Susceptibility Mapping (QSM). *Zeitschrift Fur Medizinische Physik*, *26*(1), 6–34. https://doi.org/10.1016/j.zemedi.2015.10.002

Stankiewicz, J. M., Neema, M., & Ceccarelli, A. (2014). Iron and multiple sclerosis. In *Neurobiology of Aging* (Vol. 35, Number SUPPL.2). Elsevier Inc. https://doi.org/10.1016/j.neurobiolaging.2014.03.039

Stüber, C., Pitt, D., & Wang, Y. (2016). Iron in multiple sclerosis and its noninvasive imaging with quantitative susceptibility mapping. In *International Journal of Molecular Sciences* (Vol. 17, Number 1). MDPI AG. https://doi.org/10.3390/ijms17010100

Syková, E., & Nicholson, C. (2008). Diffusion in Brain Extracellular Space. *Physiological Reviews*, *88*(4), 1277–1340. https://doi.org/10.1152/physrev.00027.2007

Tadayon, E., Khayati, R. M., Karami, V., & Nabavi, S. M. (2016). A novel method for automatic classification of multiple sclerosis lesion subtypes using diffusion tensor mr images. *Biomedical Engineering: Applications, Basis and Communications*, *28*(05), 1650038. https://doi.org/10.4015/S1016237216500381

Tait, M. J., Saadoun, S., Bell, B. A., & Papadopoulos, M. C. (2008). Water movements in the brain: role of aquaporins. *Trends in Neurosciences*, *31*(1), 37–43. https://doi.org/10.1016/j.tins.2007.11.003

Takao, H., Hayashi, N., & Ohtomo, K. (2011). Effect of scanner in asymmetry studies using diffusion tensor imaging. *NeuroImage*, *54*(2), 1053–1062. https://doi.org/10.1016/j.neuroimage.2010.09.023

Tranfa, M., Pontillo, G., Petracca, M., Brunetti, A., Tedeschi, E., Palma, G., & Cocozza, S. (2022). Quantitative MRI in Multiple Sclerosis: From Theory to Application. *American Journal of Neuroradiology*, *43*(12), 1688–1695. https://doi.org/10.3174/ajnr.A7536

Uddin, M. N., Figley, T. D., & Figley, C. R. (2018). Effect of echo time and T2-weighting on GRASE-based T1w/T2w ratio measurements at 3T. *Magnetic Resonance Imaging*, *51*, 35–43. https://doi.org/10.1016/j.mri.2018.04.012

Uddin, M. N., Figley, T. D., Kornelsen, J., Mazerolle, E. L., Helmick, C. A., O'Grady, C. B., Pirzada, S., Patel, R., Carter, S., Wong, K., Essig, M. R., Graff, L. A., Bolton, J. M., Marriott, J. J., Bernstein, C. N., Fisk, J. D., Marrie, R. A., & Figley, C. R. (2022). The comorbidity and cognition in multiple sclerosis (CCOMS) neuroimaging protocol: Study rationale, MRI acquisition, and minimal image processing pipelines. *Frontiers in Neuroimaging*, *1*. https://doi.org/10.3389/fnimg.2022.970385

Uddin, M. N., Figley, T. D., Marrie, R. A., & Figley, C. R. (2018). Can T_1 w/ T_2 w ratio be used as a myelin-specific measure in subcortical structures? Comparisons between FSE-based T_1 w/ T_2 w ratios, GRASE-based T_1 w/ T_2 w ratios and multi-echo GRASE-based myelin water fractions. *NMR in Biomedicine*, *31*(3), e3868. https://doi.org/10.1002/nbm.3868

Uddin, M. N., Figley, T. D., Solar, K. G., Shatil, A. S., & Figley, C. R. (2019). Comparisons between multi-component myelin water fraction, T1w/T2w ratio, and diffusion tensor imaging measures in healthy human brain structures. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-39199-x

Uddin, M. N., Lebel, R. M., & Wilman, A. H. (2016). Value of transverse relaxometry difference methods for iron in human brain. *Magnetic Resonance Imaging*, *34*(1), 51–59. https://doi.org/10.1016/j.mri.2015.09.002

van Gelderen, P., de Zwart, J. A., Lee, J., Sati, P., Reich, D. S., & Duyn, J. H. (2012). Nonexponential T_2* decay in white matter. *Magnetic Resonance in Medicine*, *67*(1), 110–117. https://doi.org/10.1002/mrm.22990

Vargas, W. S., Monohan, E., Pandya, S., Raj, A., Vartanian, T., Nguyen, T. D., Hurtado Rúa, S. M., & Gauthier, S. A. (2015). Measuring longitudinal myelin water fraction in new multiple sclerosis lesions. *NeuroImage: Clinical*, *9*, 369–375. https://doi.org/10.1016/j.nicl.2015.09.003

Vavasour, I. M., Laule, C., Li, D. K. B., Oger, J., Moore, G. R. W., Traboulsee, A., & MacKay, A. L. (2009). Longitudinal changes in myelin water fraction in two MS patients with active disease. *Journal of the Neurological Sciences*, *276*(1–2), 49–53. https://doi.org/10.1016/j.jns.2008.08.022

Vercellino, M., Masera, S., Lorenzatti, M., Condello, C., Merola, A., Mattioda, A., Tribolo, A., Capello, E., Mancardi, G. L., Mutani, R., Giordana, M. T., & Cavalla, P. (2009). Demyelination, Inflammation, and Neurodegeneration in Multiple Sclerosis Deep Gray Matter. *Journal of Neuropathology & Experimental Neurology*, *68*(5), 489–502. https://doi.org/10.1097/NEN.0b013e3181a19a5a

Wang, Y., & Liu, T. (2015). Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. *Magnetic Resonance in Medicine*, *73*(1), 82–101. https://doi.org/10.1002/mrm.25358

Weingärtner, S., Desmond, K. L., Obuchowski, N. A., Baessler, B., Zhang, Y., Biondetti, E., Ma, D., Golay, X., Boss, M. A., Gunter, J. L., Keenan, K. E., & Hernando, D. (2022). Development, validation, qualification, and dissemination of quantitative MR methods: Overview and recommendations by the ISMRM quantitative MR study group. *Magnetic Resonance in Medicine*, *87*(3), 1184–1206. https://doi.org/10.1002/mrm.29084

Werring, D. J., Clark, C. A., Barker, G. J., Thompson, A. J., & Miller, D. H. (1999). Diffusion tensor imaging of lesions and normal-appearing white matter in multiple sclerosis. *Neurology*, *52*(8), 1626–1626. https://doi.org/10.1212/WNL.52.8.1626

Whittall, K. P., Mackay, A. L., Graeb, D. A., Nugent, R. A., Li, D. K. B., & Paty, D. W. (1997). In vivo measurement ofT2 distributions and water contents in normal human brain. *Magnetic Resonance in Medicine*, *37*(1), 34–43. https://doi.org/10.1002/mrm.1910370107

Williams, R., Buchheit, C. L., Berman, N. E. J., & Levine, S. M. (2012). Pathogenic implications of iron accumulation in multiple sclerosis. In *Journal of Neurochemistry* (Vol. 120, Number 1, pp. 7–25). https://doi.org/10.1111/j.1471-4159.2011.07536.x

Wood, J. C., Enriquez, C., Ghugre, N., Tyzka, J. M., Carson, S., Nelson, M. D., & Coates, T. D. (2005). MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. *Blood*, *106*(4), 1460–1465. https://doi.org/10.1182/blood-2004-10-3982

Yaghmaie, N., Syeda, W. T., Wu, C., Zhang, Y., Zhang, T. D., Burrows, E. L., Brodtmann, A., Moffat, B. A., Wright, D. K., Glarin, R., Kolbe, S., & Johnston, L. A. (2021). QSMART:

Quantitative susceptibility mapping artifact reduction technique. *NeuroImage*, *231*, 117701. https://doi.org/10.1016/j.neuroimage.2020.117701

Yang, X. (2024). Improving the reliability of T_2 measurement in magnetic resonance imaging. *BioRxiv*, 2024.06.09.598128. https://doi.org/10.1101/2024.06.09.598128

Young, P. N. E., Estarellas, M., Coomans, E., Srikrishna, M., Beaumont, H., Maass, A., Venkataraman, A. v., Lissaman, R., Jiménez, D., Betts, M. J., McGlinchey, E., Berron, D., O'Connor, A., Fox, N. C., Pereira, J. B., Jagust, W., Carter, S. F., Paterson, R. W., & Schöll, M. (2020). Imaging biomarkers in neurodegeneration: current and future practices. *Alzheimer's Research & Therapy*, *12*(1), 49. https://doi.org/10.1186/s13195-020-00612-7

Zhang, Y., Gauthier, S. A., Gupta, A., Chen, W., Comunale, J., Chiang, G. C.-Y., Zhou, D., Askin, G., Zhu, W., Pitt, D., & Wang, Y. (2016). Quantitative Susceptibility Mapping and R2* Measured Changes during White Matter Lesion Development in Multiple Sclerosis: Myelin Breakdown, Myelin Debris Degradation and Removal, and Iron Accumulation. *American Journal of Neuroradiology*, *37*(9), 1629–1635. https://doi.org/10.3174/ajnr.A4825

Zhong, G., & Lou, M. (2016). Multimodal imaging findings in normal-appearing white matter of leucoaraiosis: a review. *BMJ*, *1*(2), 59–63. https://doi.org/10.1136/svn-2016-000021

Zhou, L., Li, Y., Sweeney, E. M., Wang, X. H., Kuceyeski, A., Chiang, G. C., Ivanidze, J., Wang, Y., Gauthier, S. A., de Leon, M. J., & Nguyen, T. D. (2023). Association of brain tissue cerebrospinal fluid fraction with age in healthy cognitively normal adults. *Frontiers in Aging Neuroscience*, *15*. https://doi.org/10.3389/fnagi.2023.1162001

Zhou, L., Li, Y., Wang, X., Sweeney, E., Zhang, H., Tanzi, E. B., Prince, J., Su-Ortiz, V. A., Gauthier, S. A., & Nguyen, T. D. (2021). *Increasing age is independently associated with higher*

free water in non-active MS brain - A multi-compartment analysis using FAST-T2. https://doi.org/10.1101/2021.02.01.429067

Zierfuss, B., Wang, Z., Jackson, A. N., Moezzi, D., & Yong, V. (2023). Iron in multiple sclerosis – Neuropathology, immunology, and real-world considerations. *Multiple Sclerosis and Related Disorders*, *78*, 104934. https://doi.org/10.1016/j.msard.2023.104934

Chapter 3

Volumetric MRI measures correlate with quantitative MRI measures in multiple sclerosis between white matter lesions, normal appearing white matter and deep gray matter regions

Publication status

This chapter (and Chapter 2) are part of the abstract entitled '*Correlation Between Microstructural and Volumetric MRI Measures in Multiple Sclerosis White Matter Lesions, Normal Appearing White Matter, and Deep Gray Matter Regions'* accepted for presentation at the **Society for Neuroscience** conference in 2022. The manuscript is currently under internal review before submission to a peer-review journal.

3.1 Abstract

Despite the fact that brain volumetry in multiple sclerosis (MS) exhibits clinical value in assessing disease progression, the relation between the regional volume of MS brain tissues and the microstructural integrity remains underexplored. With recent advances, quantitative MRI (qMRI) has demonstrated the ability to characterize the underlying microstructural status of brain tissues, offering promising opportunities for clinical application in MS. In this cross-sectional study, we aim to examine the correlations between 10 qMRI metrics (axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD), fractional anisotropy (FA), myelin water fraction (MWF), intra- and extra- cellular water fraction (IEWF), geometric mean T2 (gT2), T1-weighted-by-T2-weighted (T1w/T2w) ratio, R2* and QSM)) and regional brain volumetric (Vol) measures. These relationships were evaluated across three tissue regions of interest (NAWM, WML and DGM). With a sample size of 79, Spearman rank correlation analyses were performed, and statistical

significance was globally adjusted for multiple comparisons using a false discovery rate correction. A total of 235 (out of 528 tested) correlations remained significant (i.e., $p_{BH\text{-}FDR} < 0.05$). Among these, 56 of 93 correlations related to volumetric measure were significant. Volumetric measures showed the highest number of significant correlations among the qMRI metrics. Specifically, WML Vol exhibited a positive association with worsening microstructural status, as measured by qMRI metrics. This finding highlights that qMRI can support the cross-sectional uses of volumetry for MS in the clinic and even potentially provide more information about the underlying degree of microstructural damage to better characterize MS.

3.2 Introduction

Multiple sclerosis (MS) is an autoimmune neurodegenerative disease widely recognized by its pathological feature of demyelination (Korn, 2008). Demyelination is often observed along with axonal degeneration (Bitsch, 2000; Bjartmar et al., 2003; Trapp et al., 1998). In addition, the dyregulated ferrous iron leads to the formation of free radicals, which increase oxidative stress and promote further neuronal damage (Haider et al., 2014; Stankiewicz et al., 2014; Stephenson et al., 2014; Williams et al., 2012; Zierfuss et al., 2023). Given that myelin and white matter account for a large portion of the brain (Morell & Norton, 1980; Sampaio-Baptista & Johansen-Berg, 2017), brain volume changes (especially brain atrophy) in MS have long been investigated (Andravizou et al., 2019; Fox et al., 2000; Giorgio & De Stefano, 2013; Jagust & Noseworthy, 2000; Zivadinov & Bakshi, 2004). Recent volumetric studies have even found significant relationships between whole brain and regional brain atrophy, and disease progression and clinical scores/assessments (Barnett et al., 2020; Preziosa et al., 2022; Radue et al., 2015). Notably, brain volumetric measures in these studies all rely on magnetic resonance imaging.

The volumetric measure (Vol), unlike other quantitative metrics, is neither an innovative metric nor does it require advanced sequences or carefully designed post-processing pipelines. Indeed, with studies investigating the test-retest reliability of MRI-based brain volumetry (Bartzokis et al., 1993; Luft et al., 1996; Maclaren et al., 2014) and accessible software to automate the volumetry calculation process (Harkey et al., 2022; J. Y. Lee et al., 2021; Tavares et al., 2020), MRI-based brain volumetry has become common. In MS, autoimmune responses can take place in both white matter (WM) and gray matter (GM) in the central nervous system (CNS), leading to cellular-level dysregulation and axon degeneration (Andravizou et al., 2019), therefore causing brain volume changes either in short-term fluctuations (Rovaris, 2001; Uher et al., 2021) or long-

term reductions (Chard, 2003). Medication can also play a modulating affect in brain volume changes (Andravizou et al., 2019; Dieleman et al., 2017). To explore the relationship between different compartmental brain Vol in MS, a cross-sectional study with MS (N=75) and healthy control (N=21) analyzed 56 regions across the brain (including, but not limited to cortical and subcortical subregions, ventricle subregions, and WM and non-WM hypointensities) and found that many regional volumes (normalized by subject's head size) were significantly correlated to disease severity (i.e., EDSS, and that not all of them exhibited the same sensitivity among clusters (which has been qualitatively shown to be relevant to different clinical phenotypes) by the hierarchical clustering analysis, highlighting an early evidence of different degrees of regional atrophy and the unique sensitivity of the WM abnormalities in MS (Fujimori et al., 2020). Interestingly, a longitudinal study particularly looked at brain atrophy in different MS phenotypes, and the result revealed distinct spatiotemporal patterns of the brain atrophy, and they were phenotypically relevant (i.e., ventricular enlargement in relapsing remitting forms vs. cortical and subcortical atrophy in progressive forms), underpinning the heterogeneity in MS and the clinical value of regional brain volume (Pagani et al., 2005). Pairing with these observations, brain Vol has long been studied for its clinical feasibility. Expanded disability status scale (EDSS) has been shown to be associated with whole brain atrophy (Fisher et al., 2002; Rocca et al., 2017; Rudick et al., 2000), baseline gray matter Vol (Lavorgna et al., 2014), T2-lesion Vol (in combination with other brain Vol) (Popescu et al., 2013). Also, gray matter atrophy and T2-lesion Vol has also been found to be associated with depression, and long-term cognition impairment, respectively (Pravatà et al., 2017). And especially, high T2 lesion Vol is one of many indicators of a poor MS prognosis (Rotstein & Montalban, 2019). Nevertheless, despite its promising clinical applicability, the complexity of MS pathophysiology, as outlined, in contrast to the macroscopic nature of the Vol,

there is still a pressing need to further characterize brain Vol using more sensitive microstructural status to enhance its interpretability to enable precision in medicine.

Some specialized MRI sequences and associated processing pipelines are developed to produce quantitative MRI (qMRI) metrics. The scope of interest with qMRI metrics is much narrower than that of conventional MRI for examining specific properties of microstructural changes (Granziera et al., 2021). Diffusion tensor imaging (DTI) can probe diffusivity in brain tissues using the gradient dephasing/rephasing mechanism. It provides measures sensitive to compromised microstructural integrity related to diffusivity in MS cases (Sbardella et al., 2013). Myelin water imaging (MWI) estimates the entire T2 spectrum of a voxel using empirical cutoff values derived from healthy tissues. Subsequently, different compartmental water fractions (e.g., myelin water fraction) can be determined (MacKay & Laule, 2016), and these have been shown to provide high specificity to myelin via comparisons with histological staining (Laule et al., 2008; Sandrone et al., 2023). The T1-weighted-divided-by-T2-weighted (T1w/T2w) ratio is a straightforward method for probing microstructural integrity by normalizing T1-weighted images to T2-weighted images and enabling direct quantitative comparison of brain voxels (Glasser & van Essen, 2011), although its specificity to any particular biomarker (e.g., myelin concentration) has been debated, particularly within WM regions (Sandrone et al., 2023; Uddin et al., 2018, 2019). However, although not myelin-specific, a recent longitudinal study showed that the T1w/T2w ratio had high sensitivity across MS brain regions (Boaventura et al., 2022). Quantitative susceptibility mapping (QSM) utilizes phase accumulation caused by local field inhomogeneity due to susceptibility and maps the susceptibility for each voxel correspondingly (Schweser et al., 2016). Because iron has a strong positive susceptibility, QSM is a widely used imaging technique for mapping iron content in the brain, especially in the deep gray matter regions of MS (Langkammer

et al., 2013). R2* (or $1/T2^*$) mapping is another popular method for detecting elevated iron levels in the brain (Wood et al., 2005) and is beneficial for understanding the effects of iron in MS.

qMRI measures, contrary to clinical assessments, are more directly sensitive to biological changes in the brain. Due to the lack of literature investigating the relationship between different quantitative measures and MS brains' compartmental volume, this cross-sectional study aims to quantify relationships between volumetric measures and different qMRI measures in three major brain regions in MS: normal appearing white matter (NAWM), white matter lesions (WML) and deep gray matter (DGM) by conducting a statistical correlation analysis.

3.3 Methods

Dataset summary

This study was approved by the University of Manitoba Health Research Ethics Board and consent from the participants was collected before any scanning/analysis. The data set is a subset of the complete dataset, with a full battery of MRI sequences collected (Participants = 79, 67 are females), and it belongs to the follow-up visit of the Comorbidity and Cognition in Multiple Sclerosis (CCOMS) Study (Uddin et al., 2022). Participants' median expanded disability status scale is 3.5 (IQR: [3,5.5]), and mean age is 50.6 (range: 22.8 – 79.0). They were all previously diagnosed with multiple sclerosis, with conditions met under McDonald's criteria. For more details on the inclusion criteria, one may refer to our previously published imaging protocol (Uddin et al., 2022).

qMRI map pre-processing and ROI mask computation

This study aimed to investigate the correlation between volumetric (Vol) measures and ten qMRI metrics: axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD), fractional anisotropy (FA), myelin water fraction (MWF), intra- and extra- cellular water fraction (IEWF),

geometric mean T2 (gT2), T1w/T2w ratio, R2* and QSM. Figure 3-1 displays ten qMRI maps from a single axial slice for one representative participant. ROI analysis is employed with three brain regions as mentioned above: 1) NAWM, 2) WML, and 3) DGM. The ROI masks are illustrated in Figure 3-2. The imaging sequence specifications and image pre-processing pipeline details for generating each of the MNI-space qMRI maps can be found in a previously published protocol paper (Uddin et al., 2022). In addition, specific steps for ROI mask computation can be found in a parallel study from the same lab that uses the same sets ROI (Chapter 2).

Figure 3-1: Sample axial slices of 10 quantitative maps of one random MS participant of the study.

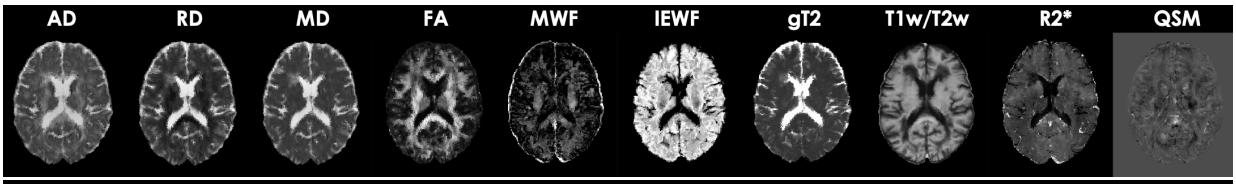
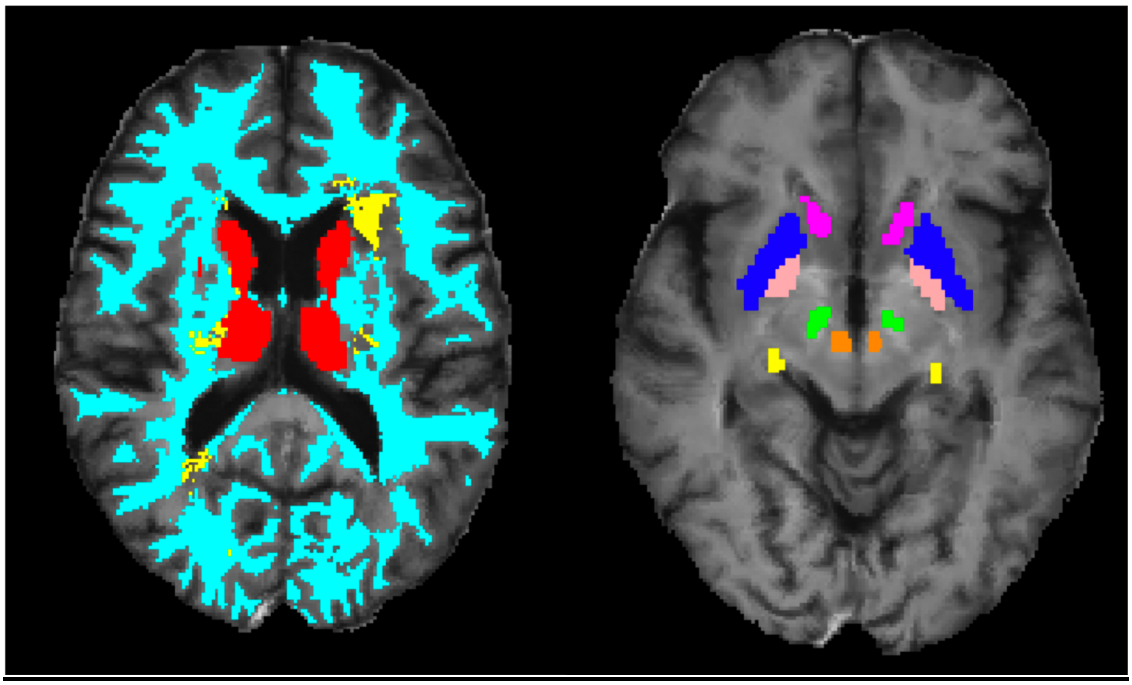


Figure 3-2: Illustration of NAWM, WML and DGM for one random MS participant.



(Left) color coded tissue ROIs overlaid on the participant's T1-weighted image, cyan: NAWM, yellow: WML, red: DGM. (Right) DGM mask is made up of six DGM regions, magenta: Caudate nucleus, blue: Putamen, pink: Globus pallidus, green: Substantia nigra, orange: Red nucleus and yellow, thalamus.

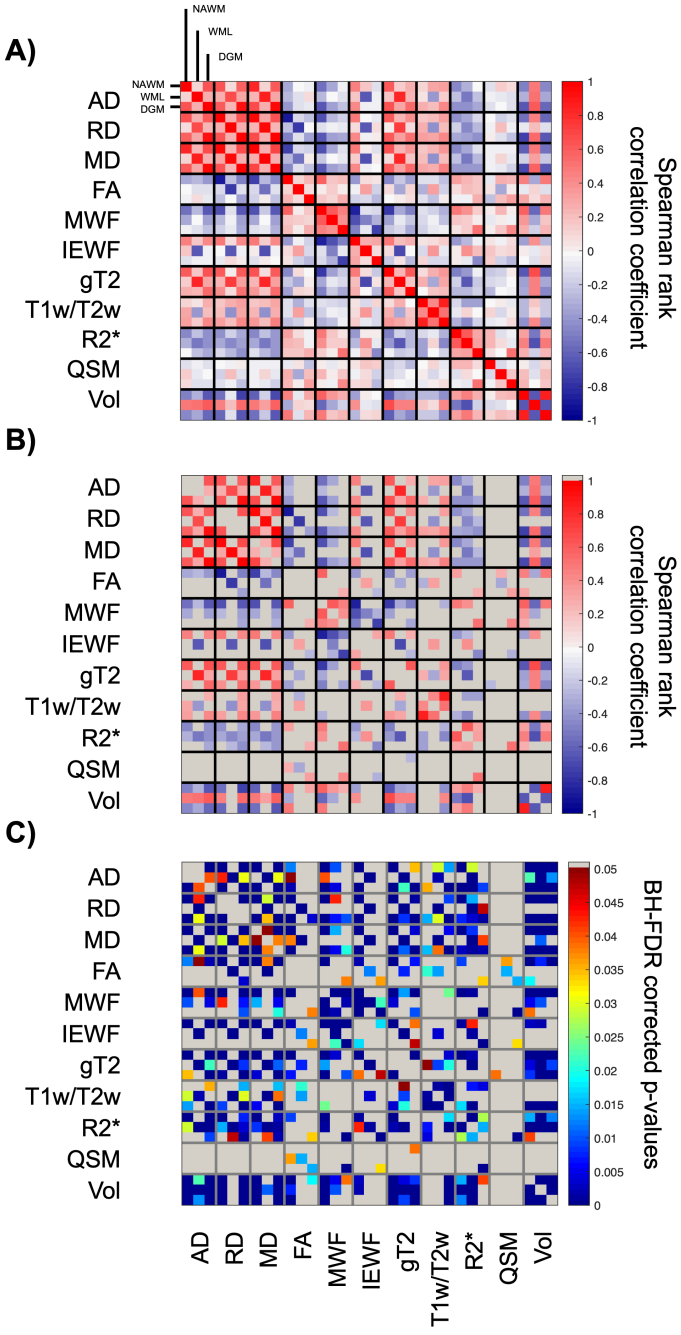
Statistical analyses

All ROI Vols were estimated by summing the ROI's voxel count for each subject in the subject space. For example, the Vol of DGM was estimated by summing voxel counts after reverting the "EvePM" Atlas (Lim et al., 2013) from MNI-152 space to individual subject space using the subject-specific backward deformation map, which is precomputed as part of the tissue segmentation output following the steps in previously published imaging protocol (Uddin et al., 2022) via Computational Anatomy Toolbox (CAT12 version r1318; http://www.neuro.uni-jena.de/cat/). In summary, these ROI were eventually brought back to and being analyzed in the template space. Averaging the metrics within each ROI yielded a total of 33 variables in the data set (i.e., 3 ROIs x 11 qMRI measurements = 33 variables). Box-and-whisker plots and the associated descriptive statistics (i.e., mean, standard deviation (Std) and coefficient of variation (CoV)) for 33 variables were computed to provide an overview of the data distribution across all qMRI metrics in all three tissue classes.

The correlations of qMRI metrics and volumetric measures among MS participants were computed, identical to the parallel work, in a pairwise manner for each ROI using Spearman rank correlation, with age and sex included as covariates for the statistical correlation analysis. Note that, despite the 33-by-33 matrix obtained by exhaustively pairing the metrics, the current study focused solely on volumetric measures and their relationships with other qMRI. That is, only a

subset of relationships (the last three rows/columns) from the full correlation matrix (Figure 3-3A) was being analyzed in this study, as the rest of the matrix was addressed and discussed in the parallel manuscript (Chapter 2). Nonetheless, to maintain the statistical integrity as the mentioned manuscript in parallel with investigating the rest of the relationship among qMRI in the correlation matrix, all p-values from unique relationships (including the ones that are not being looked at in the current manuscript) were adjusted by using the step-up Benjamini-Hochberg false discovery rate (BH-FDR) procedure to correct type-I errors from multiple comparisons (Benjamini & Hochberg, 1995), with the threshold of $p < 0.05$ after adjustment. The analyses above were performed in MATLAB (version R2020a; The MathWorks Inc., Natick, MA, USA). The number of associated significant correlations of each metric, and their shared percentage, calculated as the number of the associated significant correlations of each metric divided by the total unique significant correlations from the correlation map among qMRI-Vol correlations only, were computed to gain an entry-level understanding of how the significant correlation was distributed.

Figure 3-3: Age- and sex-corrected Spearman rank correlation matrices of 11 qMRI across 3 ROIs.



(A) Sex and age adjusted Spearman rank correlation matrix, (B) significant Spearman rank correlation matrix and associated (C) significant BH-FDR corrected p-values. Gray boxes indicate the entry is no longer significant after the BH-FDR correction.

3.4 Results

Box-and-whisker plots for the ten qMRI metrics and Vol are displayed in Figure 3-4. Descriptive statistics (i.e., mean, Std and CoV) of all distributions are summarized in Table 3-1. Particularly, for the brain volume measure, WML Vol shows greater variability across participants with a coefficient of variation of 1.13, compared with NAWM Vol (0.17) and DGM Vol (0.13). The total number (i.e., including qMRI-qMRI pairwise correlation) of significant Spearman rank correlation is 235 (out of 528), and the distribution of the number of associated significant correlation and shared percentage (over 235), of each metric is visualized in Figure 3-5, where Vol shows the most significant correlations among all (56, partially shared 23.8% of significant pairs with others). Individual age- and sex-corrected Spearman rank correlations between Vol and each quantitative metric, with corresponding significant p-values after BH-FDR correction and significant correlations are visualized in Figures 3-6, 3-7 and 3-8, respectively. The entire correlation matrix, its associated corrected p-values and significant Spearman rank correlations can be found in Figure 3-3. After the global BH-FDR correction and controlling for age and sex effects, 56 out of 93 unique correlations (i.e., all correlations except the intra-correlation and replicate of the inter-ROI correlation in Figure 3-6) remain significant (Figure 3-8). Of those that are significant after BH-FDR correction, AD, MD, gT2 and R2* have the most correlations with Vol, while IEWF and QSM have the least.

Figure 3-4: Box-and-whisker plots of 11 qMRI metrics.

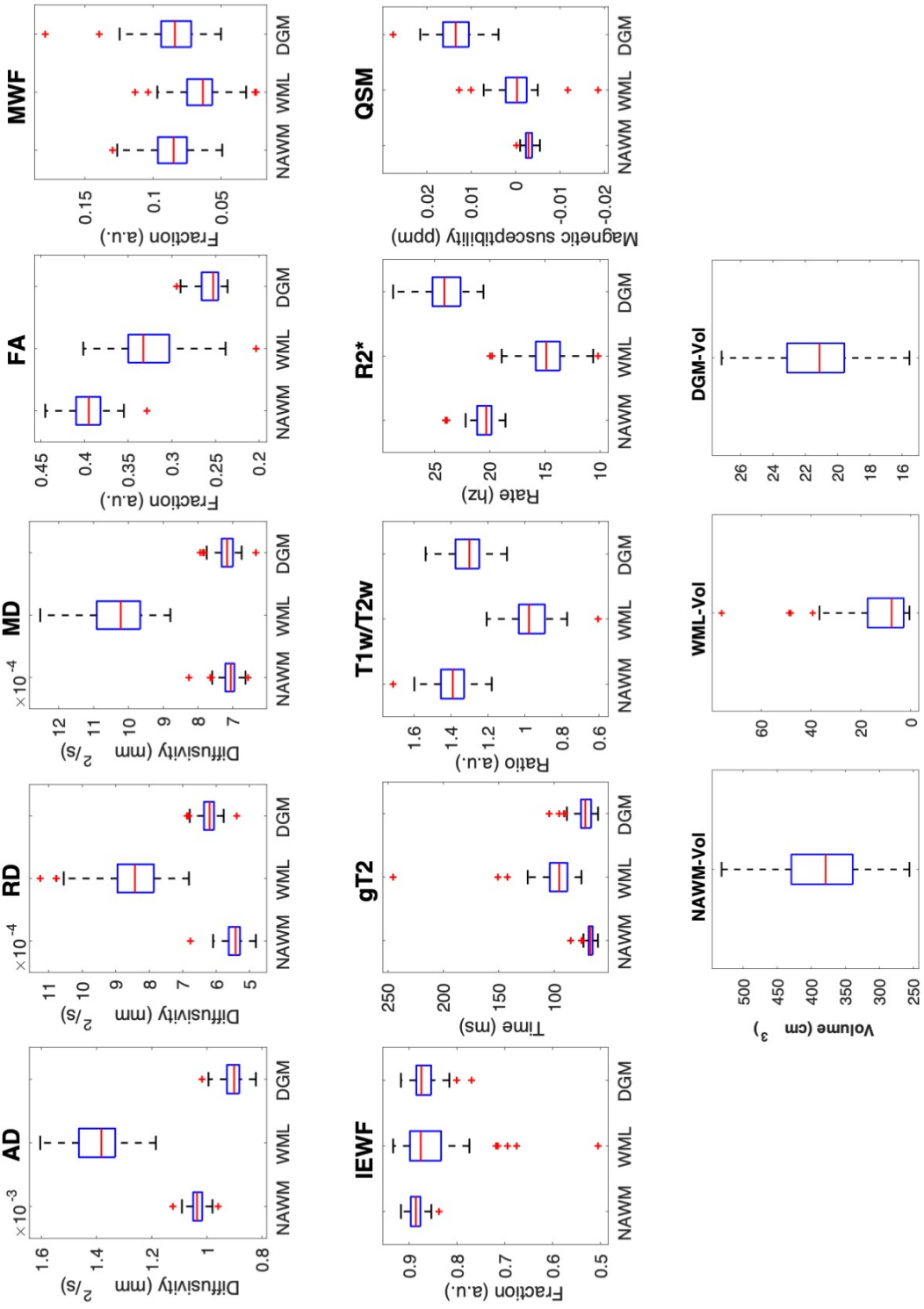


Table 3-1: Mean, standard deviation and coefficient of variation of all 33 qMRI-ROI pairings across participants.

Quantitative metrics (units)	Tissue classes	mean	Sample standard deviation	Coefficient of variation
AD (mm^2/s)	*NAWM*	0.0010	0.00003	0.026
	WML	0.0014	0.00009	0.062
	DGM	0.0009	0.00004	0.040
RD (mm^2/s)	*NAWM*	0.0005	0.00003	0.055
	WML	0.0009	0.00009	0.103
	DGM	0.0006	0.00003	0.045
MD (mm^2/s)	*NAWM*	0.0007	0.00003	0.038
	WML	0.0010	0.00008	0.079
	DGM	0.0007	0.00003	0.042
FA (a.u.)	*NAWM*	0.3939	0.02138	0.054
	WML	0.3258	0.03950	0.121
	DGM	0.2564	0.01330	0.052
MWF (a.u.)	*NAWM*	0.0865	0.01587	0.183
	WML	0.0650	0.01628	0.250
	DGM	0.0868	0.02058	0.237
IEWF (a.u.)	*NAWM*	0.8851	0.01489	0.017
	WML	0.8559	0.06669	0.078
	DGM	0.8671	0.02642	0.030
gT2 (ms)	*NAWM*	67.6399	3.64860	0.054
	WML	98.7070	21.02011	0.213
	DGM	72.9367	8.13183	0.111
T1w/T2w ratio (a.u.)	*NAWM*	1.3908	0.10274	0.074
	WML	0.9641	0.10224	0.106
	DGM	1.3107	0.10000	0.076
R2* (Hz)	*NAWM*	20.4878	1.01227	0.049
	WML	14.7693	1.93013	0.131
	DGM	24.0461	1.75999	0.073
QSM (ppm)	*NAWM*	-0.0030	0.00102	*
	WML	-0.0002	0.00420	*
	DGM	0.0137	0.00466	*
Vol (cm^3)	*NAWM*	384.2619	63.99337	0.167
	WML	11.7425	13.24742	1.128
	DGM	21.2311	2.71241	0.128

*BH-FDR corrected Wilcoxon Signed-Rank tests' p-values are shown on the right of the Table, where significant results were shown in bold fonts *Note: QSM's mean ranges from negative to positive, therefore the coefficient of variation was not calculated.*

Figure 3-5: Bar graph of 11 qMRI metric correlations and their shared percentages within 235 (out of 528) total unique significant correlations.

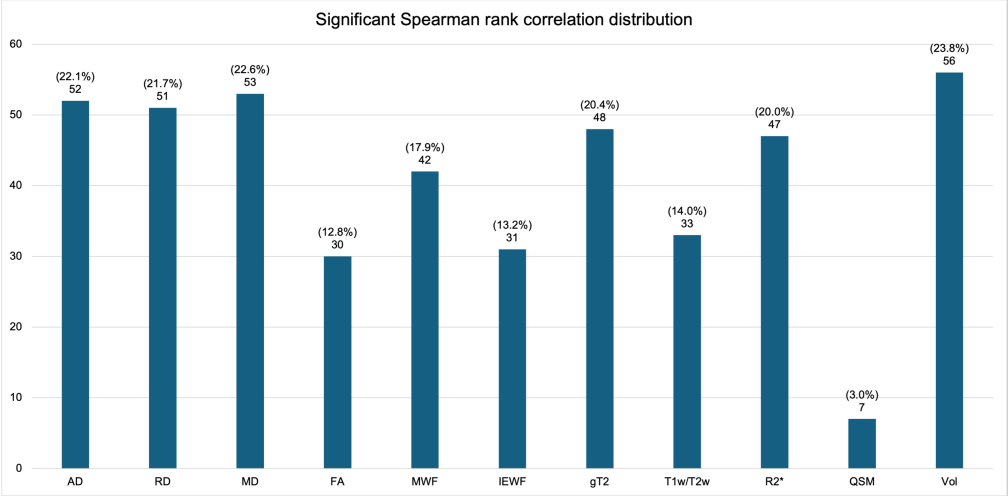


Figure 3-6: Age- and sex-corrected Spearman rank correlations of Vol with 10 qMRI metrics across 3 ROIs.

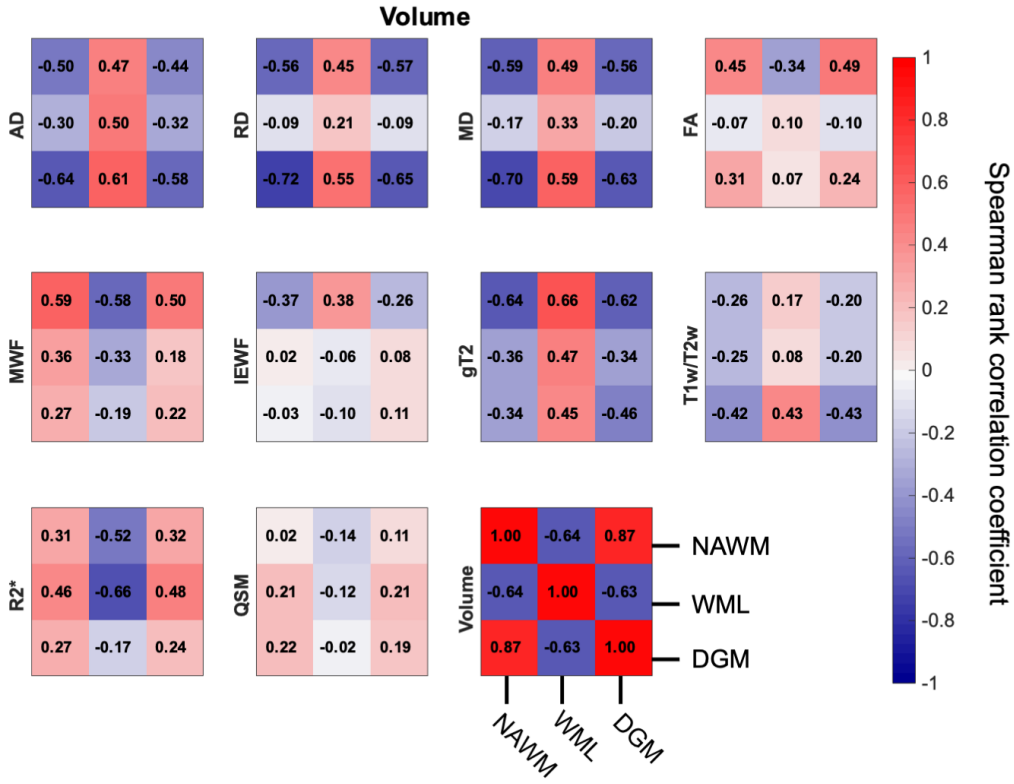
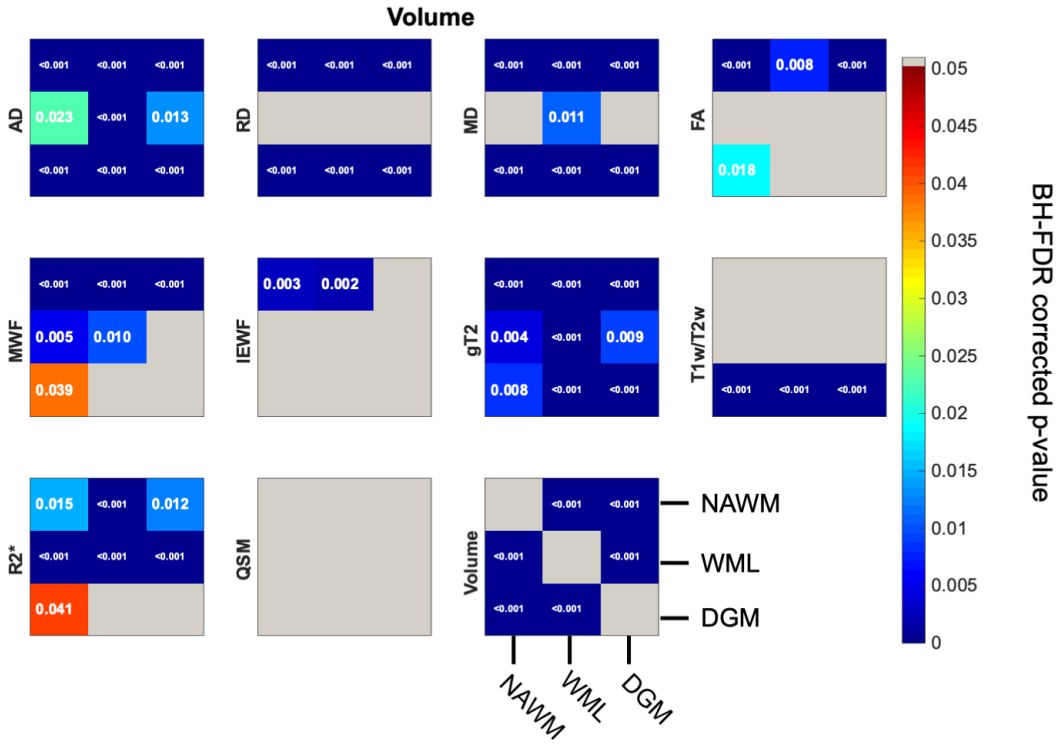
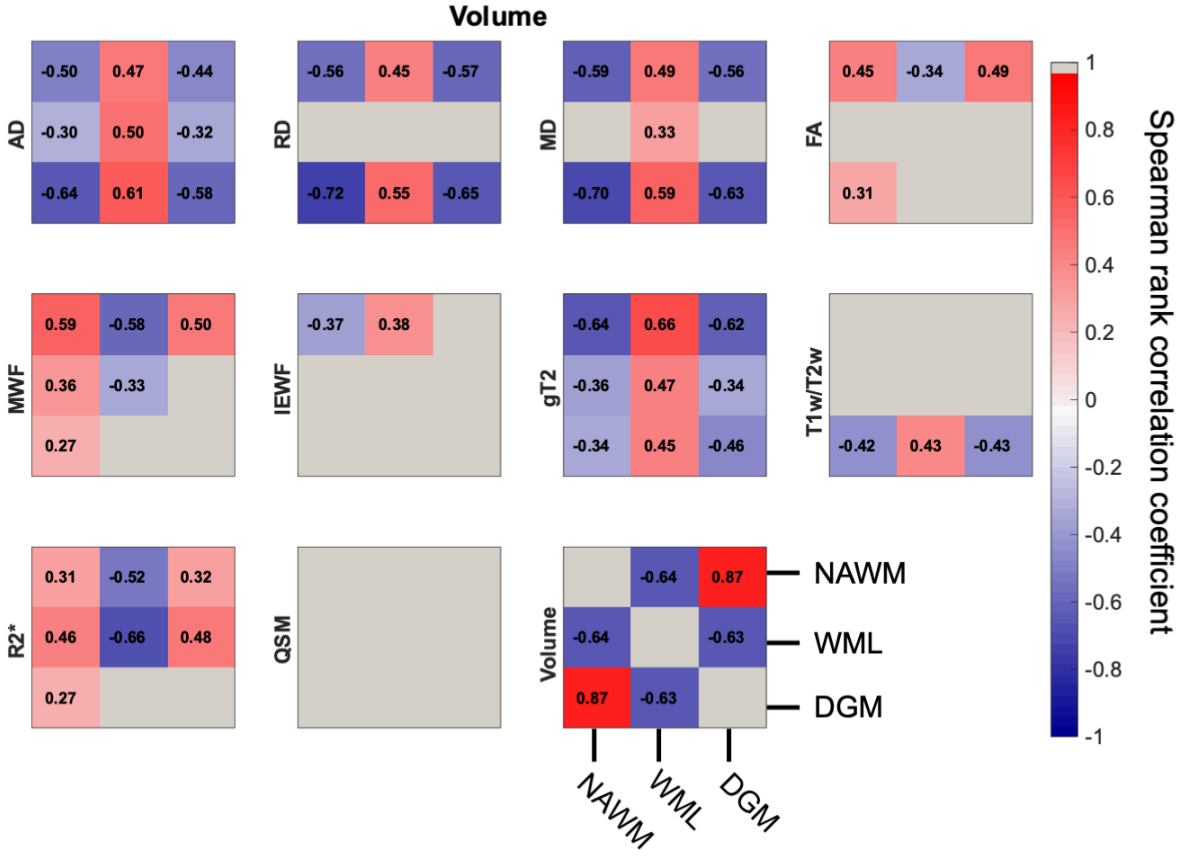


Figure 3-7: BH-FDR corrected p-values of the Spearman rank correlations of Vol with 10 qMRI metrics across 3 ROIs.



Gray boxes indicate the entry is no longer significant after the BH-FDR correction using entire matrix (Figure 3-3A).

Figure 3-8: Significant Spearman rank correlations of Vol with 10 quantitative metrics across 3 ROIs.



Gray boxes indicate the entry is no longer significant after the BH-FDR correction.

3.5 Discussion

Vol has the most shared percentage of significant correlations among the group of qMRI metrics (Figure 3-5), and that is nearly two-thirds (56/93) of the unique correlations with itself and the other 10 qMRI metrics remain significant after correction for multiple comparisons (Figure 3-8). These significant correlations with their correlation polarities, especially intra-ROI ones, suggest that there is a pronounced association between microstructural integrity and local brain volume

based on our interpretations and understanding of the biology and physics of the metrics. The off-diagonal correlations have shown the inter-regional dynamics of the correlation between volumetric measures and other quantitative measures. However, it is worth noting that the diagonal correlations (i.e., intra-ROI correlations) are neither all significant nor necessarily the strongest.

Volumetric measures correlate with a wide range of MRI metrics in MS

The finding that Vol is a metric that strongly and densely correlates with the other qMRI metrics was interesting yet anticipated by many clinical studies on the relationship between brain Vol (especially WML Vol and DGM Vol) and clinical scores and presentations. To date, no studies have systematically examined the correlation between volumes of different brain tissue classes and an extensive and multimodal set of qMRI metrics. However, one particular study showed a significant correlation between NAWM FA and cerebral volume (Larassati et al., 2022). Another study observed an association between the atrophied QSM rim+ lesion and MWF, but quantitative volume measures were not taken into account (Yao et al., 2018). In Figure 3-5, the magnitude of the significant correlation varies, but AD, RD, MD, gT2, and T1w/T2w ratio are more closely correlated with Vol, both in magnitude and density, than the other metrics. Interestingly, FA, MWF, IEWF, R2* and QSM, in contrast, show either weaker or fewer significant relationships with the volumes of these brain tissue classes. The line between the qMRI groups, which show either more or less correlation with the Vol groups, is unclear. However, we know that AD, RD, MD (diffusion-based) and gT2 and T1w/T2w ratio (relaxometry-based) are sensitive measures for detecting microstructural changes in the brain, but are less specific than, for example, MWF and QSM. Therefore, it remains an ongoing effort to further examine the exact relationship between volume

measures and underlying biological properties in MS neural repair and neural degeneration in association with MR readings.

Intra-ROI correlations between Vol and quantitative metrics in MS

The intra-ROI regions (diagonal entries) implied that a worse microstructural status of the ROI is associated with larger WML size and smaller normal tissue (i.e., NAWM and DGM) size. For example, each significant WML correlation in AD, RD, MD and gT2 is shown to be positive (i.e., red blocks), whereas MWF and R2* show significant negative correlations (i.e., blue blocks). This means that across multiple measures, individuals with higher WML volumes also had lower microstructural status (e.g., higher diffusivity, longer relaxation time and lower myelin water fraction), while individuals with lower WML volumes tended to have comparatively higher microstructural status within their lesions. The negative correlation between R2* and Vol in WML (Figure 3-8) is particularly interesting, in a way similar to an argument toward MWF that it may be specific to both myelin and iron (Birkl et al., 2019). R2* (extending the concept to R2') may also be sensitive to myelin in addition to iron, with orientation dependency (J. Lee et al., 2017; Paling et al., 2012; Weber et al., 2020). In the context of white matter, iron is usually found with macrophages and microglia (Kamma et al., 2022), and the density of activated phenotypes of the two is higher in active lesion borders (Tham et al., 2021). It was shown in this study that, on one hand, smaller WMLs tend to have higher R2* values, which may be due to their hyperactivity at the early phase or a relatively higher myelin being more likely to be conserved at early stages. Or on the other hand, larger WMLs tend to have a lower R2* value, which may be due to a combination of myelin breakdown, iron depletion, potentially at a later stage. However, a definitive conclusion cannot be made given that this is a cross-sectional analysis and lesion types were not

identified in the current study. In addition, NAWM and DGM Vol are significantly related to quantitative measures within ROI, mostly with a polarity opposite to that of WML.

Inter-ROI correlations between Vol and quantitative metrics in MS

One would anticipate that regional volume fluctuations and the microstructural changes in the same region, as described in the previous section, are more likely to be correlated than cross-region relationships. However, this dataset shows that the inter-ROI associations are equally shown to be significant. Since the significant correlation occurrence across ROI (i.e., 38 significant inter-ROI correlations divided by 63 total inter-ROI correlations = 60.3%) is slightly higher than the correlation occurrence within ROI (i.e., 18 significant intra-ROI correlations divided by 30 total intra-ROI correlations = 60.0%). Those significant correlations have suggested that: 1) people with higher Vol in NAWM, tend to have, higher MWF and R2*, and lower AD, gT2 and Vol in WML; higher FA, MWF, R2*and Vol, and lower AD, RD, MD, gT2 and T1w/T2w ratio in DGM 2) people with higher Vol in WML, tend to have, higher AD, RD, MD, IEWF and gT2, and lower FA, MWF, R2* and Vol (replicated) in NAWM; higher AD, RD, MD, gT2 and T1w/T2w ratio, and lower Vol in DGM. 3) people with higher Vol in DGM, tend to have, higher FA, MWF, R2* and Vol (replicated), and lower AD, RD, MD and gT2 in NAWM; higher R2*, and lower AD, gT2 and Vol (replicated) in WML. Although the significant findings of these inter-ROI correlations appear to be scattered (i.e., significance and strength of the correlations), and in extreme case, QSM does not show any correlations with Vol at all, suggesting the differences in qMRI metrics' sensitivity in general and across ROIs, the correlational matrices share pattern-wise structural similarity (Figure 3-8, and the pattern is more apparent in Figure 3-6) – the directionality of the correlations between Vol for one ROI and qMRI metric in the other two ROIs remain identical except for Vol-

Vol correlation (e.g., AD-Vol inter-ROI correlations: Vol in NAWM is negatively correlated to AD in both WML and DGM; Vol in WML is positively correlated to AD in both NAWM and DGM; Vol in DGM is negatively correlated to AD in both NAWM and WML). To the author's knowledge, this is the first study to discuss the dynamic association between MS brain regional volumes and quantitative measures. The topic of inter-ROI correlation between pairs of qMRI metrics was further discussed in our parallel manuscript (Chapter 2). These findings and the indirect evidence have motivated research to investigate the phenomenon or further elucidate the mechanism. However, it is clear that MS pathology and neural activities, including repair and neurodegeneration, are not confined to radiologically apparent regions (e.g., focal lesion, inflammation) but are instead globally distributed. Acknowledging also that qMRI metrics are sensitive but not generally specific, and their interpretation is region-dependent, this inter-ROI and "unidirectional" pattern indicates tissue microstructural damages exist globally, possibly in the same direction, at least as shown consistently by the majority of the qMRI metrics. This also reinforced the NAWM naming convention, highlighting the underlying subtle abnormalities in WM (Ludwin, 2006).

Study limitations

Since this study design shares the same framework as the study from the same lab that focuses solely on the correlations between qMRI metrics (the rest of the relationships of the Spearman rank correlation) (Chapter 2), the majority of the limitations are already outlined in the parallel manuscript (Chapter 2). Those include the choices of ROI computation, imaging sequence parameters, metric processing pipelines, and the decomposition of the Spearman rank correlation matrix. Nevertheless, speaking of ways to decompose matrices, the finding of the study was based

on the robust decomposition of the whole Spearman rank correlation onto individual metrics, that is volumetric measures in this case. The other manuscript (Chapter 2) can be theoretically split and organized in the same way. However, this will lead to redundant findings (i.e., discussion from qMRI metrics' perspective in the opposite ways) because of the symmetry of the matrix, potentially scattering the findings and confusing readers. Since qMRI-qMRI and qMRI-Vol matrices together form a complete basis of the original Spearman rank correlation matrix, it minimizes the complexity of analyses and avoids redundant findings. Ultimately, despite the fact that both study analyses were done in a similar way, the research questions and focuses remain different, BH-FDR was applied to the whole matrix to preserve statistical and scientific rigor.

3.6 Conclusions

Volumetric measures were found to correlate with most qMRI metrics both within and, intriguingly, across the three tissue classes. Besides significant associations between Vol and qMRI metrics within WML (the best-known site of investigation in MS), Vol also shows strong relationships with qMRI metrics in both NAWM and DGM. This is the first paper to characterize the relationship between volumetric and a large and diverse set of qMRI metrics (at least in those bulk compartments) in a large MS cohort, and to the best of our knowledge, it is the first MRI study to empirically show a significant positive correlation between total lesion volume and average lesion severity/activities. We hope the result raises awareness of emerging qMRI methods and highlights their association with volumetric measures in MS. Most excitingly, this adds to the driving force for future studies to leverage the relationship between quantitative MR metrics and Vol to gain a better understanding of the underlying mechanism and yield a more specific measure to inform clinical decisions in MS.

3.7 References

Andravizou, A., Dardiotis, E., Artemiadis, A., Sokratous, M., Siokas, V., Tsouris, Z., Aloizou, A.-M., Nikolaidis, I., Bakirtzis, C., Tsivgoulis, G., Deretzi, G., Grigoriadis, N., Bogdanos, D. P., & Hadjigeorgiou, G. M. (2019). Brain atrophy in multiple sclerosis: mechanisms, clinical relevance and treatment options. *Autoimmunity Highlights*, *10*(1), 7. https://doi.org/10.1186/s13317-019-0117-5

Barnett, Y., Garber, J. Y., & Barnett, M. H. (2020). MRI biomarkers of disease progression in multiple sclerosis: old dog, new tricks? *Quantitative Imaging in Medicine and Surgery*, *10*(2), 527–532. https://doi.org/10.21037/qims.2020.01.04

Bartzokis, G., Mintz, J., Marx, P., Osborn, D., Gutkind, D., Chiang, F., Phelan, C. K., & Marder, S. R. (1993). Reliability of in vivo volume measures of hippocampus and other brain structures using MRI. *Magnetic Resonance Imaging*, *11*(7), 993–1006. https://doi.org/10.1016/0730-725X(93)90218-3

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Birkl, C., Birkl-Toeglhofer, A. M., Endmayr, V., Höftberger, R., Kasprian, G., Krebs, C., Haybaeck, J., & Rauscher, A. (2019). The influence of brain iron on myelin water imaging. *NeuroImage*, *199*, 545–552. https://doi.org/10.1016/j.neuroimage.2019.05.042

Bitsch, A. (2000). Acute axonal injury in multiple sclerosis: Correlation with demyelination and inflammation. *Brain*, *123*(6), 1174–1183. https://doi.org/10.1093/brain/123.6.1174

Bjartmar, C., Wujek, J. R., & Trapp, B. D. (2003). Axonal loss in the pathology of MS: consequences for understanding the progressive phase of the disease. *Journal of the Neurological Sciences*, *206*(2), 165–171. https://doi.org/10.1016/S0022-510X(02)00069-2

Boaventura, M., Sastre-Garriga, J., Garcia-Vidal, A., Vidal-Jordana, A., Quartana, D., Carvajal, R., Auger, C., Alberich, M., Tintoré, M., Rovira, À., Montalban, X., & Pareto, D. (2022). T1/T2-weighted ratio in multiple sclerosis: A longitudinal study with clinical associations. *NeuroImage: Clinical*, *34*, 102967. https://doi.org/10.1016/j.nicl.2022.102967

Chard, D. T. (2003). The longitudinal relation between brain lesion load and atrophy in multiple sclerosis: a 14 year follow up study. *Journal of Neurology, Neurosurgery & Psychiatry*, *74*(11), 1551–1554. https://doi.org/10.1136/jnnp.74.11.1551

Dieleman, N., Koek, H. L., & Hendrikse, J. (2017). Short-term mechanisms influencing volumetric brain dynamics. *NeuroImage: Clinical*, *16*, 507–513. https://doi.org/10.1016/j.nicl.2017.09.002

Fisher, E., Rudick, R. A., Simon, J. H., Cutter, G., Baier, M., Lee, J.-C., Miller, D., Weinstock-Guttman, B., Mass, M. K., Dougherty, D. S., & Simonian, N. A. (2002). Eight-year follow-up study of brain atrophy in patients with MS. *Neurology*, *59*(9), 1412–1420. https://doi.org/10.1212/01.WNL.0000036271.49066.06

Fox, N. C., Jenkins, R., Leary, S. M., Stevenson, V. L., Losseff, N. A., Crum, W. R., Harvey, R. J., Rossor, M. N., Miller, D. H., & Thompson, A. J. (2000). Progressive cerebral atrophy in MS: A serial study using registered, volumetric MRI. *Neurology*, *54*(4), 807–812. https://doi.org/10.1212/WNL.54.4.807

Fujimori, J., Fujihara, K., Ogawa, R., Baba, T., Wattjes, M., & Nakashima, I. (2020). Patterns of regional brain volume loss in multiple sclerosis: a cluster analysis. *Journal of Neurology*, *267*(2), 395–405. https://doi.org/10.1007/s00415-019-09595-4

Giorgio, A., & De Stefano, N. (2013). Clinical use of brain volumetry. *Journal of Magnetic Resonance Imaging*, *37*(1), 1–14. https://doi.org/10.1002/jmri.23671

Glasser, M. F., & van Essen, D. C. (2011). Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *Journal of Neuroscience*, *31*(32), 11597–11616. https://doi.org/10.1523/JNEUROSCI.2180-11.2011

Granziera, C., Wuerfel, J., Barkhof, F., Calabrese, M., De Stefano, N., Enzinger, C., Evangelou, N., Filippi, M., Geurts, J. J. G., Reich, D. S., Rocca, M. A., Ropele, S., Rovira, À., Sati, P., Toosy, A. T., Vrenken, H., Gandini Wheeler-Kingshott, C. A. M., Kappos, L., Barkhof, F., … Yousry, T. (2021). Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. *Brain*, *144*(5), 1296–1311. https://doi.org/10.1093/brain/awab029

Haider, L., Simeonidou, C., Steinberger, G., Hametner, S., Grigoriadis, N., Deretzi, G., Kovacs, G. G., Kutzelnigg, A., Lassmann, H., & Frischer, J. M. (2014). Multiple sclerosis deep grey matter: the relation between demyelination, neurodegeneration, inflammation and iron. *Journal of Neurology, Neurosurgery & Psychiatry*, *85*(12), 1386–1395. https://doi.org/10.1136/jnnp-2014-307712

Harkey, T., Baker, D., Hagen, J., Scott, H., & Palys, V. (2022). Practical methods for segmentation and calculation of brain volume and intracranial volume: a guide and comparison. *Quantitative Imaging in Medicine and Surgery*, *12*(7), 3748–3761. https://doi.org/10.21037/qims-21-958

Jagust, W. J., & Noseworthy, J. H. (2000). Brain atrophy as a surrogate marker in MS: Faster, simpler, better? *Neurology*, *54*(4), 782–783. https://doi.org/10.1212/WNL.54.4.782

Kamma, E., Lasisi, W., Libner, C., Ng, H. S., & Plemel, J. R. (2022). Central nervous system macrophages in progressive multiple sclerosis: relationship to neurodegeneration and therapeutics. *Journal of Neuroinflammation*, *19*(1), 45. https://doi.org/10.1186/s12974-022-02408-y

Korn, T. (2008). Pathophysiology of multiple sclerosis. In *Journal of Neurology* (Vol. 255, Number SUPPL. 6, pp. 2–6). https://doi.org/10.1007/s00415-008-6001-2

Langkammer, C., Liu, T., Khalil, M., Enzinger, C., Jehna, M., Fuchs, S., Fazekas, F., Wang, Y., & Ropele, S. (2013). Quantitative Susceptibility Mapping in Multiple Sclerosis. *Radiology*, *267*(2), 551–559. https://doi.org/10.1148/radiol.12120707

Larassati, H., Pandelaki, J., Estiasari, R., Prihartono, J., Firdausia, S., Yunus, R. E., & Mulyadi, R. (2022). Diffusion magnetic resonance imaging of normal-appearing white matter in multiple sclerosis: correlation with brain volume and clinical disability. *Journal of Central Nervous System Disease*, *14*, 117957352210981. https://doi.org/10.1177/11795735221098147

Laule, C., Kozlowski, P., Leung, E., Li, D. K. B., MacKay, A. L., & Moore, G. R. W. (2008). Myelin water imaging of multiple sclerosis at 7 T: Correlations with histopathology. *NeuroImage*, *40*(4), 1575–1580. https://doi.org/10.1016/j.neuroimage.2007.12.008

Lavorgna, L., Bonavita, S., Ippolito, D., Lanzillo, R., Salemi, G., Patti, F., Valentino, P., Coniglio, G., Buccafusca, M., Paolicelli, D., d'Ambrosio, A., Bresciamorra, V., Savettieri, G., Zappia, M., Alfano, B., Gallo, A., Simone, I., & Tedeschi, G. (2014). Clinical and magnetic resonance imaging predictors of disease progression in multiple sclerosis: a nine-year follow-up study. *Multiple Sclerosis Journal*, *20*(2), 220–226. https://doi.org/10.1177/1352458513494958

Lee, J., Shin, H.-G., Jung, W., Nam, Y., Oh, S.-H., & Lee, J. (2017). An R2* model of white matter for fiber orientation and myelin concentration. *NeuroImage*, *162*, 269–275. https://doi.org/10.1016/j.neuroimage.2017.08.050

Lee, J. Y., Oh, S. W., Chung, M. S., Park, J. E., Moon, Y., Jeon, H. J., & Moon, W.-J. (2021). Clinically Available Software for Automatic Brain Volumetry: Comparisons of Volume Measurements and Validation of Intermethod Reliability. *Korean Journal of Radiology*, *22*(3), 405. https://doi.org/10.3348/kjr.2020.0518

Lim, I. A. L., Faria, A. V., Li, X., Hsu, J. T. C., Airan, R. D., Mori, S., & van Zijl, P. C. M. (2013). Human brain atlas for automated region of interest selection in quantitative susceptibility mapping: Application to determine iron content in deep gray matter structures. *NeuroImage*, *82*, 449–469. https://doi.org/10.1016/j.neuroimage.2013.05.127

Ludwin, S. K. (2006). The Pathogenesis of Multiple Sclerosis. *Journal of Neuropathology and Experimental Neurology*, *65*(4), 305–318. https://doi.org/10.1097/01.jnen.0000225024.12074.80

Luft, A. R., Skalej, M., Welte, D., Kolb, R., & Klose, U. (1996). Reliability and exactness of MRI-based volumetry: A phantom study. *Journal of Magnetic Resonance Imaging*, *6*(4), 700–704. https://doi.org/10.1002/jmri.1880060421

MacKay, A. L., & Laule, C. (2016). Magnetic Resonance of Myelin Water: An in vivo Marker for Myelin. *Brain Plasticity*, *2*(1), 71–91. https://doi.org/10.3233/bpl-160033

Maclaren, J., Han, Z., Vos, S. B., Fischbein, N., & Bammer, R. (2014). Reliability of brain volume measurements: A test-retest dataset. *Scientific Data*, *1*(1), 140037. https://doi.org/10.1038/sdata.2014.37

Morell, P., & Norton, W. T. (1980). Myelin. *Scientific American*, *242*(5), 88–114. https://doi.org/10.1038/scientificamerican0580-88

Pagani, E., Rocca, M. A., Gallo, A., Rovaris, M., Martinelli, V., Comi, G., & Filippi, M. (2005). Regional brain atrophy evolves differently in patients with multiple sclerosis according to clinical phenotype. *AJNR. American Journal of Neuroradiology*, *26*(2), 341–346. http://www.ncbi.nlm.nih.gov/pubmed/15709132

Paling, D., Tozer, D., Wheeler-Kingshott, C., Kapoor, R., Miller, D. H., & Golay, X. (2012). Reduced R_2' in multiple sclerosis normal appearing white matter and lesions may reflect decreased myelin and iron content. *Journal of Neurology, Neurosurgery & Psychiatry*, *83*(8), 785–792. https://doi.org/10.1136/jnnp-2012-302541

Popescu, V., Agosta, F., Hulst, H. E., Sluimer, I. C., Knol, D. L., Sormani, M. P., Enzinger, C., Ropele, S., Alonso, J., Sastre-Garriga, J., Rovira, A., Montalban, X., Bodini, B., Ciccarelli, O., Khaleeli, Z., Chard, D. T., Matthews, L., Palace, J., Giorgio, A., … Vrenken, H. (2013). Brain atrophy and lesion load predict long term disability in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, *84*(10), 1082–1091. https://doi.org/10.1136/jnnp-2012-304094

Pravatà, E., Rocca, M. A., Valsasina, P., Riccitelli, G. C., Gobbi, C., Comi, G., Falini, A., & Filippi, M. (2017). Gray matter trophism, cognitive impairment, and depression in patients with multiple sclerosis. *Multiple Sclerosis Journal*, *23*(14), 1864–1874. https://doi.org/10.1177/1352458517692886

Preziosa, P., Pagani, E., Meani, A., Moiola, L., Rodegher, M., Filippi, M., & Rocca, M. A. (2022). Slowly Expanding Lesions Predict 9-Year Multiple Sclerosis Disease Progression. *Neurology - Neuroimmunology Neuroinflammation*, *9*(2), e1139. https://doi.org/10.1212/NXI.0000000000001139

Radue, E.-W., Barkhof, F., Kappos, L., Sprenger, T., Haring, D. A., de Vera, A., von Rosenstiel, P., Bright, J. R., Francis, G., & Cohen, J. A. (2015). Correlation between brain volume loss and clinical and MRI outcomes in multiple sclerosis. *Neurology*, *84*(8), 784–793. https://doi.org/10.1212/WNL.0000000000001281

Rocca, M. A., Comi, G., & Filippi, M. (2017). The Role of T1-Weighted Derived Measures of Neurodegeneration for Assessing Disability Progression in Multiple Sclerosis. *Frontiers in Neurology*, *8*, 433. https://doi.org/10.3389/fneur.2017.00433

Rotstein, D., & Montalban, X. (2019). Reaching an evidence-based prognosis for personalized treatment of multiple sclerosis. *Nature Reviews Neurology*, *15*(5), 287–300. https://doi.org/10.1038/s41582-019-0170-8

Rovaris, M. (2001). Short-term brain volume change in relapsing-remitting multiple sclerosis: Effect of glatiramer acetate and implications. *Brain*, *124*(9), 1803–1812. https://doi.org/10.1093/brain/124.9.1803

Rudick, R. A., Fisher, E., Lee, J.-C., Duda, J. T., & Simon, J. (2000). Brain atrophy in relapsing multiple sclerosis: relationship to relapses, EDSS, and treatment with interferon β-1a. *Multiple Sclerosis Journal*, *6*(6), 365–372. https://doi.org/10.1177/135245850000600601

Sampaio-Baptista, C., & Johansen-Berg, H. (2017). White Matter Plasticity in the Adult Brain. *Neuron*, *96*(6), 1239–1251. https://doi.org/10.1016/j.neuron.2017.11.026

Sandrone, S., Aiello, M., Cavaliere, C., Thiebaut de Schotten, M., Reimann, K., Troakes, C., Bodi, I., Lacerda, L., Monti, S., Murphy, D., Geyer, S., Catani, M., & Dell'Acqua, F. (2023). Mapping myelin in white matter with T1-weighted/T2-weighted maps: discrepancy with histology and other myelin MRI measures. *Brain Structure and Function*, *228*(2), 525–535. https://doi.org/10.1007/s00429-022-02600-z

Sbardella, E., Tona, F., Petsas, N., & Pantano, P. (2013). DTI Measurements in Multiple Sclerosis: Evaluation of Brain Damage and Clinical Implications. *Multiple Sclerosis International*, *2013*, 1–11. https://doi.org/10.1155/2013/671730

Schweser, F., Deistung, A., & Reichenbach, J. R. (2016). Foundations of MRI phase imaging and processing for Quantitative Susceptibility Mapping (QSM). *Zeitschrift Fur Medizinische Physik*, *26*(1), 6–34. https://doi.org/10.1016/j.zemedi.2015.10.002

Stankiewicz, J. M., Neema, M., & Ceccarelli, A. (2014). Iron and multiple sclerosis. In *Neurobiology of Aging* (Vol. 35, Number SUPPL.2). Elsevier Inc. https://doi.org/10.1016/j.neurobiolaging.2014.03.039

Stephenson, E., Nathoo, N., Mahjoub, Y., Dunn, J. F., & Yong, V. W. (2014). Iron in multiple sclerosis: roles in neurodegeneration and repair. *Nature Reviews Neurology*, *10*(8), 459–468. https://doi.org/10.1038/nrneurol.2014.118

Tavares, V., Prata, D., & Ferreira, H. A. (2020). Comparing SPM12 and CAT12 segmentation pipelines: a brain tissue volume-based age and Alzheimer's disease study. *Journal of Neuroscience Methods*, *334*, 108565. https://doi.org/10.1016/j.jneumeth.2019.108565

Tham, M., Frischer, J. M., Weigand, S. D., Fitz-Gibbon, P. D., Webb, S. M., Guo, Y., Adiele, R. C., Robinson, C. A., Brück, W., Lassmann, H., Furber, K. L., Pushie, M. J., Parisi, J. E., Lucchinetti, C. F., & Popescu, B. F. (2021). Iron Heterogeneity in Early Active Multiple Sclerosis Lesions. *Annals of Neurology*, *89*(3), 498–510. https://doi.org/10.1002/ana.25974

Trapp, B. D., Peterson, J., Ransohoff, R. M., Rudick, R., Mörk, S., & Bö, L. (1998). Axonal Transection in the Lesions of Multiple Sclerosis. *New England Journal of Medicine*, *338*(5), 278–285. https://doi.org/10.1056/NEJM199801293380502

Uddin, M. N., Figley, T. D., Kornelsen, J., Mazerolle, E. L., Helmick, C. A., O'Grady, C. B., Pirzada, S., Patel, R., Carter, S., Wong, K., Essig, M. R., Graff, L. A., Bolton, J. M., Marriott, J. J., Bernstein, C. N., Fisk, J. D., Marrie, R. A., & Figley, C. R. (2022). The comorbidity and cognition in multiple sclerosis (CCOMS) neuroimaging protocol: Study rationale, MRI acquisition, and minimal image processing pipelines. *Frontiers in Neuroimaging*, *1*. https://doi.org/10.3389/fnimg.2022.970385

Uddin, M. N., Figley, T. D., Marrie, R. A., & Figley, C. R. (2018). Can T_1 w/ T_2 w ratio be used as a myelin-specific measure in subcortical structures? Comparisons between FSE-based T_1 w/ T_2 w ratios, GRASE-based T_1 w/ T_2 w ratios and multi-echo GRASE-based myelin water fractions. *NMR in Biomedicine*, *31*(3), e3868. https://doi.org/10.1002/nbm.3868

Uddin, M. N., Figley, T. D., Solar, K. G., Shatil, A. S., & Figley, C. R. (2019). Comparisons between multi-component myelin water fraction, T1w/T2w ratio, and diffusion tensor imaging measures in healthy human brain structures. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-39199-x

Uher, T., Bergsland, N., Krasensky, J., Dwyer, M. G., Andelova, M., Sobisek, L., Havrdova, E. K., Horakova, D., Zivadinov, R., & Vaneckova, M. (2021). Interpretation of Brain Volume Increase in Multiple Sclerosis. *Journal of Neuroimaging*, *31*(2), 401–407. https://doi.org/10.1111/jon.12816

Weber, A. M., Zhang, Y., Kames, C., & Rauscher, A. (2020). Myelin water imaging and R_2^* mapping in neonates: Investigating R_2^* dependence on myelin and fibre orientation in whole brain white matter. *NMR in Biomedicine*, *33*(3). https://doi.org/10.1002/nbm.4222

Williams, R., Buchheit, C. L., Berman, N. E. J., & Levine, S. M. (2012). Pathogenic implications of iron accumulation in multiple sclerosis. In *Journal of Neurochemistry* (Vol. 120, Number 1, pp. 7–25). https://doi.org/10.1111/j.1471-4159.2011.07536.x

Wood, J. C., Enriquez, C., Ghugre, N., Tyzka, J. M., Carson, S., Nelson, M. D., & Coates, T. D. (2005). MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. *Blood*, *106*(4), 1460–1465. https://doi.org/10.1182/blood-2004-10-3982

Yao, Y., Nguyen, T. D., Pandya, S., Zhang, Y., Hurtado Rúa, S., Kovanlikaya, I., Kuceyeski, A., Liu, Z., Wang, Y., & Gauthier, S. A. (2018). Combining Quantitative Susceptibility Mapping with Automatic Zero Reference (QSM0) and Myelin Water Fraction Imaging to Quantify Iron-Related Myelin Damage in Chronic Active MS Lesions. *American Journal of Neuroradiology*, *39*(2), 303–310. https://doi.org/10.3174/ajnr.A5482

Zierfuss, B., Wang, Z., Jackson, A. N., Moezzi, D., & Yong, V. (2023). Iron in multiple sclerosis – Neuropathology, immunology, and real-world considerations. *Multiple Sclerosis and Related Disorders*, *78*, 104934. https://doi.org/10.1016/j.msard.2023.104934

Zivadinov, R., & Bakshi, R. (2004). Central Nervous System Atrophy and Clinical Status in Multiple Sclerosis. *Journal of Neuroimaging*, *14*(3), 27–35. https://doi.org/10.1177/1051228404266266

Chapter 4

Mapping quantitative MRI changes as a function of distance from the boundaries of multiple sclerosis white matter lesions

Publication status

Parts of this chapter were presented in the abstract entitled '*Investigating the Signal Transition from White Matter Lesion Core to Perilesional White Matter in Multiple Sclerosis Using Quantitative Magnetic Resonance Imaging Techniques'* at the conferences: **Manitoba Student Health Research Forum**, **Manitoba Neuroscience Network** and **EndMS Conference** in 2023, and **Americas Committee for Treatment and Research in Multiple Sclerosis** in 2024. The manuscript is currently under internal review before submission to a peer-reviewed journal.

Peer-review conference abstract

Wong, K., Uddin, M. N., Figley, T. D., Kornelsen, J., Helmick, C. A., O'Grady, C. B., Mazerolle, E. L., Patel, R., Bernstein, C. N., & Fisk, J. D. (2024). Investigating Quantitative Magnetic Resonance Imaging Transitions from Multiple Sclerosis White Matter Lesions to Perilesional Normal Appearing White Matter. *MULTIPLE SCLEROSIS JOURNAL*, *30*(1), 128–129.

4.1 Abstract

White matter lesions (WMLs) are a key pathological characteristic of multiple sclerosis (MS). With extended spatial sensitivity shown by quantitative MRI (qMRI), this study aimed to investigate the spatial gradients of 10 different advanced qMRI metrics (axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD), fractional anisotropy (FA), myelin water fraction

(MWF), intra- and extra- cellular water fraction (IEWF), geometric mean T2 (gT2), T1-weighted-by-T2-weighted (T1w/T2w) ratio, R2* and quantitative susceptibility mapping (QSM)) extending from WML cores to perilesional normal appearing white matter (pNAWM) in an MS cohort (N=79). The analysis was conducted at the individual-lesion level (N=926, separated by voxel connectedness). Pre-computed WML mask of each WML was eroded (-) or dilated (+) with a discrete step size of 1 mm and ranges from -4 to +10. Each shell was used to extract the average across 10 quantitative MRI maps. The Mann-Whitney U-test was employed to assess how rapidly the adjacent shells transit, and Spearman rank correlation was used to examine both inter-shell correlations within qMRI metrics and intra-shell correlations between qMRI metrics. All qMRI exhibited mostly a linear-like signal trend from the inner WML core to the pNAWM, reaching a plateau at the outer pNAWM. The Mann-Whitney test result revealed a significant transition not only at (10/10 qMRI metrics), but also around (9/10 qMRI metrics) the WML-pNAWM boundary, even after correcting for multiple comparisons using the Bonferroni method (i.e., $p_{FWE} < 0.05$). Inter-shell Spearman rank correlation analysis has found significant relationships are more likely to cluster in either tissue classes for each qMRI metric as compared to across tissue class and intra-shell Spearman rank correlation analysis between qMRI highlighted the heterogeneous pattern of correlation around the WML-pNAWM boundary, with all p-values adjusted by the false discovery rate (i.e., $p_{BH\text{-}FDR} < 0.05$). Together, the results underscored the importance of the tissue interface region and suggested that the including pNAWM could yield additional information beyond WML in MS.

4.2 Introduction

One of the defining features of multiple sclerosis (MS) is the presence of focal white matter lesions (WMLs) throughout the brain (Lassmann, 2014; Lassmann et al., 2007). Studies have shown that the inflammatory response in the central nervous system is triggered by the infiltration of immune cells through the blood-brain barrier and is mediated by glial cells in the early stage of the disease (Korn, 2008; Lassmann, 2014). Among paraclinical tools, magnetic resonance imaging (MRI) plays a key role in identifying lesion location and type to inform better clinical decisions (Bakshi et al., 2008; Filippi & Rocca, 2007; Hemond & Bakshi, 2018).

However, while WML characterization is important, so-called normal appearing white matter (NAWM) may provide clues about tissue transition, since both in vivo studies using advanced MRI sequences and ex vivo studies using pathological staining have demonstrated that NAWM in MS has often already experienced degrees of microstructural damage and tissue deterioration (Miller et al., 2003; Moll et al., 2011). One research group previously found that among people with MS (in the relapsing-remitting form), NAWM permeability and inflammation severity are associated with clinical presentation: processing speed and cognitive decline (Eftekhari et al., 2017). Interestingly, another study from a different group subsequently showed that early NAWM inflammation disrupts nodes of Ranvier, affecting signal conduction speed (Gallego-Delgado et al., 2020). Retrospectively, it often received less attention as compared to WML. Given that WM is a significant part of the brain, both volumetrically and functionally, understanding early pathologies in NAWM may not only help elucidate the mechanisms underlying early lesion formation but also inform the interpretation of the clinical presentation (Miller et al., 2003). Furthermore, despite the growing number of studies on the MS prodromal phase (Harding et al., 2024; Makhani & Tremlett, 2021), predicting conversion to MS remains

difficult due to the limited sensitivity of current tools. However, a few studies that collected serial scans of people with MS with longitudinal quantitative MRI (qMRI) sequences concluded that qMRI signals offer insight into the specific spatial heterogeneity in NAWM and each of which shows distinct patterns of transitions from NAWM to lesions, highlighting the potential of MRI to reveal lesion precursors in the landscape of NAWM (Elliott et al., 2021; Goodkin et al., 1998).

It is worth noting that WMLs are constantly evolving during the course of MS, both pathologically and morphologically (Guttmann et al., 1995; Inglese et al., 2005; McDonald et al., 1992; Rovira et al., 2013). Macroscopically, tissue transitions (both from NAWM to WML and WML to NAWM) are visible on longitudinal MRI, particularly around the WML periphery and perilesional NAWM (pNAWM) (Barkhof et al., 2003; Rovira et al., 2013). This begs the questions: how closely are the two tissue classes related to each other microscopically, and how sharp is the gradient between the two tissue classes (and if smooth, how far into WMLs and out into NAWM are noticeable changes observed)? A study back in 2006 examined the pNAWM at varying distances from the lesion using T1 and magnetic transfer ratio and found that, pNAWM in general was not significantly correlated with generic clinical measures, such as EDSS and disease duration (Vrenken et al., 2006), but this study did not tap into the relationship between different pNAWM bands and WML. Recently, a study attempted to probe the relationship between pNAWM and WML (specifically the chronic black holes) and found a significant association between the two classes in axonal volume fraction (a multi-compartment diffusion MRI-derived metric) (Clarke et al., 2021). However, in both studies, pNAWM were not systematically and consistently delineated (i.e., lacking three-dimensional coverage and requiring expert human raters). Furthermore, WML was treated as a single region of interest (ROI) in the analyses, yet there is a significant heterogeneity in lesion subtypes (Lucchinetti et al., 2000) and evolutionary patterns. Recent

attention to paramagnetic rim lesions and the concentrically expanding patterns underscores the importance of the WML boundary (Martire et al., 2022; Preziosa et al., 2022; Sethi et al., 2017). Not long ago, a study using quantitative T1rho (i.e., a quantitative T1 metric that utilizes the MR spin lock mechanism) and DTI metrics (AD, RD, MD and FA) in pNAWM regions dilated using individual WMLs in a relapsing-remitting MS cohort has shown that those qMRI metrics not only are sensitive to the pNAWM heterogeneity, but also have different sensitivity profiles that share information with each other (Wang et al., 2025). Therefore, this study aims to adopt similar study designs but with a broader range of qMRI metrics, at finer resolution, to expand our understanding of the relationship between pNAWM and WML. In this study, individual WML surfaces will serve as the basis for region growing/shrinking, resembling a distance-based analysis. Compared to most of the prior two-dimensional analyses, this study will also use three-dimensional shells (versus 2D bands) computations both inside and outside WML, in a stepwise fashion (i.e., 1 mm^3 per step). Last but not least, the current study is conducted at the level of individual lesions.

4.3 Methods

Dataset summary

The approval of the study was gained from the University of Manitoba Health Research Ethics Board and participants provided their written consent prior to any scanning/analysis. The cross-sectional data used for this analysis entails the full battery of qMRI metrics (i.e., axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD), fractional anisotropy (FA), myelin water fraction (MWF), intra- and extra- cellular water fraction (IEWF), geometric mean T2 (gT2), T1-weighted-by-T2-weighted (T1w/T2w) ratio, R2* and quantitative susceptibility mapping (QSM)), which belongs to the follow-up dataset of the MS cohort in the Comorbidity and Cognition in

Multiple Sclerosis (CCOMS) Study MRI dataset (Uddin et al., 2022). The seventy-nine (67 are females) participants' median expanded disability status scale is 3.5 (IQR: [3,5.5]), and the mean age is 50.6 (range: 22.8 – 79.0). They were all previously diagnosed with multiple sclerosis, with conditions met under McDonald's criteria. For more details on the inclusion criteria, one may refer to our previously published imaging protocol (Uddin et al., 2022).

qMRI map pre-processing and ROI mask computation

Specific details of the image pre-processing pipelines for each qMRI technique can also be viewed in our previously published imaging protocol (Uddin et al., 2022). Binary tissue masks (i.e., global WM and WML) were computed using the same method reported in one of our recent analyses (Chapter 2). Briefly, the subject-space WM tissue probability map from each participant was and these were thresholded at 70% to exclude confounding regions with low probabilities were computed by Computational Anatomy Toolbox (CAT12 version r1318; http://www.neuro.uni-jena.de/cat/) as previously discussed in the imaging protocol for the CCOMS Study (Uddin, 2022). The binary subject-space WML tissue mask was created by selecting the overlaps of the WML masks output by two independent software packages (Uddin et al., 2022). Global WM and WML masks were then nonlinearly warped to the standard MNI152 template using the forward deformation file for each participant already included in the CCOMS dataset (Uddin, 2022), and the NAWM masks were then generated by subtracting each person's WML mask from their global WM mask. However, because of the inherently reduced field of view of the brain near the inferior cerebellar region when collecting the myelin water imaging (MWI) data, an apparent brain mask on each subject is needed to not only facilitate an accurate shell ROI (see computation below) mean calculation (i.e., virtually excluding zeros entries from the background) in MWF, IEWF, and

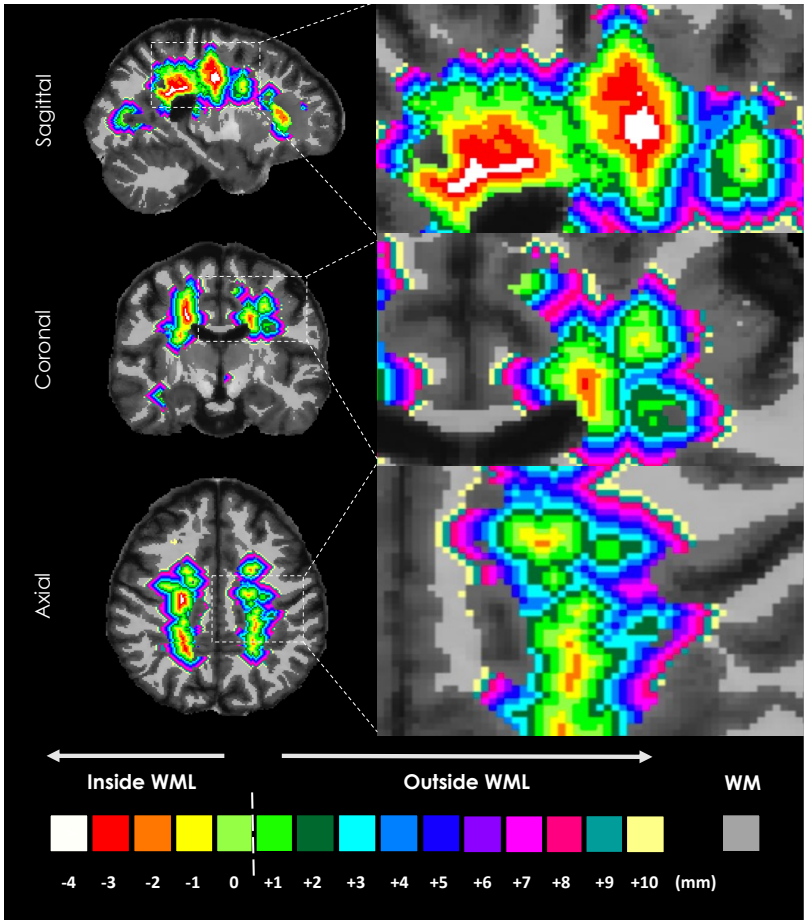
gT2, but to also exclude any separated WMLs (see below section) globally for each subject that are completely out of view. This apparent brain mask was computed by applying the 'imclose' function in MATLAB (version R2020a; The MathWorks Inc., Natick, MA, USA) with a spherical element of a radius of 10 voxels on the MWF map (Chapter 2). While both the subject and template spaces are valid approaches, the latter aligns with our recent study investigating the relationships between different qMRI metrics (Chapter 2). However, the interpretation of the results should therefore be considered using this standardized distance scale (rather than an absolute distance scale), since the same template voxel distance can correspond a different subject space voxel distance due to individual variation in brain size and local structural morphometry.

WML separation and shell computation

In an in-house script, the ROI computation was done as follows: Individual WML masks were separated using the MATLAB function 'bwlabeln' by choosing the '26' as the voxel connectivities (version R2020a; The MathWorks Inc., Natick, MA, USA). The equidistant shells of interest in width of 1 mm for each WML range from - 4 to +10 mm from the lesion boundary (3D surface). WMLs that were under 34 mm^3 (equivalent to a sphere with a radius $\leq$ 2 mm) and were outside of the apparent brain mask were excluded from the analysis. As illustrated in Figure 4-1, a total of 15 mutually exclusive shells were created using MATLAB functions 'imerode' and 'imdilate' in 1 mm increments for each separated WML. To prevent potential overlaps between pNAWM shells of any given WML with other WMLs and their associated pNAWM shells, other WMLs (including those excluded due to small sizes) and the associated pNAWM shell (i.e., from lesion core to + 10 mm) on the same subject were excluded from the pNAWM shells that were being computed at each time. That is, when each dilated shell of a WML was being computed, a composite WML

mask by the other WMLs was also dilated with the same distance. The overlapping region between the two was discarded, so that any voxel in the 'x' mm pNAWM shells of a particular WML will not be less than 'x' mm distance away from any other WMLs.

Figure 4-1: Schematic visualization of the erosion/dilation shells computations with a random MS MS participant.



They are displayed in sagittal (Top), coronal (Middle) and Axial (Bottom) views. The shells are overlaid on top of a participant's T1w/T2w ratio map. It ranges from -4 to +10 mm following the color scale shown at the bottom. WM mask is also displayed (Gray). Dash line: WML boundary. (WM: white matter; WML: white matter lesion)

Statistical analyses

All 10 qMRI values were extracted from the voxel within each eroded/dilated shell, and then averaged to produce a single value. However, due to the lesion sizes and morphologies, not all WMLs had the -4, -3, or even -2 mm shells (e.g., not large enough, not perfectly spherical). According to the violation of some distribution's normality by the Kolmogorov–Smirnov test, non-parametric tests were chosen for subsequent statistical analyses.

Firstly, to investigate how each qMRI measure changes as a function of distance from WML surfaces, every pair of adjacent shells with a distance of 1 mm was visualized using line plots, box-and-whisker plots, and tested for the statistical differences, using two-tailed Wilcoxon Rank-Sum tests with Bonferroni correction for multiple comparison (Bonferroni, 1936) with a threshold of $p_{FWE} < 0.05$ with the total of 140 tests.

Secondly, to examine how signals from the shell correlate with one another, Spearman rank correlations were computed for all possible pairs of shells for the same qMRI metric. Specifically, this analysis was performed using the 'corr' function in MATLAB (version R2020a; The MathWorks Inc., Natick, MA, USA), with the 'pairwise' option to handle cases where inner shells have missing values. In addition, as one of our previous findings when exploring the relationship between qMRI metrics across three brain ROIs (which include global NAWM and global WML masks) in MS, the significant correlations was found to have the highest mean magnitude (i.e., having the strongest correlation) between different qMRI metrics (the same ones recruited in this study) within the same ROI compared to when they are located in different ROIs (Chapter 2). Since this study has the same set of qMRI metrics, and similar ROIs, except now these ROIs were selectively computed based on the WML surfaces into finer ROIs, therefore, Spearman rank correlation analysis was also performed at inter-qMRI correlations within the same shell

(following the same programming procedure as mentioned above), to further investigate how do they relate to each other throughout the transition from WML core to outer pNAWM. All p-values for both Spearman rank correlation analyses (a total of 1725 correlations = 1050 inter-shell correlation for the ten qMRI metrics + 675 intra-shell correlation between the ten qMRI metrics) were further adjusted using the step-up process of Benjamini-Hochberg false discovery rate (BH-FDR) (Benjamini & Hochberg, 1995) with a threshold of $p_{BH\text{-}FDR} < 0.05$. The above ROI computations and analyses were done in MATLAB (version R2020a; The MathWorks Inc., Natick, MA, USA).

4.4 Results

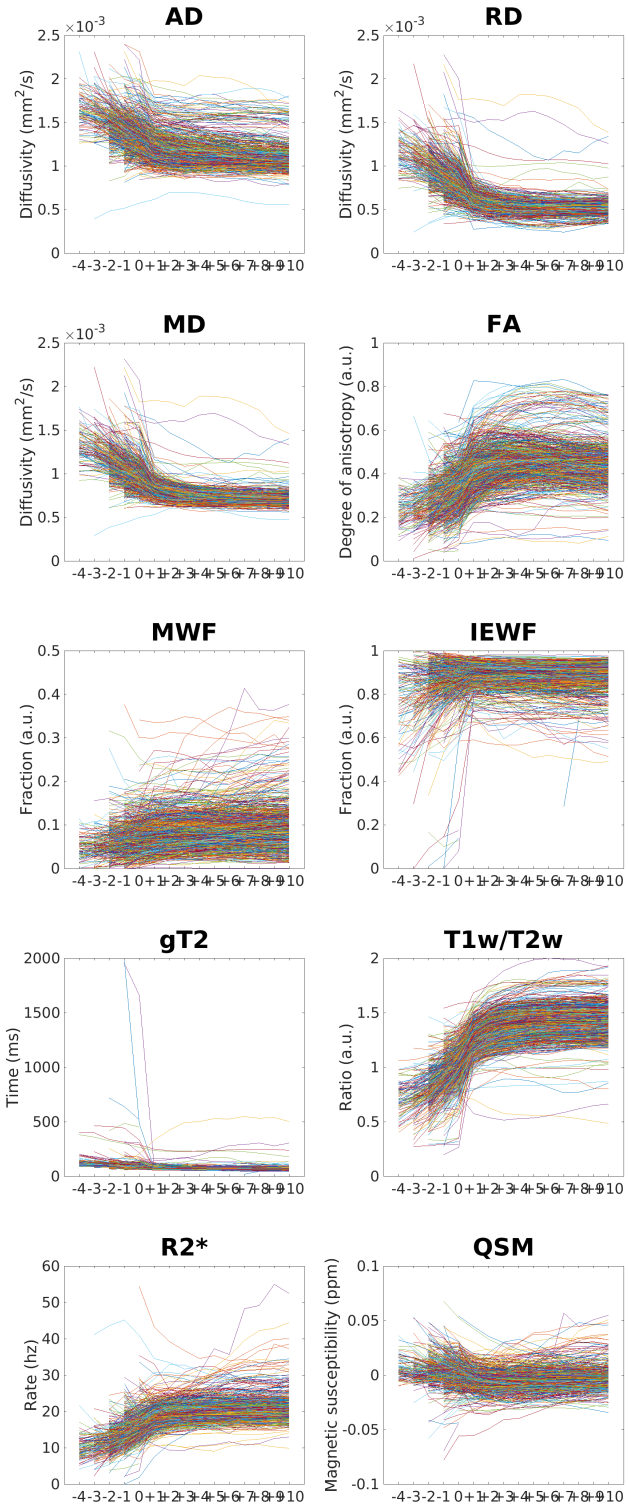
Across all 79 persons with MS, 926 individual WMLs were identified with volumes greater than 34 mm^3 and within their apparent brain mask in standardized MNI space. However, due to a combination of the constraints of the WML volume, morphometry and spatial location (e.g., immediately adjacent to cortical or subcortical gray matter), not all WMLs have the complete set of erosion shells (i.e., -4, -3, -2, -1) as quantified in Table 4-1. Signal intensities within each shell and for each metric are visualized in line plots and box-and-whisker plots in Figures 4-2 and 4-3, respectively. As they reach the pNAWM regions and extend outward, a plateau was observed in most metrics, with less heterogeneity than in the core. The result of the two-tailed Wilcoxon Rank-Sum test after Bonferroni correction indicated that 55/140 adjacent pairs of metrics remained significant. Most notably, the significant results cluster at the tissue transition from WML to pNAWM with different sensitivity across metrics. Especially at the tissue transition (i.e., the comparison between the 0 and +1 shells), all tests were significant.

The BH-FDR corrected Spearman rank correlation of the inter-shells (with the same qMRI metric) and intra-shells (with different qMRI metrics) was also visualized in Figures 4-4 and 4-5, respectively. Across inter-shell correlation analysis, all shells' correlation within the same tissue class (WML or NAWM) were strong and remained significant, while the shells' correlation between tissue classes (WML and NAWM) were relatively weaker and not all of them remained significant. In the intra-shell correlation analysis among qMRI metrics, each qMRI metric exhibited a distinct map of significant correlations with others. The majority of the direction and strength of these significant correlations remained uniform at least within the same tissue class. But for those who experienced correlation transition (i.e., a change in direction, weakening of strength), the transitions occurred in and around the WML-pNAWM boundary.

Table 4-1: The number of WMLs that have shells available at certain distances from the WML-pNAWM boundary.

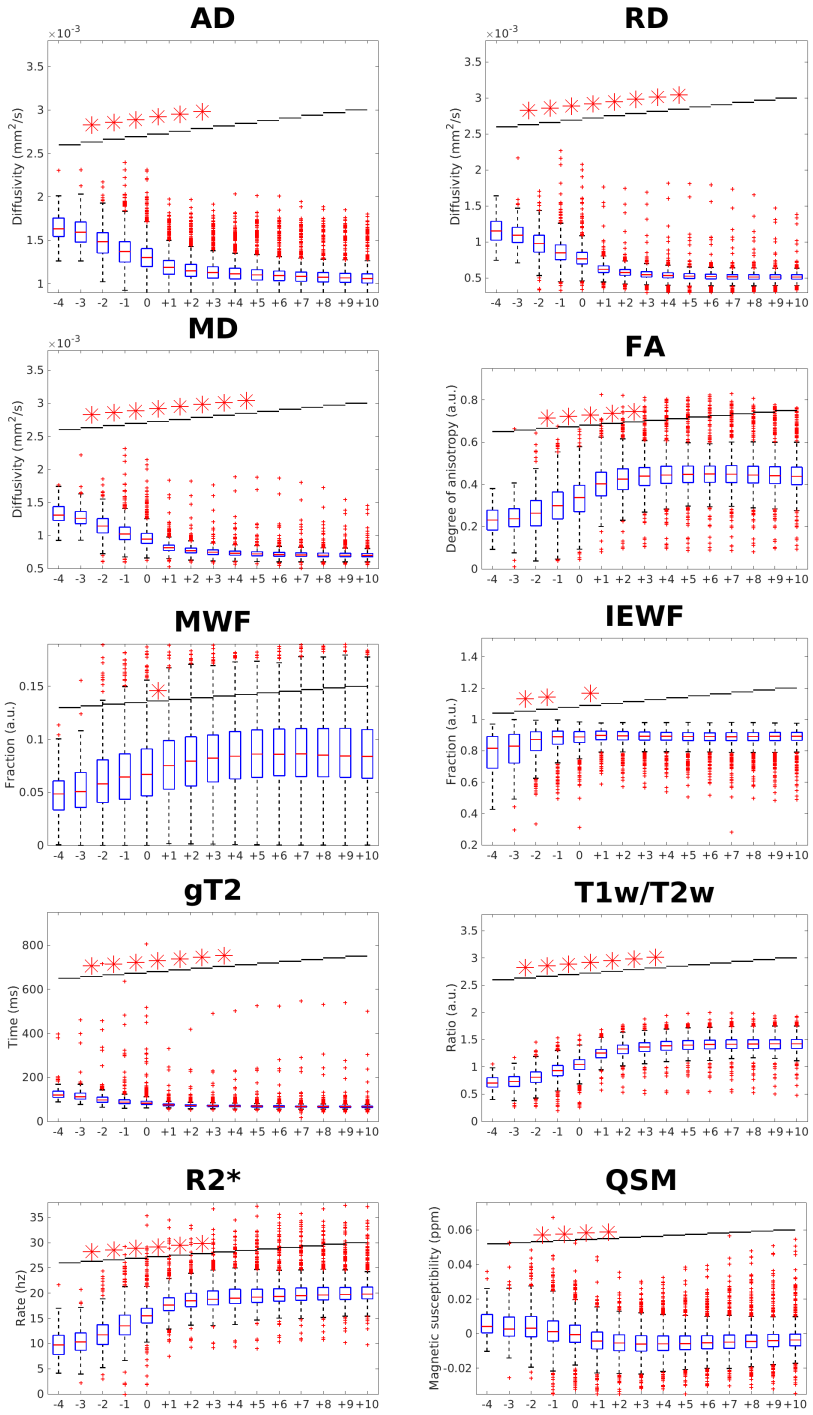
Distance (mm)	Numbers of WMLs
-4	84
-3	152
-2	404
-1	766
0	926
+1	926
+2	926
+3	926
+4	926
+5	926
+6	926
+7	926
+8	926
+9	926
+10	926

Figure 4-2: Line plots of the shell transitions for each qMRI metrics from -4 mm to + 10 mm.



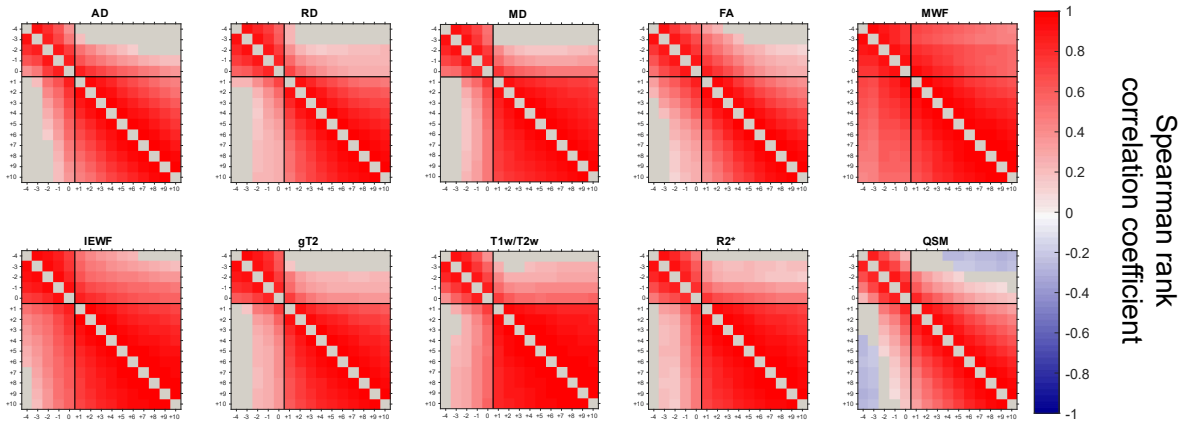
Different colors are associated with different WMLs.

Figure 4-3: The Box and whisker plots of the shell transition for different qMRI metrics from -4 mm to + 10 mm.



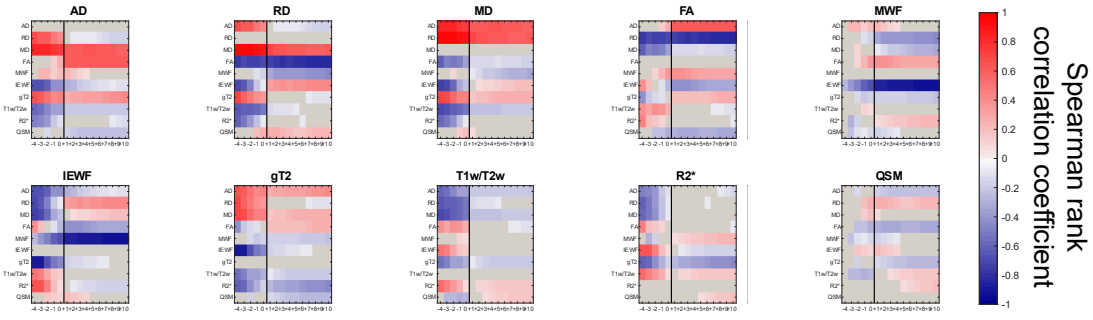
The significant results of the two-tailed Wilcoxon Rank-Sum test are indicated by the star signs on top. These results are corrected for multiple comparisons using the Bonferroni method.

Figure 4-4: The inter-shell Spearman rank correlation matrices for each qMRI metric from -4 mm to + 10 mm.



Shell distances are displayed on both horizontal and vertical axes. The shown elements are corrected by the Benjamini-Hochberg false discovery rate method.

Figure 4-5: The intra-shell Spearman rank correlation matrix between qMRI metrics at shell distances from -4 mm to + 10 mm.



The shown elements are corrected by the Benjamini-Hochberg false discovery rate method.

4.5 Discussion

The signal transition as a function of distance from the WML boundary and the signal distribution reveal several notable findings. The trend across qMRI signals (from inner WML to outer pNAWM) was a smooth and monotonic change within WMLs, with the most severe microstructural damage within and closer to the WML cores, with improving microstructural measures continuing out (in some cases several mm) into NAWM before an eventual plateau further into NAWM regions. Aside from a handful of outliers in the distribution, WMLs were overall more spatially heterogeneous than pNAWM (wider inter-quartile range in Figure 4-3), except for the case of MWF and FA. The inter-shell Spearman rank correlation analyses displayed a contrast between within-class and between-class correlations. The intra-shell Spearman rank correlation revealed information sharing between qMRI metrics throughout the shell transition, and particularly highlighted the significance of the regions around the WML-pNAWM boundary.

Transitions happen in a wider range of regions beyond the WML-pNAWM boundary

Aligned with the findings from the work by Wang et al. (Wang et al., 2025) but with shells of finer steps, all qMRI metrics can detect differences when they cross the boundary of the WML and enter the pNAWM regions, highlighting the great sensitivity of qMRI metrics. And this signal transition even extends beyond the WML-pNAWM boundary for most qMRI metrics, reaching as far as the +4 mm and +5 mm pairs (RD and MD). Therefore, the so-called “normal appearing” white matter on conventional MRI appears to be abnormal in many cases across multiple qMRI metrics – especially for white matter regions immediately adjacent to WMLs. Including those pNAWM regions may aid in better characterization of the disease activities. According to a study, NAWM is in a fine balance between neurodegenerative and neuroprotective processes (Zeis et al., 2007).

This sheds light on, for example, the escalated forms of glial cells and their significant roles in determining the fate of the lesion, especially at the periphery of WML or adjacent to pNAWM (Healy et al., 2022; Kooistra & Schirmer, 2025). In fact, studies have demonstrated the increasingly important diagnostic value (Martire et al., 2022; Montalban et al., 2025) and potential for prognostic applications (Absinta et al., 2019) of lesion rim features such as iron-rich rims indicated by the hyperintensity around lesions using susceptibility-sensitizing methods. In addition, the significance of pNAWM in measuring the level of neuroprotection and repair (Clarke et al., 2021), the apparent rim gradient transitions with qMRI measures here highlighted the potentials in expanding the scope of tissue degradation beyond WML to deepen our understanding of the disease, and specifically, differences in the transition sensitivity in different qMRI measures showcase a multi-perspective of microstructural degradation and their spatial extensions, broadening the spatial definition of rim.

WMLs have worse microstructural measures in deeper layers

In addition to the pNAWM regions, this study aims to probe the sensitivity of qMRI within WML using finer steps. As shown in Figures 4-2 and 4-3, microstructural integrity as measured by qMRI progressively worsens with increasing depth from the WML boundary toward the core. This underscores that larger WMLs' (with the deeper shells available) microstructural integrity is more severely compromised, as suggested by the qMRI metrics, in a concentric pattern by definition of the study design. Although the volumetric measure is not quantitatively shown, this partially explains why our recent finding – the WML volumetric measure is negatively correlated with microstructural measures (Chapter 3). The differences between the adjacent pairs within WML were also statistically significant, reaching as deep as the -3 mm and -2 mm pair for AD, RD, MD,

gT2, T1w/T2w ratio and R2*. But no significant differences were found between the -4 mm and -3 mm pair. In the literature on chronic lesion ultrastructure, the lesion core was found to exhibit varying degrees of demyelination and an expanded extracellular space (BARNES et al., 1991) and depending on lesion type and progression, the glial cell populations may vary as well (Kooistra & Schirmer, 2025). Another longitudinal study had investigated the changes of the blood-brain barrier patterns of enhancing lesions using dynamic enhancing MRI and found that as enhanced WML ages, it not only imposes a larger volume and more apparent ring enhancement, but the enhancing pattern of gadolinium uptake in a short period becomes shallower in the centre as compared to the periphery (Gaitán et al., 2011). Nonetheless, the contrast within WML as sensitized by qMRIs, can potentially be utilized for spatial identification or classification of lesion subtypes.

The heterogeneity in outer pNAWM is more subtle

On the other hand, as one traverses the shells toward the outer pNAWM regions (+6 mm and beyond), all qMRI measures reach a plateau with no significant differences between any adjacent pairs of shells. A similar plateau trend has also been shown in the work (quantitative T1 and magnetic transfer ratio in a 2D band-stepping fashion) by Vrenken et al. (Vrenken et al., 2006) and the closely related work (as discussed) by Wang et al. (Wang et al., 2025). However, the outer pNAWM shells can still exhibit subtle heterogeneity. The subtle heterogeneity in the distal pNAWM can be due to the subtle difference between pre-lesional NAWM (e.g., juxtacortical or distal location of the WML of interest) and stable NAWM, given prior evidence of NAWM precursor signal to WML (Elliott et al., 2021; Goodkin et al., 1998), which these qMRIs are not yet able to detect, highlighting the need to develop a more sensitive imaging marker. But it is also

worth highlighting that the region diffusely abnormal white matter (DAWM), which is radiologically characterized by the intermediate intensity between NAWM and WML on conventional MRI images in MS (Cairns et al., 2022), may offer insights into this subtle heterogeneity. Despite the fact that DAWM is not easily segmented and still poorly defined across the literature, DAWM is seen in all MS phenotypes and shows clinical potential in MS (Cairns et al., 2022). Nevertheless, given the complementarity between qMRI metrics in characterizing brain structures demonstrated by previous studies with healthy (Uddin et al., 2019) and MS cohorts (Chapter 2; Lipp et al., 2019), future work may investigate the feasibility of the multimodal approach in DAWM characterization.

Information in pNAWM and WML may complement each other

The contrast between within-class correlation and between-class correlation after thresholding for multiple comparisons has suggested the complementarity between the information held by the two tissue classes, especially given how sensitive qMRI metrics are to detecting the differences as they traverse from inner WML to outer pNAWM around the boundary. Although no clear evidence of independence in the tissue pathology has been shown so far, there is evidence of independence between two seemingly related events, such as axonal degeneration and demyelination within individual tissue classes (DeLuca, 2006; Rahmanzadeh et al., 2021; Silber & Sharief, 1999). The subtle balance between the autoimmune response and neuroprotective effect in NAWM (Zeis et al., 2007) may provide another hypothesis for the independence. Different from the rest of the inter-shell correlation metrics, negative correlations emerge at the edges (correlation between inner WML and outer pNAWM shells) of the QSM's matrix and all correlation entries were shown to be statistically significant in MWF's matrix. QSM is sensitive to both diamagnetic and

paramagnetic changes. And in MS, this can manifest as macroscopic tissue changes, such as alterations in myelin and iron concentrations (Deh et al., 2018; Y. Wang & Liu, 2015; Wisnieff et al., 2015). With a bidirectional scale, it can likely capture both positive and negative correlations with the pathological dynamics of these events. For MWF, one interpretation is that the myelin level in the CNS, especially in the MS brain, is probably changing in the same direction in bulk (repairing vs degenerating). This partially aligns with a study looking at the degenerating aspects of the longitudinal relapsing and remitting form of MS (Vavasour et al., 2018). However, MWF is sensitive not only to myelin water but also to iron (Birkl et al., 2019), let alone to the type of myelin being detected (intact vs. debris), making interpretation in this case challenging (MacKay & Laule, 2016). Nonetheless, the widely "diffused" pattern of shells for those metrics implies a global linkage of microstructural changes.

qMRI metrics are correlated in the same regions, even in finer subregions

Consistent with the findings from our previous work (Chapter 2), qMRI metrics were sharing information with each other within the same regions (even at finer divisions – shells in this context), and again the work by Wang et al. (Wang et al., 2025). Most of the correlations from previous work (Chapter 2) (when investigating the relationship between different qMRI metrics from different ROIs, in cases of NAWM and WML) reappeared (same directions) in the current work, with some exceptions likely due to the different choice of ROI (individual concentric shells vs. global masks). T1w/T2w ratio and QSM previously didn't have any significant relationship in NAWM (Chapter 2) but they are significantly correlated in this study. Despite the relatively lower strengths in those correlations, they again underpin the importance of ROI selection in affecting the result.

The significance of the regions around WML-pNAWM

Despite the distinct patterns of the correlations between qMRI metrics within the same shells, most of these patterns exhibited a gradient transition throughout the whole range of distance (-4 mm to +10 mm) (Figure 4-5). Most gradient transitions exhibit a weakening effect at and around WML-pNAWM, and some gradient transitions even have a change in direction from WML to pNAWM (Figure 4-5), indicating a lower degree of information sharing between qMRI metrics. The weakening effect in some of the correlations between qMRI metrics when going toward the outer pNAWM shells is also observed in the work by Wang et al., albeit not addressed in depth (Wang et al., 2025). Those observations reinforce the points raised in previous discussions with high relevance in regions around the WML-pNAWM boundary, suggesting a complex pathophysiology at the rim of WML in MS. Incorporating more qMRI metrics may yield better characterization, such as at the lesion rim or peripheral regions.

Study limitations

This study comes with several limitations. WMLs were separated by an automated algorithm with minimal human intervention. There will be a discrepancy between those individual WMLs, and the ones identified by human expert raters, therefore hindering the translatability of the current work to contemporary clinical settings. Also, among these WMLs, lesion subtype (e.g., acute, chronic inactive and chronic active) is not considered in the current study and requires further analysis (using additional conventional MRI sequences such as Gadolinium-enhanced T1, and collecting multiple time points) to elucidate the mixed effect.

ROI analysis was chosen to enhance the findings while minimizing noise. But the ROI size was not specifically controlled in the study. And when there is a high degree of heterogeneity in large ROIs, ROI analysis may not be the best option. Also, because of the combination of how we handle the shell computation of two neighboring lesions when they overlap (i.e., discard), and how individual WMLs were distributed in their WM space (e.g., juxtacortical, periventricular), shells associated with any two WMLs (even from the same subject and the same distance from the WML boundary), may have drastically different ROI sizes. Considering the anatomical features, given the proximity between WM and gray matter (cortical and subcortical), there is a chance (even higher with outer shells) of finding shells that are directly located adjacent to gray matter, which can contaminate ROI analysis due to potential partial volume effects. Especially, it was known in MS that WMLs are more preferentially located in juxtacortical and periventricular regions (Bouman et al., 2022; Filippi et al., 2019; Gauthier, 2023; Pirpamer et al., 2022; Pitt et al., 2022; Tonietto et al., 2023; Vaneckova et al., 2022; Vellinga et al., 2009). Nonetheless, the inherent WM masks were all thresholded at 70%, which already minimizes the impact of the partial volume effect.

4.6 Conclusions

This cross-sectional study uses the contour of the individual WML as a basis for analyzing qMRI sensitivities as a function of distance from the WML boundary, ranging from from WML core (-4 mm) to pNAWM (+10 mm). It revealed that both pNAWM and WML exhibit a gradient of moderate tissue deterioration as measured by qMRI metrics, which this effect most apparent around the WML-pNAWM boundary. The independence of microstructural activities between WML and pNAWM measured by qMRI sheds light on the untapped balance and thresholds of

neurodegeneration and the repair process over the course of the disease, especially beyond lesion territories. Future studies should consider using qMRI metrics (potentially more advanced and specific ones) to better understand the heterogeneity in pNAWM, and especially the extent of microstructural damage in regions around the WML-pNAWM boundary, thereby offering an additional degree of information to better characterize the disease and inform clinical decisions.

4.7 References

Absinta, M., Sati, P., Masuzzo, F., Nair, G., Sethi, V., Kolb, H., Ohayon, J., Wu, T., Cortese, I. C. M., & Reich, D. S. (2019). Association of Chronic Active Multiple Sclerosis Lesions With Disability In Vivo. *JAMA Neurology*, *76*(12), 1474. https://doi.org/10.1001/jamaneurol.2019.2399

Bakshi, R., Thompson, A. J., Rocca, M. A., Pelletier, D., Dousset, V., Barkhof, F., Inglese, M., Guttmann, C. R., Horsfield, M. A., & Filippi, M. (2008). MRI in multiple sclerosis: current status and future prospects. *The Lancet Neurology*, *7*(7), 615–625. https://doi.org/10.1016/S1474-4422(08)70137-6

Barkhof, F., Brück, W., De Groot, C. J. A., Bergers, E., Hulshof, S., Geurts, J., Polman, C. H., & van der Valk, P. (2003). Remyelinated Lesions in Multiple Sclerosis. *Archives of Neurology*, *60*(8), 1073. https://doi.org/10.1001/archneur.60.8.1073

BARNES, D., MUNRO, P. M. G., YOUL, B. D., PRINEAS, J. W., & MCDONALD, W. I. (1991). THE LONGSTANDING MS LESION: A QUANTITATIVE MRI AND ELECTRON MICROSCOPIC STUDY. *Brain*, *114*(3), 1271–1280. https://doi.org/10.1093/brain/114.3.1271

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Birkl, C., Birkl-Toeglhofer, A. M., Endmayr, V., Höftberger, R., Kasprian, G., Krebs, C., Haybaeck, J., & Rauscher, A. (2019). The influence of brain iron on myelin water imaging. *NeuroImage*, *199*, 545–552. https://doi.org/10.1016/j.neuroimage.2019.05.042

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni Del R Istituto Superiore Di Scienze Economiche e Commericiali Di Firenze*, *8*, 3–62.

Bouman, P. M., Strijbis, V. I., Jonkman, L. E., Hulst, H. E., Geurts, J. J., & Steenwijk, M. D. (2022). Artificial double inversion recovery images for (juxta)cortical lesion visualization in multiple sclerosis. *Multiple Sclerosis Journal*, *28*(4), 541–549. https://doi.org/10.1177/13524585211029860

Cairns, J., Vavasour, I. M., Traboulsee, A., Carruthers, R., Kolind, S. H., Li, D. K. B., Moore, G. R. W., & Laule, C. (2022). Diffusely abnormal white matter in multiple sclerosis. *Journal of Neuroimaging*, *32*(1), 5–16. https://doi.org/10.1111/jon.12945

Clarke, M. A., Lakhani, D. A., Wen, S., Gao, S., Smith, S. A., Dortch, R., Xu, J., & Bagnato, F. (2021). Perilesional neurodegenerative injury in multiple sclerosis: Relation to focal lesions and impact on disability. *Multiple Sclerosis and Related Disorders*, *49*, 102738. https://doi.org/10.1016/j.msard.2021.102738

Deh, K., Ponath, G. D., Molvi, Z., Parel, G. T., Gillen, K. M., Zhang, S., Nguyen, T. D., Spincemaille, P., Ma, Y., Gupta, A., Gauthier, S. A., Pitt, D., & Wang, Y. (2018). Magnetic susceptibility increases as diamagnetic molecules breakdown: Myelin digestion during multiple sclerosis lesion formation contributes to increase on QSM. *Journal of Magnetic Resonance Imaging*, *48*(5), 1281–1287. https://doi.org/10.1002/jmri.25997

DeLuca, G. C. (2006). The contribution of demyelination to axonal loss in multiple sclerosis. *Brain*, *129*(6), 1507–1516. https://doi.org/10.1093/brain/awl074

Eftekhari, E., Hojjat, S.-P., Vitorino, R., Carroll, T. J., Cantrell, C. G., Lee, L., Taylor, M. W., Morrow, S. A., Benhabib, H., Aviv, R. I., & Kassner, A. (2017). Normal appearing white matter permeability: a marker of inflammation and information processing speed deficit among relapsing remitting multiple sclerosis patients. *Neuroradiology*, *59*(8), 771–780. https://doi.org/10.1007/s00234-017-1862-7

Elliott, C., Momayyezsiahkal, P., Arnold, D. L., Liu, D., Ke, J., Zhu, L., Zhu, B., George, I. C., Bradley, D. P., Fisher, E., Cahir-McFarland, E., Stys, P. K., Geurts, J. J. G., Franchimont, N., Gafson, A., & Belachew, S. (2021). Abnormalities in normal-appearing white matter from which multiple sclerosis lesions arise. *Brain Communications*, *3*(3), fcab176. https://doi.org/10.1093/braincomms/fcab176

Filippi, M., Preziosa, P., Banwell, B. L., Barkhof, F., Ciccarelli, O., De Stefano, N., Geurts, J. J. G., Paul, F., Reich, D. S., Toosy, A. T., Traboulsee, A., Wattjes, M. P., Yousry, T. A., Gass, A., Lubetzki, C., Weinshenker, B. G., & Rocca, M. A. (2019). Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. In *Brain* (Vol. 142, Number 7, pp. 1858–1875). Oxford University Press. https://doi.org/10.1093/brain/awz144

Filippi, M., & Rocca, M. A. (2007). Conventional MRI in Multiple Sclerosis. *Journal of Neuroimaging*, *17*(s1). https://doi.org/10.1111/j.1552-6569.2007.00129.x

Gaitán, M. I., Shea, C. D., Evangelou, I. E., Stone, R. D., Fenton, K. M., Bielekova, B., Massacesi, L., & Reich, D. S. (2011). Evolution of the blood–brain barrier in newly forming multiple sclerosis lesions. *Annals of Neurology*, *70*(1), 22–29. https://doi.org/10.1002/ana.22472

Gallego-Delgado, P., James, R., Browne, E., Meng, J., Umashankar, S., Tan, L., Picon, C., Mazarakis, N. D., Faisal, A. A., Howell, O. W., & Reynolds, R. (2020). Neuroinflammation in the normal-appearing white matter (NAWM) of the multiple sclerosis brain causes abnormalities at the nodes of Ranvier. *PLOS Biology*, *18*(12), e3001008. https://doi.org/10.1371/journal.pbio.3001008

Gauthier, S. A. (2023). Location, location, location: myelin repair and proximity to ventricular CSF in multiple sclerosis. *Brain*, *146*(1), 10–12. https://doi.org/10.1093/brain/awac448

Goodkin, D. E., Rooney, W. D., Sloan, R., Bacchetti, P., Gee, L., Vermathen, M., Waubant, E., Abundo, M., Majumdar, S., Nelson, S., & Weiner, M. W. (1998). A serial study of new MS lesions and the white matter from which they arise. *Neurology*, *51*(6), 1689–1697. https://doi.org/10.1212/WNL.51.6.1689

Guttmann, C. R., Ahn, S. S., Hsu, L., Kikinis, R., & Jolesz, F. A. (1995). The evolution of multiple sclerosis lesions on serial MR. *AJNR. American Journal of Neuroradiology*, *16*(7), 1481–1491. http://www.ncbi.nlm.nih.gov/pubmed/7484637

Harding, K. E., Kreft, K. L., Ben-Shlomo, Y., & Robertson, N. P. (2024). Prodromal multiple sclerosis: considerations and future utility. *Journal of Neurology*, *271*(4), 2129–2140. https://doi.org/10.1007/s00415-023-12173-4

Healy, L. M., Stratton, J. A., Kuhlmann, T., & Antel, J. (2022). The role of glial cells in multiple sclerosis disease progression. *Nature Reviews Neurology*, *18*(4), 237–248. https://doi.org/10.1038/s41582-022-00624-x

Hemond, C. C., & Bakshi, R. (2018). Magnetic Resonance Imaging in Multiple Sclerosis. *Cold Spring Harbor Perspectives in Medicine*, *8*(5), a028969. https://doi.org/10.1101/cshperspect.a028969

Inglese, M., Grossman, R. I., & Filippi, M. (2005). Magnetic Resonance Imaging Monitoring of Multiple Sclerosis Lesion Evolution. *Journal of Neuroimaging*, *15*(s4), 22S-29S. https://doi.org/10.1177/1051228405282243

Kooistra, S. M., & Schirmer, L. (2025). Multiple Sclerosis: Glial Cell Diversity in Time and Space. *Glia*, *73*(3), 574–590. https://doi.org/10.1002/glia.24655

Korn, T. (2008). Pathophysiology of multiple sclerosis. In *Journal of Neurology* (Vol. 255, Number SUPPL. 6, pp. 2–6). https://doi.org/10.1007/s00415-008-6001-2

Lassmann, H. (2014). Mechanisms of white matter damage in multiple sclerosis. *Glia*, *62*(11), 1816–1830. https://doi.org/10.1002/glia.22597

Lassmann, H., Brück, W., & Lucchinetti, C. F. (2007). The Immunopathology of Multiple Sclerosis: An Overview. *Brain Pathology*, *17*(2), 210–218. https://doi.org/10.1111/j.1750-3639.2007.00064.x

Lipp, I., Jones, D. K., Bells, S., Sgarlata, E., Foster, C., Stickland, R., Davidson, A. E., Tallantyre, E. C., Robertson, N. P., Wise, R. G., & Tomassini, V. (2019). Comparing MRI metrics to quantify white matter microstructural damage in multiple sclerosis. *Human Brain Mapping*, hbm.24568. https://doi.org/10.1002/hbm.24568

Lucchinetti, C., Brück, W., Parisi, J., Scheithauer, B., Rodriguez, M., & Lassmann, H. (2000). Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. *Annals of Neurology*, *47*(6), 707–717. https://doi.org/10.1002/1531-8249(200006)47:6<707::AID-ANA3>3.0.CO;2-Q

MacKay, A. L., & Laule, C. (2016). Magnetic Resonance of Myelin Water: An in vivo Marker for Myelin. *Brain Plasticity*, *2*(1), 71–91. https://doi.org/10.3233/bpl-160033

Makhani, N., & Tremlett, H. (2021). The multiple sclerosis prodrome. *Nature Reviews Neurology*, *17*(8), 515–521. https://doi.org/10.1038/s41582-021-00519-3

Martire, M. S., Moiola, L., Rocca, M. A., Filippi, M., & Absinta, M. (2022). What is the potential of paramagnetic rim lesions as diagnostic indicators in multiple sclerosis? *Expert Review of Neurotherapeutics*, *22*(10), 829–837. https://doi.org/10.1080/14737175.2022.2143265

McDonald, W. I., MILLER, D. H., & Barnes, D. (1992). The pathological evolution of multiple sclerosis. *Neuropathology and Applied Neurobiology*, *18*(4), 319–334. https://doi.org/10.1111/j.1365-2990.1992.tb00794.x

Miller, D. H., Thompson, A. J., & Filippi, M. (2003). Magnetic resonance studies of abnormalities in the normal appearing white matter and grey matter in multiple sclerosis. *Journal of Neurology*, *250*(12), 1407–1419. https://doi.org/10.1007/s00415-003-0243-9

Moll, N. M., Rietsch, A. M., Thomas, S., Ransohoff, A. J., Lee, J., Fox, R., Chang, A., Ransohoff, R. M., & Fisher, E. (2011). Multiple sclerosis normal-appearing white matter: Pathology–imaging correlations. *Annals of Neurology*, *70*(5), 764–773. https://doi.org/10.1002/ana.22521

Montalban, X., Lebrun-Frénay, C., Oh, J., Arrambide, G., Moccia, M., Pia Amato, M., Amezcua, L., Banwell, B., Bar-Or, A., Barkhof, F., Butzkueven, H., Ciccarelli, O., Chataway, J., Cohen, J. A., Comi, G., Correale, J., Deisenhammer, F., Filippi, M., Fiol, J., … Thompson, A. J. (2025). Diagnosis of multiple sclerosis: 2024 revisions of the McDonald criteria. *The Lancet Neurology*, *24*(10), 850–865. https://doi.org/10.1016/S1474-4422(25)00270-4

Pirpamer, L., Kincses, B., Kincses, Z. T., Kiss, C., Damulina, A., Khalil, M., Stollberger, R., Schmidt, R., Enzinger, C., & Ropele, S. (2022). Periventricular magnetisation transfer

abnormalities in early multiple sclerosis. *NeuroImage: Clinical*, *34*, 103012. https://doi.org/10.1016/j.nicl.2022.103012

Pitt, D., Lo, C. H., Gauthier, S. A., Hickman, R. A., Longbrake, E., Airas, L. M., Mao-Draayer, Y., Riley, C., De Jager, P. L., Wesley, S., Boster, A., Topalli, I., Bagnato, F., Mansoor, M., Stuve, O., Kister, I., Pelletier, D., Stathopoulos, P., Dutta, R., & Lincoln, M. R. (2022). Toward Precision Phenotyping of Multiple Sclerosis. *Neurology Neuroimmunology & Neuroinflammation*, *9*(6). https://doi.org/10.1212/NXI.0000000000200025

Preziosa, P., Pagani, E., Meani, A., Moiola, L., Rodegher, M., Filippi, M., & Rocca, M. A. (2022). Slowly Expanding Lesions Predict 9-Year Multiple Sclerosis Disease Progression. *Neurology - Neuroimmunology Neuroinflammation*, *9*(2), e1139. https://doi.org/10.1212/NXI.0000000000001139

Rahmanzadeh, R., Lu, P.-J., Barakovic, M., Weigel, M., Maggi, P., Nguyen, T. D., Schiavi, S., Daducci, A., La Rosa, F., Schaedelin, S., Absinta, M., Reich, D. S., Sati, P., Wang, Y., Bach Cuadra, M., Radue, E.-W., Kuhle, J., Kappos, L., & Granziera, C. (2021). Myelin and axon pathology in multiple sclerosis assessed by myelin water and multi-shell diffusion imaging. *Brain*, *144*(6), 1684–1696. https://doi.org/10.1093/brain/awab088

Rovira, À., Auger, C., & Alonso, J. (2013). Magnetic resonance monitoring of lesion evolution in multiple sclerosis. *Therapeutic Advances in Neurological Disorders*, *6*(5), 298–310. https://doi.org/10.1177/1756285613484079

Sethi, V., Nair, G., Absinta, M., Sati, P., Venkataraman, A., Ohayon, J., Wu, T., Yang, K., Shea, C., Dewey, B. E., Cortese, I. C., & Reich, D. S. (2017). Slowly eroding lesions in multiple sclerosis. *Multiple Sclerosis Journal*, *23*(3), 464–472. https://doi.org/10.1177/1352458516655403

Silber, E., & Sharief, M. K. (1999). Axonal degeneration in the pathogenesis of multiple sclerosis. *Journal of the Neurological Sciences*, *170*(1), 11–18. https://doi.org/10.1016/S0022-510X(99)00178-1

Tonietto, M., Poirion, E., Lazzarotto, A., Ricigliano, V., Papeix, C., Bottlaender, M., Bodini, B., & Stankoff, B. (2023). Periventricular remyelination failure in multiple sclerosis: a substrate for neurodegeneration. *Brain*, *146*(1), 182–194. https://doi.org/10.1093/brain/awac334

Uddin, M. N., Figley, T. D., Kornelsen, J., Mazerolle, E. L., Helmick, C. A., O'Grady, C. B., Pirzada, S., Patel, R., Carter, S., Wong, K., Essig, M. R., Graff, L. A., Bolton, J. M., Marriott, J. J., Bernstein, C. N., Fisk, J. D., Marrie, R. A., & Figley, C. R. (2022). The comorbidity and cognition in multiple sclerosis (CCOMS) neuroimaging protocol: Study rationale, MRI acquisition, and minimal image processing pipelines. *Frontiers in Neuroimaging*, *1*. https://doi.org/10.3389/fnimg.2022.970385

Uddin, M. N., Figley, T. D., Solar, K. G., Shatil, A. S., & Figley, C. R. (2019). Comparisons between multi-component myelin water fraction, T1w/T2w ratio, and diffusion tensor imaging measures in healthy human brain structures. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-39199-x

Vaneckova, M., Piredda, G. F., Andelova, M., Krasensky, J., Uher, T., Srpova, B., Havrdova, E. K., Vodehnalova, K., Horakova, D., Hilbert, T., Maréchal, B., Fartaria, M. J., Ravano, V., & Kober, T. (2022). Periventricular gradient of T1 tissue alterations in multiple sclerosis. *NeuroImage: Clinical*, *34*, 103009. https://doi.org/10.1016/j.nicl.2022.103009

Vavasour, I. M., Huijskens, S. C., Li, D. K., Traboulsee, A. L., Mädler, B., Kolind, S. H., Rauscher, A., Moore, G. W., MacKay, A. L., & Laule, C. (2018). Global loss of myelin water over

5 years in multiple sclerosis normal-appearing white matter. *Multiple Sclerosis Journal*, *24*(12), 1557–1568. https://doi.org/10.1177/1352458517723717

Vellinga, M. M., Geurts, J. J. G., Rostrup, E., Uitdehaag, B. M. J., Polman, C. H., Barkhof, F., & Vrenken, H. (2009). Clinical correlations of brain lesion distribution in multiple sclerosis. *Journal of Magnetic Resonance Imaging*, *29*(4), 768–773. https://doi.org/10.1002/jmri.21679

Vrenken, H., Geurts, J. J. G., Knol, D. L., Polman, C. H., Castelijns, J. A., Pouwels, P. J. W., & Barkhof, F. (2006). Normal-appearing white matter changes vary with distance to lesions in multiple sclerosis. *AJNR. American Journal of Neuroradiology*, *27*(9), 2005–2011. http://www.ncbi.nlm.nih.gov/pubmed/17032884

Wang, L., So, T. Y., Choi, J. C. H., Lau, A. Y. L., Yu, Z., Yeung, D. K. W., Abrigo, J., King, A. D., Wáng, Y. X. J., Ai, Q. H., & Chen, W. (2025). Evaluation of the perilesional normal-appearing white matter (NAWM) microenvironment in relapsing-remitting multiple sclerosis with quantitative T1rho magnetic resonance imaging. *Quantitative Imaging in Medicine and Surgery*, *15*(10), 9479–9491. https://doi.org/10.21037/qims-2025-1075

Wang, Y., & Liu, T. (2015). Quantitative susceptibility mapping (QSM): Decoding MRI data for a tissue magnetic biomarker. *Magnetic Resonance in Medicine*, *73*(1), 82–101. https://doi.org/10.1002/mrm.25358

Wisnieff, C., Ramanan, S., Olesik, J., Gauthier, S., Wang, Y., & Pitt, D. (2015). Quantitative susceptibility mapping (QSM) of white matter multiple sclerosis lesions: Interpreting positive susceptibility and the presence of iron. *Magnetic Resonance in Medicine*, *74*(2), 564–570. https://doi.org/10.1002/mrm.25420

Zeis, T., Graumann, U., Reynolds, R., & Schaeren-Wiemers, N. (2007). Normal-appearing white matter in multiple sclerosis is in a subtle balance between inflammation and neuroprotection. *Brain*, *131*(1), 288–303. https://doi.org/10.1093/brain/awm291

Chapter 5

Comparing baseline qMRI metrics in stable & transitioning regions in and around multiple sclerosis white matter lesions

Publication status

Parts of this chapter were presented in the abstract entitled '*Comparing Baseline qMRI Measures in Stable and Transitioning Regions in/around MS White Matter Lesions'* at the conferences: **Organization of Human Brain Mapping** and **Canadian Organization of Medical Physicists Annual Scientific Meeting** in 2025. The manuscript is currently under internal review before submission to a peer-reviewed journal.

Peer-review conference abstract

Wong, K., Uddin, M. N., Figley, T., Kornelsen, J., Fisk, J., Marrie, R. A., & Figley, C. (2025). Comparing Baseline qMRI Measures in Stable & Transitioning Regions in MS White Matter Lesions. *MEDICAL PHYSICS*, *52*(8), 57.

5.1 Abstract

White matter lesions (WMLs) evolve over the course of multiple sclerosis (MS), and their radiological presentation is important for diagnosis and monitoring. Therefore, this study aimed to investigate whether baseline quantitative MRI (qMRI) metrics differed between stable and transitioning regions in and around WMLs over the subsequent two years. Axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD), fractional anisotropy (FA), myelin water fraction (MWF), intra- and extra- cellular water fraction (IEWF), geometric mean T2 (gT2), T1-weighted-

by-T2-weighted (T1w/T2w) ratio were available for the full cohort (N=90) and R2* relaxation rate (R2*) and quantitative susceptibility mapping (QSM) were available for the partial cohort (N=35). The analysis was conducted at the individual lesion level. By associating individual WML masks from two time points, three classes of white matter regions were identified: 1) primary WML regions: a) newly appearing, b) vanishing, and c) enduring, 2) subregions of the enduring WML: a) expanding, b) contracting, and c) persisting), and perilesional normal appearing white matter (pNAWM) (i.e., 2mm dilated shells from WML mask): a) transitioning and b) stable, yielding a total of eight regions to extract signal from the qMRI maps. Statistical comparisons (two-tailed Wilcoxon Rank-Sum and Signed-Rank tests) were conducted among primary WML regions/subregions, as well as between regions that underwent similar tissue transitions or had identical tissue class labels at baseline or follow-up, to test the sensitivity of baseline qMRI metrics. Across the full (partial) MS cohort, 2123(711) unique cortical WMLs were identified, with 623(252) newly appearing WMLs, 519(169) vanishing WMLs, and 978(350) enduring WMLs. Of 110 tests, 86 were found to be significant after multiple-comparison correction. In addition to statistical significance, these differences were directionally meaningful according to the contemporary understanding of each qMRI metric. The demonstrated sensitivity of baseline qMRI measures to WM heterogeneity across these regions suggests that they can be used to predict the morphological evolution of WMLs.

5.2 Introduction

Multiple sclerosis (MS) is the most prevalent demyelinating disease (Love, 2006) with diverse clinical presentations (Lublin et al., 2014; McGinley et al., 2021) Conventional magnetic resonance imaging (MRI), provides a rough map of demyelination and neurodegeneration in MS, and plays a pivotal role in clinical aspects such as diagnosis and disease monitoring (Thompson et al., 2018; Wattjes et al., 2021), but the findings do not provide sufficient information to facilitate early personalized interventions (Pozzilli et al., 2023). In the white matter (WM), MS lesions vary in type, stage and spatial location (Absinta et al., 2016; Filippi et al., 2019), and in many cases worsen alongside symptoms (Meca-Lallana et al., 2021). With the advancement of studies and the development of MRI sequences in recent years, lesions such as paramagnetic rim lesions and the central vein sign have garnered more attention from clinicians, potentially improving the diagnostic criteria for MS (Hemond et al., 2024; Martire et al., 2022; Rimkus et al., 2024; Sati et al., 2016). Although MRI can stratify lesions into categories and quantify macroscopic changes in WML, its prognostic value remains underexploited (Rocca et al., 2024). Follow-up appointments and MRI scans are required to monitor progression and reassess treatment response in people with MS, typically at intervals of several months (e.g., six months). Nonetheless, lesion evolution can occur, involving a mix of degeneration and neural protective and/or recovery responses (Mey et al., 2023; Ricigliano et al., 2022), and therapeutic interventions are often delayed until a clinical relapse with new or worsening symptoms.

Provided MRI's maturity in MS, the sensitivity to microstructural integrity makes quantitative MRI (qMRI) a potentially worthy candidate for MS prognosis (Granziera et al., 2021; Rocca et al., 2024; Tranfa et al., 2022). However, despite the potential demonstrated by qMRI in MS, it is not ready for widespread clinical implementation (Manfrini et al., 2021; Weingärtner et

al., 2022). One of the major obstacles qMRI faces here is the lack of clinically accepted interpretability and clinical relevance (Tranfa et al., 2022), including the open question of whether qMRI methods can potentially indicate tissue deterioration before new WMLs appear. (Dousset et al., 1998; Silver et al., 1998). Several studies using various qMRI methods have shown that measurable differences in normal appearing white matter (NAWM) can actually happen, ranging from months up to two years (Boaventura et al., 2022; Elliott et al., 2021; Gloor et al., 2023; Liu et al., 2012; Pike et al., 2000) before the appearance of a new WML in conventional MRI settings. One recent study has also provided strong evidence of tissue heterogeneity and suggested a finer categorization of WM using a mix of diffusion-based metrics (i.e., neurite orientation dispersion and density imaging in conjunction with conventional diffusion tensor imaging) (Caranova et al., 2025). However, most prelesional tissue monitoring studies require a significant amount of time and attention from expert human raters to identify and longitudinally associate lesions, which marks a significant barrier to implementing these methods in clinical settings. In addition, both human-induced and method-induced biases (e.g., the choice of qMRI metric) have led to inconsistent findings. Given the motivating findings and the gap in current literature, this study aims to employ longitudinal qMRI scans from a cohort of persons with MS, along with fully automated methods (without rater or human intervention in lesion segmentation and labelling), to investigate whether baseline qMRI metrics differed in the stable and transitioning regions in and around WMLs over the subsequent 2 years.

5.3 Methods

Dataset summary

The approval of the study was gained from the University of Manitoba Health Research Ethics Board and participants provided their written consent prior to any scanning/analysis. These participates belong to the MS cohort (baseline and 2-year follow-up visits) of the Comorbidity and Cognition in Multiple Sclerosis (CCOMS) Study (Uddin et al., 2022). The participants were all previously diagnosed with multiple sclerosis, with conditions met under McDonald's criteria. For more details on the inclusion criteria, one may refer to our previously published imaging protocol (Uddin et al., 2022). This study includes a battery of qMRI metrics. Axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD), fractional anisotropy (FA), myelin water fraction (MWF), intra- and extra-cellular water fraction (IEWF), geometric mean T2 (gT2), and T1-weighted-by-T2-weighted (T1w/T2w) ratio were available for the full cohort; R2* relaxation rate (R2*) and quantitative susceptibility mapping (QSM) were only available for the partial cohort. Therefore, this study consists of two levels of analysis with different statistical power. Table 5-1 summarizes the demographic information for participants in each cohort.

Table 5-1: Demographics of participant from the two levels of analysis.

	Full Cohort	**Partial Cohort**
Quantitative MRI techniques (Metrics)	Diffusion tensor imaging (**AD, RD, MD, FA**), **T1w/T2w ratio** Myelin water imaging (**MWF, IEWF, gT2**)	Iron mapping (**R2*, QSM**)
Cohort size (Female)	90 (75)	35 (28)
Mean age (Std)	48.98 (12.62)	45.50 (10.89)
Median baseline EDSS (IQR)	3.5 (2.0) (2.0 – 4.0)	3.0 (1.9) (2.0 – 3.9)
Median change in EDSS (IQR)	0.5 (1.0) (0.0 – 1.0)	0.5 (1.5) (0.0 – 1.5)

qMRI map pre-processing and ROI mask computation

The region of interest (ROI) level analysis is chosen for this particular study design. The primary tissue masks (i.e., global WM, WML and NAWM), from both time points were computed using the same methodology as in our previous work (Chapter 2; Chapter 3; Chapter 4). It should be noted that because of the inherently reduced field of view of the brain near the inferior cerebellar region when collecting the myelin water imaging (MWI) data, an apparent brain mask on each subject is needed to not only facilitate an accurate ROI mean calculation (i.e., virtually excluding zeros entries from the background) in MWF, IEWF, and gT2, but to also exclude any associated WMLs from both time points (see below section) globally for each subject that are completely out of view. This apparent brain mask was computed by applying the 'imclose' function in MATLAB (version R2020a; The MathWorks Inc., Natick, MA, USA) with a spherical element of a radius of 10 voxels on the MWF map (Chapter 2). The quantitative maps and brain masks were first coregistered and normalized to the same MNI-152 template space using individual lesion-filled T1-weighted images at each time point, following the coregistration and nonlinear normalization parameters outlined in the protocol paper (Uddin et al., 2022). In this way, any brain movement between sequences is controlled, and potential brain atrophy over the two-year window (Andravizou et al., 2019; Bermel & Bakshi, 2006) will at least partly be corrected by using a common stereotactic space (i.e., MNI152 template).

Lesion separation for lesion-level analysis

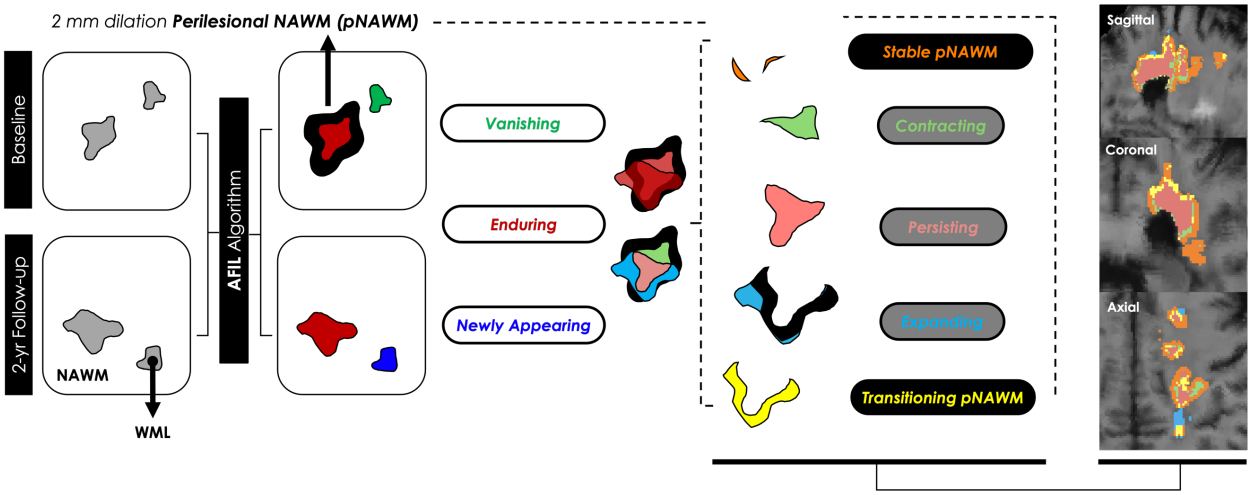
The whole-brain WML masks at both time points from each participant were first overlapped with the apparent brain mask from both time points to exclude any out-of-field portion of WML, and they were evaluated at the individual lesion level using the Automatic Follow-up of Individual Lesion (AFIL) algorithm (Köhler et al., 2019) with default settings. In brief, this algorithm first

three-dimensionally splits the lesions with 6-voxel connectivity. It then overlaps the lesions from each time point to associate them with a unified numbering system for analysis, based on their spatiotemporal overlap (e.g., a confluent lesion in which two or more separate lesions grow into one single lesion and vice versa for the separated lesions with the reversed scenario). Using the default settings, small WMLs (i.e., < 8.25 mm^3) were also neglected from the original WML masks to minimize statistical uncertainties introduced to the study (Köhler et al., 2019).

Transitioning and stable WM ROIs definitions

After individual WMLs were automatically identified across the two time points, they were defined into eight stable and transitioning WM ROIs. First, the three primary WML regions signifying the global transition patterns were defined as: 1) newly appearing WML, 2) vanishing WML and 3) enduring WML. Second, three sub-regions of enduring WML regions signifying local transition patterns were further split into: 1) expanding WML, 2) contracting WML and 3) persisting WML. Lastly, a 2-voxel wide shell is computed by dilating the enduring WML in the baseline to help define: 1) stable and 2) transitioning perilesional NAWM (pNAWM), by excluding or overlapping the "expanding" subregions of the enduring WML, respectively. With that, all transition patterns between the two classes (i.e., NAWM and WML) were systematically captured. Figure 5-1 illustrates how each of the eight regions was defined.

Figure 5-1: A simplified workflow of obtaining the eight ROIs for current study.



The primary category of ROIs (i.e., naming regions with white backgrounds) are direct outputs from the AFIL algorithm. The ROIs from the secondary category (i.e., naming regions with gray and black backgrounds) are defined by overlapping longitudinally associated (using AFIL algorithm) WMLs from both time points (baseline and 2-year follow-up). The ROIs from the secondary category are displayed on a random MS participant.

Statistical analyses

The eight WM ROIs were used to compute the average values for each baseline qMRI metric. Then, intra-metric comparisons were conducted at two levels: all possible pairs within ROI groups (i.e., primary, secondary and pNAWM groups) and selected pairs across ROI groups (e.g., primary vs secondary regions). The only across-group comparison is selected for regions that undergo the same tissue transitions or have identical tissue class labels at baseline or follow-up. All comparisons in the study were outlined and rationalized in Table 5-2. Statistical tests involving primary WML types (with different sample sizes – i.e., different numbers of WMLs for each type) were performed using unpaired Wilcoxon Rank-Sum tests. In contrast, tests between enduring

WML subregions and NAWM regions (with paired samples for the same lesions) were performed using paired Wilcoxon Signed-Rank tests. Because this is an exploratory analysis, all tests remained two-tailed, and the statistical probabilities were corrected using the Benjamini-Hochberg (BH-FDR) method (Benjamini & Hochberg, 1995), with a threshold of $p < 0.05$.

Table 5-2: The descriptions and a priori rationales for each ROI-ROI comparison and the corresponding statistical tests.

Intra-/Inter-group comparison (Primary/Secondary)	**ROI pairs (baseline tissue class)**	**Rationale**	**Statistical test (Unpaired/Paired)**
Intra-group (Primary)	Vanishing WML (WML) vs. Enduring WML (WML)	Investigates the ROI pairs within the primary group of WML regions	Wilcoxon's Rank-Sum (Unpaired)
Intra-group (Primary)	Enduring WML (WML) vs. Newly Appearing WML (NAWM)		Wilcoxon's Rank-Sum (Unpaired)
Intra-group (Primary)	Newly Appearing WML (NAWM) vs. Vanishing WML (WML)		Wilcoxon's Rank-Sum (Unpaired)
Intra-group (Secondary)	Contracting WML (WML) vs. Persisting WML (WML)	Investigates the ROI pairs within the secondary group of WML regions	Wilcoxon's Signed-Rank (Paired)
Intra-group (Secondary)	Persisting WML (WML) vs. Expanding WML (NAWM)		Wilcoxon's Signed-Rank (Paired)
Intra-group (Secondary)	Expanding WML (NAWM) vs. Contracting WML (WML)		Wilcoxon's Signed-Rank (Paired)

Intra-group (Secondary)	Stable pNAWM (NAWM) vs. Transitioning pNAWM (NAWM)	Investigates the ROI pair within the secondary regions that are controlled by dilating 2 mm away from the baseline WML surface	Wilcoxon's Signed-Rank (Paired)
Intra-group (Secondary)	Stable pNAWM (NAWM) vs. Contracting WML (WML)	Investigates the ROI pair between the transitioning secondary region as a whole and the stable pNAWM region	Wilcoxon's Signed-Rank (Paired)
Intra-group (Secondary)	Stable pNAWM (NAWM) vs. Expanding WML (NAWM)		Wilcoxon's Signed-Rank (Paired)
Inter-group (Primary vs. Secondary)	Vanishing WML (WML) vs. Contracting WML (WML)	Investigates the ROI pairs from primary and secondary groups that underwent the same tissue transition	Wilcoxon's Rank-Sum (Unpaired)
Inter-group (Primary vs. Secondary)	Newly Appearing WML (NAWM) vs. Expanding WML (NAWM)		Wilcoxon's Rank-Sum (Unpaired)

5.4 Results

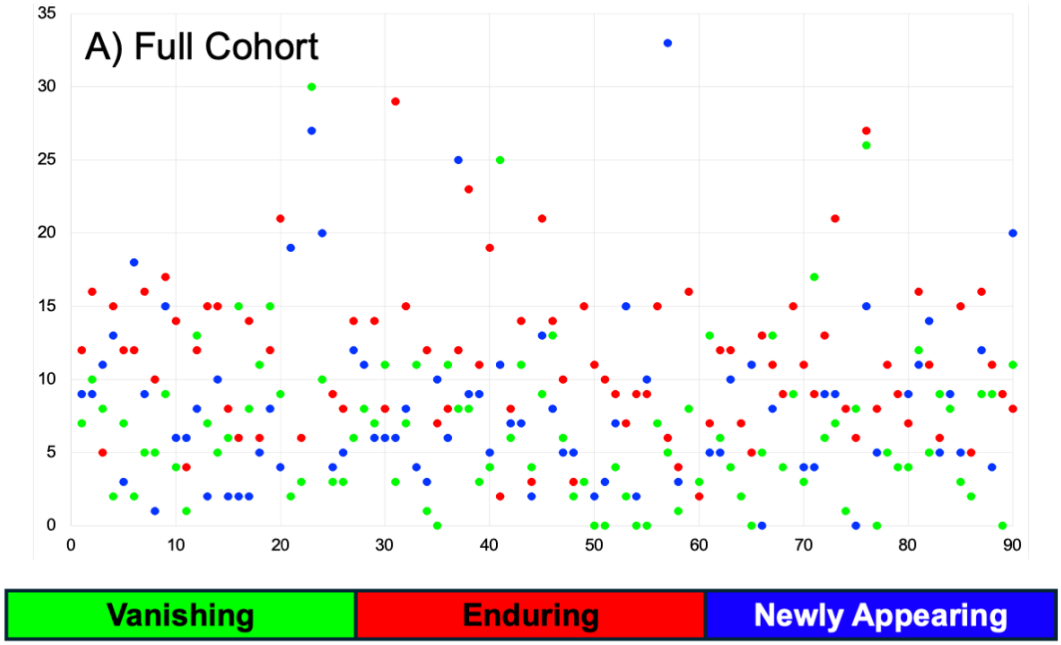
During the process of associating WMLs using the AFIL algorithm, 22910 WMLs across both time points fell under the small lesion criteria (i.e., < 8.25 mm^3) and were therefore discarded. Incidentally, one purely expanded enduring WML was found and excluded from the study (i.e., this WML has no shrinking WML region). However, none of the enduring WMLs were found to be completely unchanged or purely shrunken. Finally, a total of 2347 and 849 unique WMLs were identified across both time points for the full and partial cohorts, respectively. The number of WML

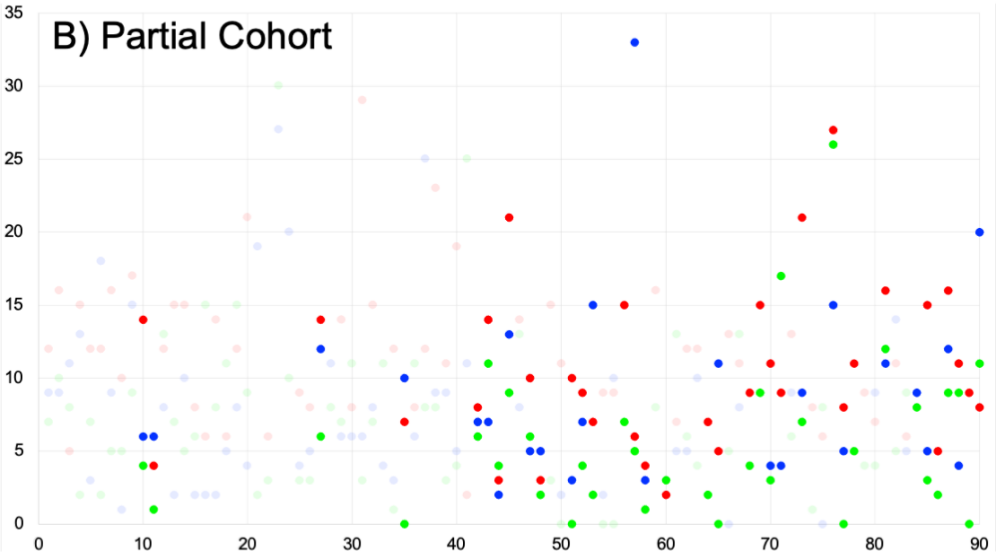
in the primary regions is summarized in Table 5-3. The breakdown of the WML in the primary regions for each participant is further summarized in Figure 5-2. Figure 5-3 shows the box-and-whisker plots for the ROIs for each qMRI metric, and the statistical comparison results. 87/110 (79%) of the comparisons in this study remain significant ($p < 0.05$) after correction by BH-FDR. BH-FDR corrected p-values are summarized in Table 5-4.

Table 5-3: A summary of WMLs subtypes after being separated and identified.

	Full Cohort	**Partial Cohort**
Total # of WML	2347	849
Total # of Vanishing WML (%)	733 (31%)	289 (34%)
Total # of Enduring WML (%)	1012 (43%)	362 (43%)
Total # of Newly Appearing WML (%)	602 (26%)	198 (23%)

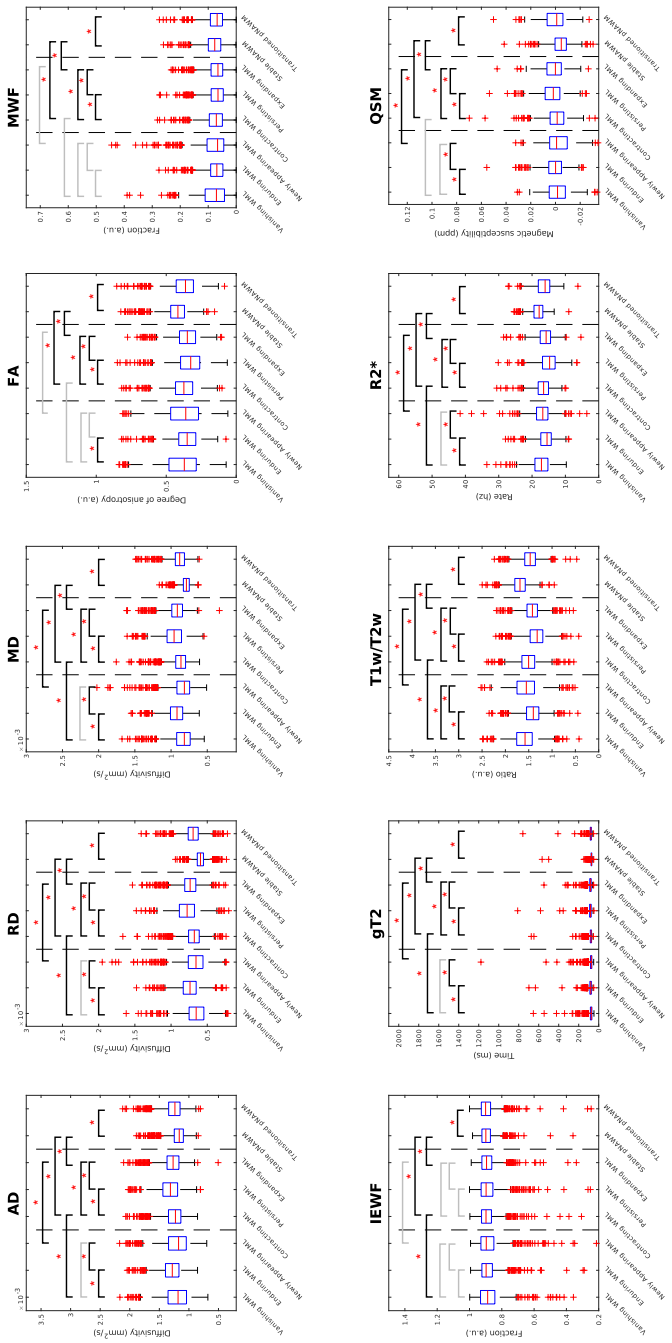
Figure 5-2: Scatter plots of the Vanishing, Enduring and Newly Appearing WMLs for each MS participant in full and partial cohort.





The number (y-axis) of the different WMLs is displayed for each MS participant (x-axis) of the A) full and B) partial cohort. (The full cohort data is overlaid on top of the partial cohort data to enable qualitative comparison).

Figure 5-3: Box-and-whisker plots of the 10 qMRI metrics from the eight WML regions defined in the current study.



Significant results after BH-FDR correction are indicated by asterisks overlaid on top of the pairs of ROI comparison. Grayed brackets indicated the statistical test is not significant.

Table 5-4: BH-FDR corrected p-values for all ROI pair comparisons of 10 qMRI metrics

ROI pair comparison	AD	RD	MD	FA	MWF	IEWF	gT2	T1w/T2w	R2*	QSM
Vanishing WML - Enduring WML	****	****	****	**	Not Sig.	Not Sig.	****	****	****	*
Enduring WML - Newly Appearing WML	****	****	****	Not Sig.	Not Sig.	Not Sig.	****	****	****	*
Newly Appearing WML - Vanishing WML	Not Sig.	Not Sig.	Not Sig.	Not Sig.	Not Sig.	Not Sig.	Not Sig.	*	Not Sig.	Not Sig.
Contracting WML - Persisting WML	****	****	****	****	****	Not Sig.	****	****	****	****
Persisting WML - Expanding WML	****	****	****	****	*	Not Sig.	***	****	****	****
Expanding WML - Contracting WML	****	****	****	****	****	Not Sig.	****	****	****	***
Stable pNAWM - Transitioning pNAWM	****	****	****	****	****	*	****	****	****	****
Stable pNAWM - Expanding pNAWM	****	****	****	****	****	***	****	****	****	****
Stable pNAWM - Contracting pNAWM	****	****	****	****	****	*	****	****	****	****
Vanishing WML - Contracting WML	****	***	****	Not Sig.	Not Sig.	*	****	****	*	Not Sig.
Newly Appearing WML - Expanding WML	****	****	****	Not Sig.	Not Sig.	Not Sig.	****	****	****	*

: <0.05, **:<0.001, *:<0.0001, ****:<0.00001, Not Sig.: Not Significant*

5.5 Discussion

Overall, more than three-quarters of the tested comparisons showed significant differences, even after the BH-FDR correction. The large number of significant results indicates that, at least at the group level, baseline qMRI metrics may capture the transition information about the WM landscape in MS over 2 years. However, it is worth reminding readers that these two-tailed statistical tests do not assume the direction of the result per se. But directional inference is possible by placing the statistically significant result in the context of the data distribution in the box-and-whisker plot (Figure 5-3).

Stable WML regions have shown the worst baseline microstructural measure, and stable pNAWM has shown the most normal looking microstructural measure

By glancing through the distributions across the WM regions, the stable regions from all qMRI metrics were most likely to be found at both ends of the spectrum, potentially highlighting how integrity and stability were "preserved" in their microstructural status. The comparisons between enduring WML and newly appearing WML, and vanishing WML, respectively, are consistently shown to be significant in AD, RD, MD, gT2, T1w/T2w ratio, R2* and QSM, except for FA (where the comparison between newly appearing WML and enduring WML is not significant). Similar findings have also been shown by comparing persisting WML (i.e., the stable sub-region of the enduring WML), and contracting and expanding WML, across all qMRI metrics, except for IEWF. At the opposite end of the spectrum, all qMRI metrics are sensitive to differences between stable pNAWM, and each of the other three ROIs: transitioning pNAWM, contracting WML and expanding WML. Higher/lower diffusivity, lower/higher MWF, higher/lower gT2, and lower/higher T1w/T2w ratio generally suggest a worse/better microstructural status of the tissue, and that aligns with all of our findings here.

Early signs of transition patterns in the baseline WM subregions

The fact that almost all contracting and expanding WMLs have a higher microstructural status than the persisting WML is not only a sign of heterogeneity among tissue classes in the WM landscape, but also a sign that heterogeneity is possibly an indicator of tissue transition. After carefully inspecting the directionality of the test between contracting and expanding WML, the authors admittedly did not expect to observe that contracting WML has a slightly better microstructural status than the expanding WML region (e.g., lower MD, higher MWF, higher T1w/T2w ratio). It is fascinating to see that the WML region directly attached to the enduring WML eventually healed at a later time already showed an early sign of recovery at baseline, in comparison to the NAWM

region (i.e., expanding WML), which undergoes neural degeneration later. With the same token (but from the NAWM end), transitioning pNAWM, as suggested by all qMRI metrics, has shown early signs of degeneration. In fact, many qMRI studies, including T1w/T2w ratio (Boaventura et al., 2022), MTR (Elliott et al., 2021; Gloor et al., 2023; Pike et al., 2000), and diffusion metrics (de Groot et al., 2013; Elliott et al., 2021; Gloor et al., 2023; Liu et al., 2012), have already shown pre-lesional areas exhibited an early sign of degeneration as compared to baseline NAWM areas. Given the fact that the two pNAWM regions in the current study have fairly similar baseline characteristics (i.e., being from the same tissue class at baseline and sharing a common concentric perilesional shell region), this finding has enhanced the spatial sensitivity of qMRI metrics for detecting early WM changes. Moreover, the same set of qMRI metrics has already demonstrated high sensitivity to the periphery of the WML particularly when using a concentric shell pattern as a basis, even extending toward the NAWM space in our recent work (Chapter 4). Together, it indicates another level of tissue heterogeneity as sensitized by these qMRI metrics in WM.

Subtle difference between regions that undergoes the same transitions

Of the 14 significant comparisons between regions that followed the same transition pattern, seven of which are from comparisons between newly appearing and expanding WMLs, and seven of which are from comparisons between vanishing and contracting WMLs. Though not all results are significant, this shows a difference between WML regions that underwent whole-lesion transition (i.e., newly appearing and vanishing) and those that underwent partial-lesion transition (i.e., expanding and contracting). The regions that underwent the partial-lesion transition also appeared to be less variable (i.e., narrower IQRs) and have worse microstructural status as indicated by the qMRI measures (e.g., higher diffusivity, lower T1w/T2w ratio). Given that those can be pairs of

distinct events that had different clinical interpretations and outcomes (not just directional changes in volume), this underscores the incredible sensitivity of qMRI metrics in WM in general.

Shared heterogeneity between NAWM and WML

Nevertheless, qMRI metrics were not quite sensitive to the difference between newly appearing and vanishing WML, with a significant result only shown in T1w/T2w ratio. The lack of a positive result has shown the WM heterogeneity and dynamics at the WML-NAWM level that are not yet captured by qMRI metrics. Despite the differences in tissue class at baseline (e.g., newly appearing WML: NAWM, vanishing WML: WML), where one would expect WML to have a worse microstructural status than NAWM as reflected in qMRI, the recovery process of vanishing WML and degeneration of WM in the newly appearing WML region are possibly bringing the microstructural integrity to a transient state where they are indistinguishable, let alone those regions can possibly show focal heterogeneity like enduring WML, which cannot be probe using current study design.

Study limitations

The positive findings from a two-year follow-up design demonstrate qMRI's sensitivity to WM heterogeneity which also inform future tissue transition. While it may be useful for prognosis, helping to minimize scan frequency in clinical settings, this study presented a few key limitations. Two years in MS is a long duration that manifests noticeable macroscopic tissue evolution in a larger region of interest, subtle lesion radiological transition can occur at rates as frequent as weeks (Guttmann et al., 1995), let alone the ongoing underlying pathophysiological activities. In validation study designs in the future, more frequent follow-up (wherever possible), and longer

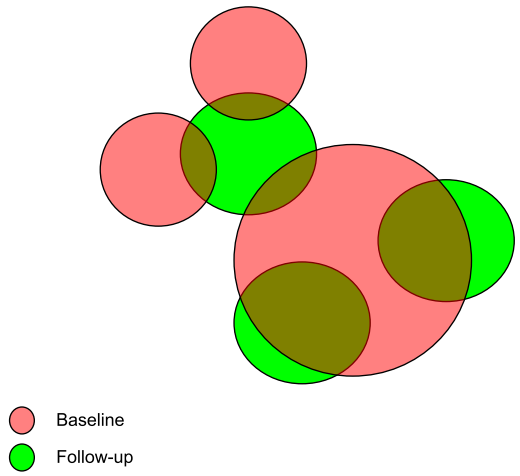
study durations (e.g., > 5 years) are required to promote a systematic understanding of long-term WM changes.

Another limitation stems from how WMLs were longitudinally separated and associated. Current literature lacks consensus or a rationale regarding the "correct" parameter for connectedness. 6-connectivity is the most conservative, but there are no fatal flaws in using it compared to the more liberal settings of 18 and 26, other than having lesions being smaller (i.e., lesion voxels are less likely to be connected), and ROI analysis will help mitigate the discrepancy that it will introduce into the study (i.e., summarizing the result and reducing the number of entries). Associating lesions from two time points is not as simple as one thinks, as there are cases of confluence (i.e., grouping of two or more lesions in the follow-up) and separation (i.e., splitting of one lesion into two or more lesions), let alone cases when lesions shifted from one place to another without overlap (which is treated as vanishing and newly appearing lesions by definition, but current study is not capable of verifying this). Note that emerging approaches aim to address visually apparent confluent lesions (Dworkin et al., 2018; Lou et al., 2021; Wynen et al., 2024, 2025). However, despite appealing results, these methods present shortcomings, such as relatively high computational requirements, human bias in rating, and reliance on assumptions (e.g., a fuzzy definition of confluent lesions). In addition, the lack of longitudinal validation of these methods makes connectedness a more naïve method for this study design. On the other hand, a limitation identified in the original toolbox (Köhler et al., 2019) is that a strict condition was imposed that separated lesions that cannot be confluent at the same time. One example is illustrated in Figure 5-4. Although it was limited by the algorithm, the authors who developed the toolbox did not discuss this specifically and this imposes an inherent ambiguity in global labelling when the WML spatial configuration is complex, especially in a 3-dimensional space with drastic evolution

patterns over a long period and multiple time points, and there is no easy fix for the current problem. However, this situation is relatively uncommon, as suggested by the post hoc analysis (i.e., 219 out of 2347 or roughly 9.3% in the full cohort data), and is mitigated by the choice of ROI analysis (e.g., only a single number is compared). However, mislabeling is related to only partial inaccuracy in ROI analysis after the averaging effect. In other words, these "mislabeled" WMLs would have overlapping voxels longitudinally with more than one WML that underwent a complicated morphological evolution (e.g., a part of a confluent WML is from a separation of another distal WML). Despite potential ambiguity, which imposes caveats on interpretation, this does not defy the spirit of the general study design given the systematic definitions of the ROI.

Interpreting the individual metric is quite difficult, given their moderate specificity for neuroinflammatory events, as illustrated by methods like diffusion tensor imaging (Jelescu & Fieremans, 2023). The finding of R2* (i.e., high R2* in the pNAWM) is another example because of its sensitivity to iron (especially in MS studies) (Neema et al., 2007) and myelin (Hametner et al., 2018; Paling et al., 2012). Similarly, the negative finding in MWF posed another barrier to the metric interpretation, which can be complicated by the mixture effect of iron and myelin in MS (Birkl et al., 2019). Further studies should aim to remove this particular barrier, for example, by using the χ separation to enhance understanding (Lee et al., 2024). Nonetheless, this limitation is widely recognized in the literature and may extend to other qMRI metrics (Granziera et al., 2021; Gulani & Seiberlich, 2020; Keenan et al., 2019; Manfrini et al., 2021; Saltarelli et al., 2025; Tranfa et al., 2022). Ultimately, this drives the multimodal approach with complementary qMRI metrics to gain rich information (Chapter 2) and mitigate potential issues such as resolution and sensitivity discrepancies arising from acquisition and pre-processing preferences across different labs, therefore helping facilitate a more decisive conclusion.

Figure 5-4: A simple demonstration of the ambiguity of lesion labelling (AFIL algorithm).



AFIL impose a condition that separated lesion cannot be confluent at the same time. The majority of the bottom right WML in red though is separated into two green WML in the follow-up in the bottom right, it also becomes part of the confluent WML in the follow-up period near the top along with the other two red WML. The level of complexity increases as the lesion evolutionary patterns is more convoluted and when considering cases in 3D instead of 2D.

5.6 Conclusions

By utilizing a set of tissue-evolution-informed ROIs, this longitudinal study has not only demonstrated the advanced sensitivity in qMRI metrics, but also shown the heterogeneity in both space and time that characterizes MS as a disease. Validation studies are highly encouraged to verify the findings. Still, this motivates future work to adapt qMRI metrics to characterize finer regions in the WM space to better understand tissue evolution in MS, potentially leading to a richer, more personalized delineation of a tissue heterogeneity map early on to better inform clinical decisions and facilitate precision in medicine for MS.

5.7 References

Absinta, M., Sati, P., & Reich, D. S. (2016). Advanced MRI and staging of multiple sclerosis lesions. *Nature Reviews Neurology*, *12*(6), 358–368. https://doi.org/10.1038/nrneurol.2016.59

Andravizou, A., Dardiotis, E., Artemiadis, A., Sokratous, M., Siokas, V., Tsouris, Z., Aloizou, A.-M., Nikolaidis, I., Bakirtzis, C., Tsivgoulis, G., Deretzi, G., Grigoriadis, N., Bogdanos, D. P., & Hadjigeorgiou, G. M. (2019). Brain atrophy in multiple sclerosis: mechanisms, clinical relevance and treatment options. *Autoimmunity Highlights*, *10*(1), 7. https://doi.org/10.1186/s13317-019-0117-5

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bermel, R. A., & Bakshi, R. (2006). The measurement and clinical relevance of brain atrophy in multiple sclerosis. *The Lancet Neurology*, *5*(2), 158–170. https://doi.org/10.1016/S1474-4422(06)70349-0

Birkl, C., Birkl-Toeglhofer, A. M., Endmayr, V., Höftberger, R., Kasprian, G., Krebs, C., Haybaeck, J., & Rauscher, A. (2019). The influence of brain iron on myelin water imaging. *NeuroImage*, *199*, 545–552. https://doi.org/10.1016/j.neuroimage.2019.05.042

Boaventura, M., Sastre-Garriga, J., Garcia-Vidal, A., Vidal-Jordana, A., Quartana, D., Carvajal, R., Auger, C., Alberich, M., Tintoré, M., Rovira, À., Montalban, X., & Pareto, D. (2022). T1/T2-weighted ratio in multiple sclerosis: A longitudinal study with clinical associations. *NeuroImage: Clinical*, *34*, 102967. https://doi.org/10.1016/j.nicl.2022.102967

Caranova, M., Soares, J. F., Pereira, D. J., Lima, A. C., Sousa, L., Batista, S., Castelo-Branco, M., & Duarte, J. V. (2025). Longitudinal Identification of Pre-Lesional Tissue in Multiple Sclerosis With Advanced Diffusion MRI. *Journal of Neuroimaging*, *35*(1). https://doi.org/10.1111/jon.70022

de Groot, M., Verhaaren, B. F. J., de Boer, R., Klein, S., Hofman, A., van der Lugt, A., Ikram, M. A., Niessen, W. J., & Vernooij, M. W. (2013). Changes in Normal-Appearing White Matter Precede Development of White Matter Lesions. *Stroke*, *44*(4), 1037–1042. https://doi.org/10.1161/STROKEAHA.112.680223

Dousset, V., Gayou, A., Brochet, B., & Caille, J.-M. (1998). Early structural changes in acute MS lesions assessed by serial magnetization transfer studies. *Neurology*, *51*(4), 1150–1155. https://doi.org/10.1212/WNL.51.4.1150

Dworkin, J. D., Linn, K. A., Oguz, I., Fleishman, G. M., Bakshi, R., Nair, G., Calabresi, P. A., Henry, R. G., Oh, J., Papinutto, N., Pelletier, D., Rooney, W., Stern, W., Sicotte, N. L., Reich, D. S., & Shinohara, R. T. (2018). An Automated Statistical Technique for Counting Distinct Multiple Sclerosis Lesions. *American Journal of Neuroradiology*, *39*(4), 626–633. https://doi.org/10.3174/ajnr.A5556

Elliott, C., Momayyezsiahkal, P., Arnold, D. L., Liu, D., Ke, J., Zhu, L., Zhu, B., George, I. C., Bradley, D. P., Fisher, E., Cahir-McFarland, E., Stys, P. K., Geurts, J. J. G., Franchimont, N., Gafson, A., & Belachew, S. (2021). Abnormalities in normal-appearing white matter from which multiple sclerosis lesions arise. *Brain Communications*, *3*(3), fcab176. https://doi.org/10.1093/braincomms/fcab176

Filippi, M., Preziosa, P., Banwell, B. L., Barkhof, F., Ciccarelli, O., De Stefano, N., Geurts, J. J. G., Paul, F., Reich, D. S., Toosy, A. T., Traboulsee, A., Wattjes, M. P., Yousry, T. A., Gass, A.,

Lubetzki, C., Weinshenker, B. G., & Rocca, M. A. (2019). Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. In *Brain* (Vol. 142, Number 7, pp. 1858–1875). Oxford University Press. https://doi.org/10.1093/brain/awz144

Gloor, M., Andelova, M., Gaetano, L., Papadopoulou, A., Burguet Villena, F., Sprenger, T., Radue, E.-W., Kappos, L., Bieri, O., & Garcia, M. (2023). Longitudinal analysis of new multiple sclerosis lesions with magnetization transfer and diffusion tensor imaging. *European Radiology*, *34*(3), 1680–1691. https://doi.org/10.1007/s00330-023-10173-6

Granziera, C., Wuerfel, J., Barkhof, F., Calabrese, M., De Stefano, N., Enzinger, C., Evangelou, N., Filippi, M., Geurts, J. J. G., Reich, D. S., Rocca, M. A., Ropele, S., Rovira, À., Sati, P., Toosy, A. T., Vrenken, H., Gandini Wheeler-Kingshott, C. A. M., Kappos, L., Barkhof, F., … Yousry, T. (2021). Quantitative magnetic resonance imaging towards clinical application in multiple sclerosis. *Brain*, *144*(5), 1296–1311. https://doi.org/10.1093/brain/awab029

Gulani, V., & Seiberlich, N. (2020). Quantitative MRI: Rationale and Challenges. In *Advances in magnetic resonance technology and applications* (Vol. 1, pp. xxxvii–li). Elsevier. https://doi.org/10.1016/B978-0-12-817057-1.00001-9

Guttmann, C. R., Ahn, S. S., Hsu, L., Kikinis, R., & Jolesz, F. A. (1995). The evolution of multiple sclerosis lesions on serial MR. *AJNR. American Journal of Neuroradiology*, *16*(7), 1481–1491. http://www.ncbi.nlm.nih.gov/pubmed/7484637

Hametner, S., Endmayr, V., Deistung, A., Palmrich, P., Prihoda, M., Haimburger, E., Menard, C., Feng, X., Haider, T., Leisser, M., Köck, U., Kaider, A., Höftberger, R., Robinson, S., Reichenbach, J. R., Lassmann, H., Traxler, H., Trattnig, S., & Grabner, G. (2018). The influence of brain iron and myelin on magnetic susceptibility and effective transverse relaxation - A

biochemical and histological validation study. *NeuroImage*, *179*, 117–133. https://doi.org/10.1016/j.neuroimage.2018.06.007

Hemond, C. C., Gaitán, M. I., Absinta, M., & Reich, D. S. (2024). New Imaging Markers in Multiple Sclerosis and Related Disorders. *Neuroimaging Clinics of North America*, *34*(3), 359–373. https://doi.org/10.1016/j.nic.2024.03.004

Jelescu, I. O., & Fieremans, E. (2023). Sensitivity and specificity of diffusion MRI to neuroinflammatory processes. In C. Laule & J. D. Port (Eds.), *Imaging Neuroinflammation* (Vol. 9, pp. 31–50). Academic Press. https://doi.org/10.1016/B978-0-323-91771-1.00010-1

Keenan, K. E., Biller, J. R., Delfino, J. G., Boss, M. A., Does, M. D., Evelhoch, J. L., Griswold, M. A., Gunter, J. L., Hinks, R. S., Hoffman, S. W., Kim, G., Lattanzi, R., Li, X., Marinelli, L., Metzger, G. J., Mukherjee, P., Nordstrom, R. J., Peskin, A. P., Perez, E., … Sullivan, D. C. (2019). Recommendations towards standards for quantitative MRI (qMRI) and outstanding needs. *Journal of Magnetic Resonance Imaging*, *49*(7). https://doi.org/10.1002/jmri.26598

Köhler, C., Wahl, H., Ziemssen, T., Linn, J., & Kitzler, H. H. (2019). Exploring individual multiple sclerosis lesion volume change over time: Development of an algorithm for the analyses of longitudinal quantitative MRI measures. *NeuroImage: Clinical*, *21*, 101623. https://doi.org/10.1016/j.nicl.2018.101623

Lee, J., Ji, S., & Oh, S.-H. (2024). So You Want to Image Myelin Using MRI: Magnetic Susceptibility Source Separation for Myelin Imaging. *Magnetic Resonance in Medical Sciences*, *23*(3), rev.2024-0001. https://doi.org/10.2463/mrms.rev.2024-0001

Liu, Y., Mitchell, P. J., Kilpatrick, T. J., Stein, M. S., Harrison, L. C., Baker, J., Ditchfield, M., Li, K., Egan, G. F., Butzkueven, H., & Kolbe, S. C. (2012). Diffusion tensor imaging of acute

inflammatory lesion evolution in multiple sclerosis. *Journal of Clinical Neuroscience*, *19*(12), 1689–1694. https://doi.org/10.1016/j.jocn.2012.03.022

Lou, C., Sati, P., Absinta, M., Clark, K., Dworkin, J. D., Valcarcel, A. M., Schindler, M. K., Reich, D. S., Sweeney, E. M., & Shinohara, R. T. (2021). Fully automated detection of paramagnetic rims in multiple sclerosis lesions on 3T susceptibility-based MR imaging. *NeuroImage: Clinical*, *32*, 102796. https://doi.org/10.1016/j.nicl.2021.102796

Love, S. (2006). Demyelinating diseases. *Journal of Clinical Pathology*, *59*(11), 1151–1159. https://doi.org/10.1136/jcp.2005.031195

Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sørensen, P. S., Thompson, A. J., Wolinsky, J. S., Balcer, L. J., Banwell, B., Barkhof, F., Bebo, B., Calabresi, P. A., Clanet, M., Comi, G., Fox, R. J., Freedman, M. S., Goodman, A. D., Inglese, M., Kappos, L., … Polman, C. H. (2014). Defining the clinical course of multiple sclerosis. *Neurology*, *83*(3), 278–286. https://doi.org/10.1212/WNL.0000000000000560

Manfrini, E., Smits, M., Thust, S., Geiger, S., Bendella, Z., Petr, J., Solymosi, L., & Keil, V. C. (2021). From research to clinical practice: a European neuroradiological survey on quantitative advanced MRI implementation. *European Radiology*, *31*(8), 6334–6341. https://doi.org/10.1007/s00330-020-07582-2

Martire, M. S., Moiola, L., Rocca, M. A., Filippi, M., & Absinta, M. (2022). What is the potential of paramagnetic rim lesions as diagnostic indicators in multiple sclerosis? *Expert Review of Neurotherapeutics*, *22*(10), 829–837. https://doi.org/10.1080/14737175.2022.2143265

McGinley, M. P., Goldschmidt, C. H., & Rae-Grant, A. D. (2021). Diagnosis and Treatment of Multiple Sclerosis. *JAMA*, *325*(8), 765. https://doi.org/10.1001/jama.2020.26858

Meca-Lallana, V., Berenguer-Ruiz, L., Carreres-Polo, J., Eichau-Madueño, S., Ferrer-Lozano, J., Forero, L., Higueras, Y., Téllez Lara, N., Vidal-Jordana, A., & Pérez-Miralles, F. C. (2021). Deciphering Multiple Sclerosis Progression. *Frontiers in Neurology*, *12*, 608491. https://doi.org/10.3389/fneur.2021.608491

Mey, G. M., Mahajan, K. R., & DeSilva, T. M. (2023). Neurodegeneration in multiple sclerosis. *WIREs Mechanisms of Disease*, *15*(1). https://doi.org/10.1002/wsbm.1583

Neema, M., Stankiewicz, J., Arora, A., Dandamudi, V. S. R., Batt, C. E., Guss, Z. D., Al-Sabbagh, A., & Bakshi, R. (2007). T1- and T2-Based MRI Measures of Diffuse Gray Matter and White Matter Damage in Patients with Multiple Sclerosis. *Journal of Neuroimaging*, *17*(s1). https://doi.org/10.1111/j.1552-6569.2007.00131.x

Paling, D., Tozer, D., Wheeler-Kingshott, C., Kapoor, R., Miller, D. H., & Golay, X. (2012). Reduced R_2' in multiple sclerosis normal appearing white matter and lesions may reflect decreased myelin and iron content. *Journal of Neurology, Neurosurgery & Psychiatry*, *83*(8), 785–792. https://doi.org/10.1136/jnnp-2012-302541

Pike, G. B., De Stefano, N., Narayanan, S., Worsley, K. J., Pelletier, D., Francis, G. S., Antel, J. P., & Arnold, D. L. (2000). Multiple Sclerosis: Magnetization Transfer MR Imaging of White Matter before Lesion Appearance on T2-weighted Images. *Radiology*, *215*(3), 824–830. https://doi.org/10.1148/radiology.215.3.r00jn02824

Pozzilli, C., Pugliatti, M., Vermersch, P., Grigoriadis, N., Alkhawajah, M., Airas, L., & Oreja-Guevara, C. (2023). Diagnosis and treatment of progressive multiple sclerosis: A position paper. *European Journal of Neurology*, *30*(1), 9–21. https://doi.org/10.1111/ene.15593

Ricigliano, V. A. G., Tonietto, M., Hamzaoui, M., Poirion, É., Lazzarotto, A., Bottlaender, M., Gervais, P., Maillart, E., Stankoff, B., & Bodini, B. (2022). Spontaneous remyelination in

lesions protects the integrity of surrounding tissues over time in multiple sclerosis. *European Journal of Neurology*, *29*(6), 1719–1729. https://doi.org/10.1111/ene.15285

Rimkus, C. de M., Otsuka, F. S., Nunes, D. M., Chaim, K. T., & Otaduy, M. C. G. (2024). Central Vein Sign and Paramagnetic Rim Lesions: Susceptibility Changes in Brain Tissues and Their Implications for the Study of Multiple Sclerosis Pathology. *Diagnostics*, *14*(13), 1362. https://doi.org/10.3390/diagnostics14131362

Rocca, M. A., Preziosa, P., Barkhof, F., Brownlee, W., Calabrese, M., De Stefano, N., Granziera, C., Ropele, S., Toosy, A. T., Vidal-Jordana, À., Di Filippo, M., & Filippi, M. (2024). Current and future role of MRI in the diagnosis and prognosis of multiple sclerosis. *The Lancet Regional Health - Europe*, *44*, 100978. https://doi.org/10.1016/j.lanepe.2024.100978

Saltarelli, G., Di Cerbo, G., Innocenzi, A., De Felici, C., Splendiani, A., & Di Cesare, E. (2025). Quantitative MRI in Neuroimaging: A Review of Techniques, Biomarkers, and Emerging Clinical Applications. *Brain Sciences*, *15*(10), 1088. https://doi.org/10.3390/brainsci15101088

Sati, P., Oh, J., Constable, R. T., Evangelou, N., Guttmann, C. R. G., Henry, R. G., Klawiter, E. C., Mainero, C., Massacesi, L., McFarland, H., Nelson, F., Ontaneda, D., Rauscher, A., Rooney, W. D., Samaraweera, A. P. R., Shinohara, R. T., Sobel, R. A., Solomon, A. J., Treaba, C. A., … Reich, D. S. (2016). The central vein sign and its clinical evaluation for the diagnosis of multiple sclerosis: a consensus statement from the North American Imaging in Multiple Sclerosis Cooperative. *Nature Reviews Neurology*, *12*(12), 714–722. https://doi.org/10.1038/nrneurol.2016.166

Silver, N. C., Lai, M., Symms, M. R., Barker, G. J., McDonald, W. I., & Miller, D. H. (1998). Serial magnetization transfer imaging to characterize the early evolution of new MS lesions. *Neurology*, *51*(3), 758–764. https://doi.org/10.1212/WNL.51.3.758

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., Correale, J., Fazekas, F., Filippi, M., Freedman, M. S., Fujihara, K., Galetta, S. L., Hartung, H. P., Kappos, L., Lublin, F. D., Marrie, R. A., Miller, A. E., Miller, D. H., Montalban, X., … Cohen, J. A. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. In *The Lancet Neurology* (Vol. 17, Number 2, pp. 162–173). Lancet Publishing Group. https://doi.org/10.1016/S1474-4422(17)30470-2

Tranfa, M., Pontillo, G., Petracca, M., Brunetti, A., Tedeschi, E., Palma, G., & Cocozza, S. (2022). Quantitative MRI in Multiple Sclerosis: From Theory to Application. *American Journal of Neuroradiology*, *43*(12), 1688–1695. https://doi.org/10.3174/ajnr.A7536

Uddin, M. N., Figley, T. D., Kornelsen, J., Mazerolle, E. L., Helmick, C. A., O'Grady, C. B., Pirzada, S., Patel, R., Carter, S., Wong, K., Essig, M. R., Graff, L. A., Bolton, J. M., Marriott, J. J., Bernstein, C. N., Fisk, J. D., Marrie, R. A., & Figley, C. R. (2022). The comorbidity and cognition in multiple sclerosis (CCOMS) neuroimaging protocol: Study rationale, MRI acquisition, and minimal image processing pipelines. *Frontiers in Neuroimaging*, *1*. https://doi.org/10.3389/fnimg.2022.970385

Wattjes, M. P., Ciccarelli, O., Reich, D. S., Banwell, B., de Stefano, N., Enzinger, C., Fazekas, F., Filippi, M., Frederiksen, J., Gasperini, C., Hacohen, Y., Kappos, L., Li, D. K. B., Mankad, K., Montalban, X., Newsome, S. D., Oh, J., Palace, J., Rocca, M. A., … Oh, J. (2021). 2021 MAGNIMS–CMSC–NAIMS consensus recommendations on the use of MRI in patients with multiple sclerosis. *The Lancet Neurology*, *20*(8), 653–670. https://doi.org/10.1016/S1474-4422(21)00095-8

Weingärtner, S., Desmond, K. L., Obuchowski, N. A., Baessler, B., Zhang, Y., Biondetti, E., Ma, D., Golay, X., Boss, M. A., Gunter, J. L., Keenan, K. E., & Hernando, D. (2022).

Development, validation, qualification, and dissemination of quantitative MR methods: Overview and recommendations by the ISMRM quantitative MR study group. *Magnetic Resonance in Medicine*, *87*(3), 1184–1206. https://doi.org/10.1002/mrm.29084

Wynen, M., Gordaliza, P. M., Istasse, M., Stölting, A., Maggi, P., Macq, B., & Bach Cuadra, M. (2025). ConfLUNet: Multiple sclerosis lesion instance segmentation in presence of confluent lesions. *Computers in Biology and Medicine*, *197*, 110919. https://doi.org/10.1016/j.compbiomed.2025.110919

Wynen, M., Istasse, M., Gordaliza, P. M., Stölting, A., Maggi, P., Macq, B., & Cuadra, M. B. (2024). Conflunet: Improving Confluent Lesion Identification In Multiple Sclerosis With Instance Segmentation. *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, 1–5. https://doi.org/10.1109/ISBI56570.2024.10635707

Chapter 6

Can multiple sclerosis white matter lesion evolution be predicted using baseline MRI? A promising attempt using a two-year follow-up dataset and a simple U-Net architecture

Publication status

Parts of this chapter was presented in the abstract entitled 'Predicting multiple sclerosis lesion evolution over two-years using baseline MRI and deep learning*'*at **Organization of Human Brain Mapping** and '*Can multiple sclerosis white matter lesion evolution be predicted using baseline MRI? A promising attempt using a two-year follow-up dataset and a simple U-Net architecture'* at **Manitoba Neuroscience Network** conferences in 2026. The manuscript is currently under internal review before submission to a peer-reviewed journal.

6.1 Abstract

Since baseline quantitative magnetic resonance imaging (qMRI) has shown promise in informing longitudinal tissue transitions in white matter (WM) in people with multiple sclerosis (pwMS), we investigated the performance of a 3D U-Net model for predicting MS white matter lesion (WML) evolution using both separate and combined conventional MRI (cMRI) and qMRI inputs. Baseline and two-year follow-up WML masks were used to compute 4-class (i.e., shrinking, stable, growing, and background) disease evolution maps (DEMs) for the MS participants (N=35 [28F], mean [std] age=45.5 [10.9] years). Three sets of baseline MRI inputs (i.e., qMRI-based [10 maps], cMRI-based [2 normalized images] and combined [all images and maps]) were separately 4-D concatenated with baseline WM and WML masks to train three separate 3D U-Nets to minimize the 4-class DEM cross-entropy. Utilizing leave-one-out cross-validation, we used ground-truth

follow-up WML masks to compute Dice similarity coefficients (DSCs) with each person's ground-truth baseline ($DSC_{baseline}$) and predicted follow-up ($DSC_{modelled}$) WML mask. Wilcoxon signed-rank tests were used to compare $DSC_{baseline}$ and $DSC_{modelled}$ (one-tailed), and between pairs of $DSC_{modelled}$ (two-tailed). After correcting for multiple comparisons, all $DSCs_{modelled}$ (median [IQR] $DSC_{modelled(qMRI)}$ = 0.72 [0.60 – 0.80], $DSC_{modelled(cMRI)}$ = 0.73 [0.64 – 0.79], $DSC_{modelled(combined)}$ = 0.73 [0.66 – 0.80]) were statistically higher than $DSC_{baseline}$ ($p_{FWE} < 0.05$), but were not statistically different from one another. Although qMRI (alone or in combination) did not yield better model performance than cMRI, these findings show that spatiotemporal patterns of MS WML evolution can be predicted two years in advance using only baseline MRIs.

6.2 Introduction

Multiple sclerosis (MS) is an autoimmune disease that causes demyelination and neurodegeneration in the central nervous system (Korn, 2008; Lassmann, 2018; Lassmann et al., 2007). The inherent pathological heterogeneity of lesions is one of the important factors that leads to a wide spectrum of clinical presentations in MS (Disanto et al., 2011; Kuhlmann et al., 2017; Lucchinetti et al., 2000). Disease progression can be slowed and halted by disease-modifying therapies (DMT) (McGinley et al., 2021), and personalization of treatment has become a general recommendation to improve patient outcomes (Comabella et al., 2016; Pathak, 2023). Notably, prognosis is very important for personalized medicine in MS. Many factors have been identified with prognostic potential, such as demographic, environmental, clinical, MRI findings and laboratory measurements (e.g., the presence of oligoclonal bands and neurofilament light chain) (Rotstein & Montalban, 2019). However, due to the lack of a reliable biomarker, the current prognosis framework is built upon a holistic approach, similar to the diagnosis process, but more challenging (to reach the same level of sensitivity and specificity) (Rotstein & Montalban, 2019).

Among these identified prognostic factors, qualitative lesion features on conventional MRI (cMRI) are increasingly important in successive revisions of MS diagnostic criteria (Montalban et al., 2025; Polman et al., 2011; Thompson et al., 2018). These are spatiotemporal features of MS lesions based on their locations (e.g., juxtacortical, periventricular, infratentorial, spinal cord, optic nerve) and appearance and morphologies (e.g., central vein sign (CVS), paramagnetic rim lesion (PRL)). Lesions can also be classified by stage (e.g., acute/chronic, active/inactive) based on their shapes and textures using designated MRI sequences (Absinta et al., 2016). Those findings, including the semi-quantitative measures (e.g., numbers of lesions and the total volume of the lesions), not only enhance the specificity of the diagnosis, but also carry implications for disease

prognosis (Hoffmann et al., 2024; Lomer et al., 2024; Rotstein & Montalban, 2019). Clearly, being able to predict where the new lesions will appear and how lesions evolve could provide valuable information to aid MS prognosis.

At the same time, quantitative MRI (qMRI) techniques have higher sensitivity and specificity to various pathophysiological processes that can potentially shed light on MS-related changes (Lavrova et al., 2021; Saltarelli et al., 2025; Tranfa et al., 2022). Several studies have found abnormality in important brain regions such as cortical gray matter in MS (Straub et al., 2023; Tardif et al., 2012) with correlation of cognitive brain functions (van Wijnen et al., 2020) and clinical disabilities (Gracien et al., 2016; Lema Dopico et al., 2021). Many qMRI measures can detect early changes in normal appearing white matter (NAWM) (Miller et al., 1989; West et al., 2014), and are associated with underlying MS pathology (Moll et al., 2011). Retrospectively, several qMRI metrics (e.g., magnetic transfer ratio metrics, diffusion tensor metrics, T1-weighted-by-T2-weighted (T1w/T2w) ratio, myelin-related metrics) have been shown to detect early changes in pre-lesional space (Boaventura et al., 2022; Elliott et al., 2021; Gloor et al., 2023; Liu et al., 2012; Tagge et al., 2022; Chapter 5).

Despite qMRI methods being so powerful, we still lack a full understanding of the relationship between qMRI and MS tissue heterogeneity, and processing and utilizing the information embedded in these 3D metrics for lesion prediction require attention to the high dimensionality, due to the increasing number of advanced MRI features. With the contemporary achievements of artificial intelligence (AI) and machine learning (ML) in big data across different fields of research, there are more and more emerging MS prognostic applications that utilize ML and MRI data (Seccia et al., 2021; Vázquez-Marrufo et al., 2023). While the majority of ML prognostic applications in MS focus on predicting the clinical outcome, for example the course of

the disease and the disease progression informed by conventional clinical scores, lesion evolution prediction is also highly relevant to MS prognosis, as previously outlined. Caba et al. demonstrated promising results using baseline radiomic features from T1- and T2-weighted images to classify prelesional (i.e., future lesional) areas of the MS brains in the baseline, apart from NAWM areas, using an ML classification approach (Caba et al., 2022). Kelly et al. had reached a similar conclusion (using T2-FLAIR images), and they even demonstrated the clinical applicability by using the best performing model (i.e., XGBoost) to produce a probability map of lesion occurrence (Kelly et al., 2024). Furthermore, a recent work that used a diffusion-based U-Net architecture has also demonstrated the location of new and enlarging T2 lesions can be learnt using baseline radiomic features (Favero et al., 2025). These findings had all underscore the baseline heterogeneity in WM and the potential of ML to stratify those regions to inform future tissue changes. However, whether or not ML methods can be utilized to predict all types of transitions (i.e., shrinking, stable, growing WML and stable NAWM) in the WM space over time remains largely unexplored, especially using qMRI maps. This study is therefore motivated to employ a deep learning (DL) convolutional neural network architecture (specifically a 3D U-Net) to predict the evolution of person-specific WM space over two years using baseline cMRI and/or multimodal qMRI data.

6.3 Methods

Dataset summary

The approval of the study was gained from the University of Manitoba Health Research Ethics Board, and participants provided their consent before any scanning/analysis. The baseline MRI data and longitudinal brain masks (i.e., 2 years apart) were from a subset of 35 persons with

MS (pwMS) within the Comorbidity and Cognition in Multiple Sclerosis (CCOMS) Study (Uddin et al., 2022). Table 6-1 summarizes the demographic information (i.e., sex, age, baseline EDSS, and change in EDSS over time) of the cohort.

Table 6-1: Demographics of the MS cohort.

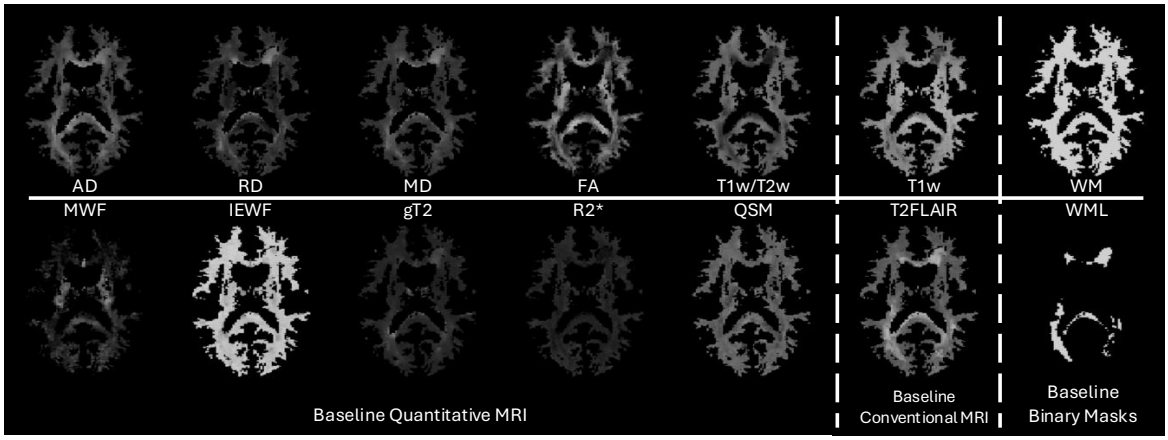
Demographic information of the participants	
Cohort size (Female)	35 (28)
Mean age (Std)	45.50 (10.89)
Median baseline EDSS (IQR)	3.0 (2 – 3.9)
Median change in EDSS (IQR)	0.5 (0 – 1.5)

qMRI map/cMRI images pre-processing and ROI mask computation

For each pwMS, the dataset includes 10 baseline qMRI maps: DTI-derived axial diffusivity (AD), radial diffusivity (RD), mean diffusivity (MD), and fractional anisotropy (FA); myelin-water-imaging-derived myelin water fraction (MWF), intra- and extra-cellular water fraction (IEWF), geometric mean T2 (gT2); T1-weighted-by-T2-weighted (T1w/T2w) ratio; multiecho gradient echo based transverse relaxation rate R2*, and quantitative susceptibility mapping (QSM) along with two baseline cMRI images: T1-weighted (T1-w) and T2-FLAIR, and global WM and WML masks from both timepoints. WML and NAWM from baseline and follow-up were computed following the same methodology as our previously published papers (Chapter 2; Chapter 3; Chapter 4; Chapter 5), except the follow-up WML mask was computed by taking the overlap between the follow-up lesion mask and the baseline WM mask to limit the view within the baseline WM space and stabilize the model performance due to the slight discrepancy between the thresholded WM masks from two time points. Specific image pre-processing pipelines were included in our recently published imaging protocol (Uddin et al., 2022), which describes how all quantitative qMRI maps were generated, how lesion masks were made, and how all maps and

images were non-linearly warped to MNI-152 template space. Moreover, given evidence of atrophy in multiple sclerosis (Andravizou et al., 2019; Bermel & Bakshi, 2006), especially over 2 years, the normalization step enables direct comparison of the brain across different time points within the same stereotactic space (i.e., MNI-152 space). Since this study focuses on using the feature maps (i.e., qMRI maps and cMRI maps) to predict the evolution of WM only, other regions (e.g., gray matter, cerebrospinal fluid) of the feature maps were cropped and discarded simply to reduce model complexity and enhance interpretability using the intermediate CAT12 tissue segmentation map from the pipeline (Gaser et al., 2024). In addition, the trade-off between the reduced field-of-view (FOV) near the inferior cerebellar region and imaging time during the myelin water imaging acquisition stage imposed an apparent brain mask on each subject to ensure a common FOV across all feature maps before they were fed into the model for training. This apparent brain mask was computed by applying the 'imclose' function in MATLAB (version R2020a; The MathWorks Inc., Natick, MA, USA) with a spherical element of a radius of 10 voxels on the MWF map. Figure 6-1 shows a panel of a participant with all the available inputs, qMRI, cMRI and tissue masks after aforementioned processing and masked by the baseline WM space.

Figure 6-1: A panel display of all baseline MRI maps/images and tissue masks as model inputs.

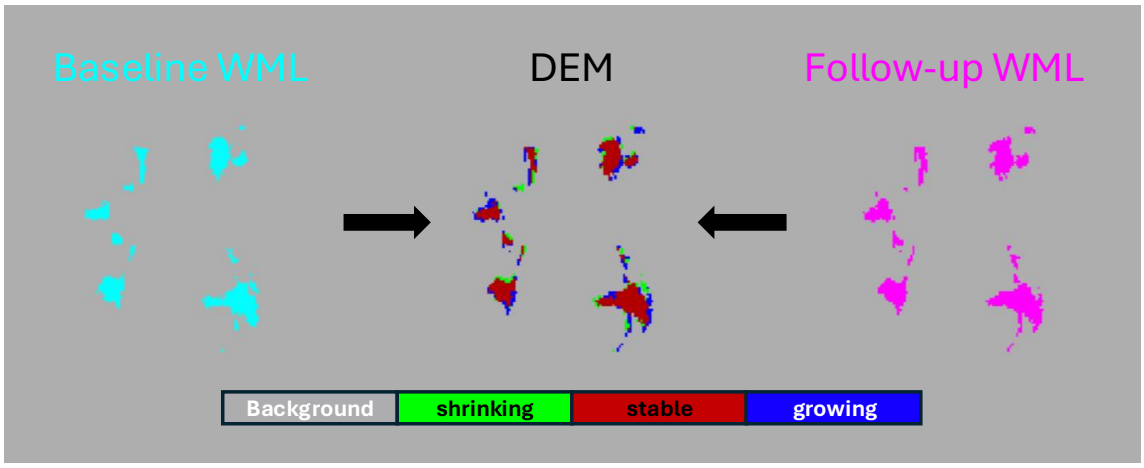


Model inputs (cropped with the subject-specific WM mask) included the baseline WML mask plus 10 baseline quantitative MRI maps (for the qMRI-based model), 2 conventional images (for the cMRI-based model), or 12 maps and images (for the combined model). All sets of MRI inputs are concatenated in the 4th dimension alongside baseline WML and WM masks.

Discrete disease evolution map computation as the model output

The concept of the disease evolution map (DEM) originated in a paper by Rachmadi et al. (Rachmadi et al., 2019), which investigated WM hyperintensity prediction. However, this generalized approach is also suitable for investigating MS WML prediction by computing a map with four classes relative to the baseline WM space: shrinking, stable, growing WML and background, as shown in Figure 6-2. Briefly, the shrinking (-1), stable (0) and growing (+1) WML regions were recognized and labelled by computing the difference between the binary follow-up WML mask and the binary baseline WML mask. It is worth noting that stable NAWM (i.e., the WM tissue that did not undergo tissue transition over time) is also categorized as background.

Figure 6-2: A panel display of WML masks from both time points and DEM.



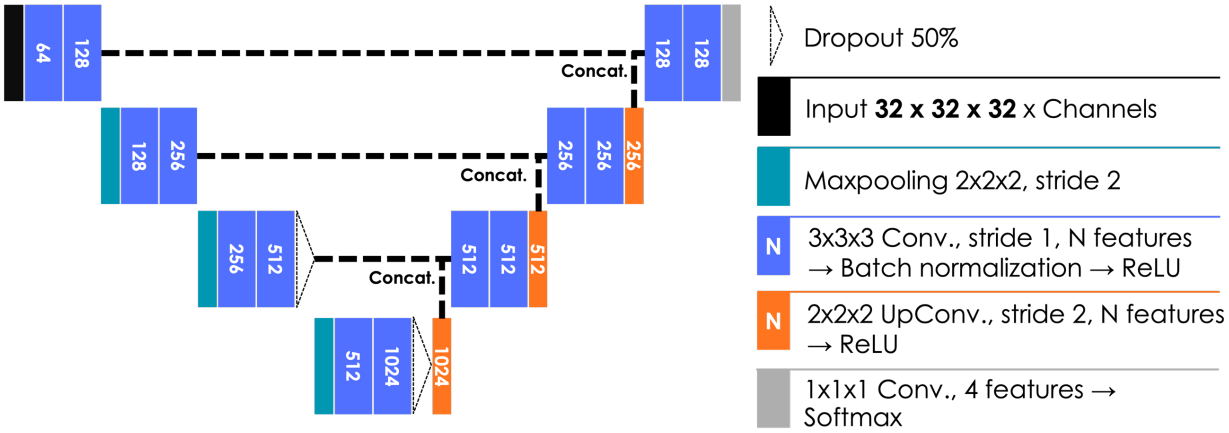
Example of ground truth baseline WML mask (left, cyan), 2-year follow-up WML mask (right, magenta), and the resulting disease evolution map (DEM; middle). (gray: background, green: shrinking, maroon: stable and blue: growing)

U-Net architecture and training

Since the task for the machine learning model is to 'segment' baseline brain regions that have different evolution trajectory in the follow-up using baseline MRI spatial features, the choice of a U-Net architecture among a large pool of DL architectures due to its well-known success in biomedical imaging segmentation (Ronneberger et al., 2015) and 3D learning is chosen to enable realistic learning beyond 2D features, which has also been demonstrated to have superior performance than the 2D counterpart (Çiçek et al., 2016). In fact, the potential of U-Net is clearly demonstrated in MS lesion segmentation applications – cross-sectional segmentation (Ma et al., 2022; Zeng et al., 2020) and longitudinal detection (Diaz-Hurtado et al., 2022). Due to the preliminary exploratory nature of this study, the authors decided to employ a relatively basic 3D U-Net architecture to investigate the feasibility of this approach and to establish a common ground for future comparisons. This 3D U-Net is autogenerated by MATLAB (version R2024b; The MathWorks Inc., Natick, MA, USA) by the function 'unet3d' with an encoder depth of 3, 64 filters in the first convolution layer, an iso-filter size of 3 and the convolution padding set to "same". This model takes a 4D input (e.g., concatenated baseline 3D MRI volumes and tissue masks) and outputs a normalized 4-class probability for each voxel using a softmax operator at the end of the network. We employed the Adam-W optimizer to yield more consistent regularization and better generalization (Loshchilov & Hutter, 2019). Despite using an 18 core (hyperthreaded to 36 core) CPU, quad-GPU workstation (4 x NVIDIA RTX8000 GPUs with NVLink interconnects); 256 GB of conventional RAM, plus 192 GB of dedicated video RAM, the entire 4D volume needed to be

subsampled into 4D blocks with spatial dimensions of 32x32x32 for each input to train the models. However, to mitigate potential biases from coverage and/or edge effects due to idiosyncratic subsampling, resampled blocks with 50% volume overlap between adjacent blocks were added to the training dataset. All 4 GPUs were run simultaneously, but with a deterministic allocation algorithm to enable result replication and controlled evaluation across models. Similarly, CPU and GPU random number streams are controlled throughout the initialization and shuffling steps to ensure replicable results. The key hyperparameters were optimized and chosen for all model training as follows: mini batch size of 128, learning rate of 5e-3, maximum epoch of 20 (with data shuffled at each epoch for each worker), weight decay factor of 0.01 (default), gradient decay factor of 0.9 (default), squared gradient decay of 0.999 (default), and epsilon of 1e-8 (default). The training was set to minimize the objective function, which was the 4-class cross-entropy between the subsampled 4D ground-truth DEM and the predicted DEM blocks (which were ultimately reassembled into a whole-brain DEM by placing each 4D block back into its original position). The full 3D U-Net architecture is displayed in Figure 6-3. All models (see below for the MRI input combinations) were trained using the leave-one-out cross-validation at the participant level (N=35). In this way, the whole-brain model-predicted DEM for each participant (i.e., the one left out from each iteration) was saved for subsequent testing and model evaluation.

Figure 6-3: Schematic of the 3D U-Net architecture used in this study.



This is a MATLAB (R2024b) generated 3D U-Net with the parameters specified in the method 'U-Net architecture and training' section, where the inputs were (subsampled blocks of dimension 32x32x32xC, where C denotes the input channels – i.e., relevant MRI data – depending on the model configuration).

Model comparison and evaluation

Using the same model architecture and parameters, 3D U-Nets were trained using three sets of inputs: one with each participants' baseline WM masks and all ten baseline qMRI maps; one with each participants' baseline WM masks and both baseline cMRI images, following intensity normalization wherein each participant's T1-w and T2-FLAIR images were rescaled between 0 and 1; and one with each participants' baseline WM masks, both normalized baseline cMRI images and all ten baseline qMRI maps. The intensity normalization step for the cMRI images was intended to reduce inter-participant scaling differences among T1-w and T2-FLAIR images because (unlike the qMRI maps), they are not inherently intensity normalized.

During every iteration of training, Dice similarity coefficients (DSCs) of each model ($DSC_{modelled(qMRI)}$, $DSC_{modelled(cMRI)}$ and $DSC_{modelled(combined)}$) were computed for each hold-out individual between the predicted follow-up WML mask and the ground truth follow-up WML (wholistically accounting for shrinking, stable, and growing WML regions). However, to provide a global control and to quantify the actual change over time, we also computed DSCs between each participants' ground truth baseline and follow-up WML masks ($DSC_{baseline}$). $DSC_{modelled(qMRI)}$, $DSC_{modelled(cMRI)}$ and $DSC_{modelled(combined)}$ were then compared to $DSC_{baseline}$ using one-tailed (assuming models would learn at least some temporal changes) Wilcoxon signed-rank tests paired for each individual; and then compared to each other (all possible pairs) using a two-tailed Wilcoxon signed-rank test. The significance threshold across all tests was then adjusted for multiple comparisons using Bonferroni correction for family-wise error ($p_{FWE} = 0.05/6 = 0.0083$). For each 3D U-Net input, DSCs were calculated for shrinking, stable, and growing WML regions based on the predicted vs. ground truth class-specific DEM for each participant, and these were then assessed within and across models. Finally, to elucidate where (spatially) the 3D U-Net

models were gleaning the most information from when predicting each DEM class, we employed, the segmentation gradient-weighted class activation map (seg-gradCAM) method (Vinogradova et al., 2020).

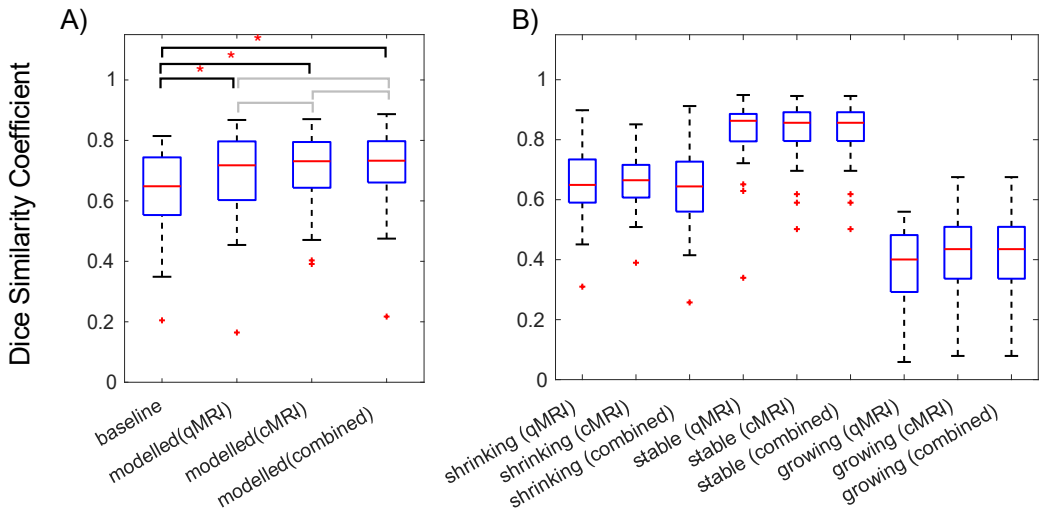
6.4 Results

The image input oversampling method yielded a total of 18,710 4D blocks across the entire 35-person dataset. Figure 6-4 shows box-and-whisker plots of DSC values (both aggregated and region-specific) for all three models. The three models (DSC median [IQR]: $DSC_{modelled(qMRI)}$ 0.72 [0.60,0.80]; $DSC_{modelled(cMRI)}$ 0.73 [0.64,0.79]; $DSC_{modelled(combined)}$ 0.73 [0.66,0.80]) all demonstrated the ability to predict better follow-up WML masks (one-tailed, $DSC_{modelled(qMRI)}$ p-value = 3e-6; $DSC_{modelled(cMRI)}$ p-value = 2e-6; $DSC_{modelled(combined)}$ p-value = 2e-6) compared to using just the baseline WML mask (or $DSC_{baseline}$ 0.64 [0.55,0.74]), as shown in Tables 6-2 and 6-3. Table 6-3 also summarizes the two-tailed tests between the trained models, which revealed no statistically significant differences after multiple-comparison correction. Among different DEM classes, stable WML regions had the best results (qMRI model = 0.86 [0.79,0.89]; cMRI model = 0.86 [0.80,0.89]; combined model = 0.86 [0.81,0.89]), followed by shrinking WML regions (qMRI model = 0.65 [0.59,0.73]; cMRI model = 0.66 [0.61,0.72]; combined model = 0.64 [0.56,0.73]), and growing WML regions had the lowest DSCs (qMRI model = 0.40 [0.29,0.48]; cMRI model = 0.44 [0.34,0.51]; combined model = 0.46 [0.37,0.53]). These trends were also apparent when comparing the actual and predicted 3D disease evolution maps (DEMs), which consistently showed high correspondence (Figure 6-5).

Similar to the block training and prediction for the model, Seg-gradCAM-derived saliency map was individually computed, rescaled to 0 – 1 (where ‘0’ means ‘not important’ and ‘1’ means

'very important') by class, and they were placed back to their original position to form the whole brain saliency map. The entire last (i.e., deepest) block of convolutional feature layers was chosen to facilitate the saliency map computation. As shown in Figure 6-6, models' attention on the spatial location within the brain block when making the prediction for a specific class was displayed on an axial slice for the participant with median $DSC_{modelled}$ performance in the combined model (DSC = 0.73 (modelled, or stable+growing), 0.56 (shrinking), 0.86 (stable) and 0.49 (growing)).

Figure 6-4: Box-and-whisker plots of the models' performance.



A) $DSC_{baseline}$ and $DSC_{modelled}$ of each model and B) DSC of the three main regions (i.e., shrinking, stable and growing) of the predictive task for each model. Significant results after BH-FDR correction are indicated by asterisks overlaid on top of the pairs of ROI comparison in A). Grayed brackets indicated the statistical test is not significant.

Table 6-2: A summary of the baseline and modelled DSCs for each type of 3D U-Net model input.

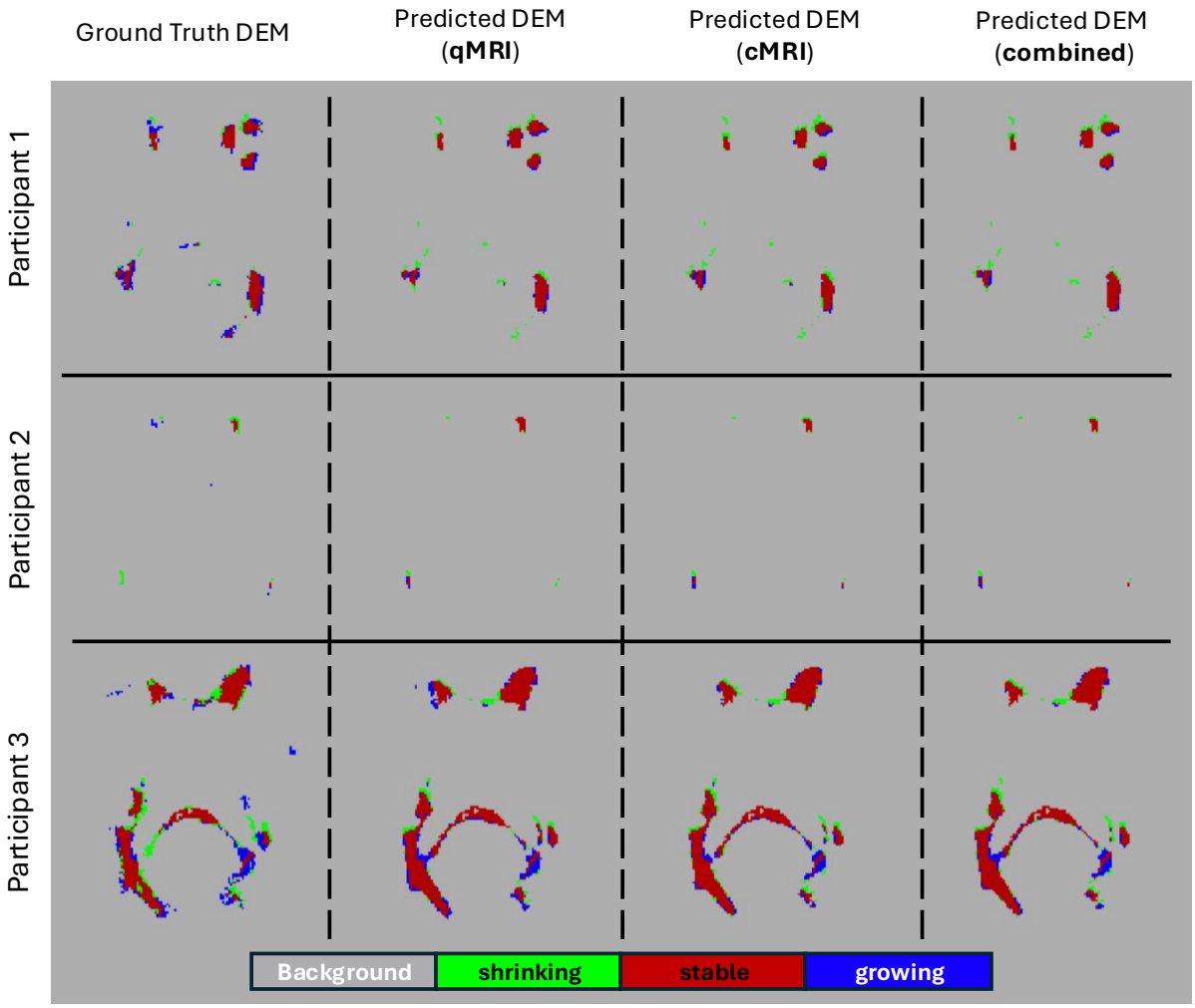
	qMRI-based	**cMRI-based**	**combined**
Median $DSC_{baseline}$ (IQR)	0.64 (0.55 – 0.74)		
Median $DSC_{modelled}$ (IQR)	0.72 (0.60 – 0.80)	0.73 (0.64 – 0.79)	0.73 (0.66 – 0.80)
Median $DSC_{shrinking}$ (IQR)	0.65 (0.59 – 0.73)	0.66 (0.61 – 0.72)	0.64 (0.56 – 0.73)
Median DSC_{stable} (IQR)	0.86 (0.79 – 0.89)	0.86 (0.80 – 0.89)	0.86 (0.81 – 0.89)
Median $DSC_{growing}$ (IQR)	0.40 (0.29 – 0.48)	0.44 (0.34 – 0.51)	0.46 (0.37 – 0.53)

Table 6-3: Wilcoxon signed-rank and rank sum test p-value results comparing $DSC_{baseline}$ and $DSCs_{modelled}$.

	baseline	qMRI-based	cMRI-based	combined
baseline				
qMRI-based	**3e-6**			
cMRI-based	**2e-6**	Not significant		
combined	**2e-6**	Not significant	Not significant	

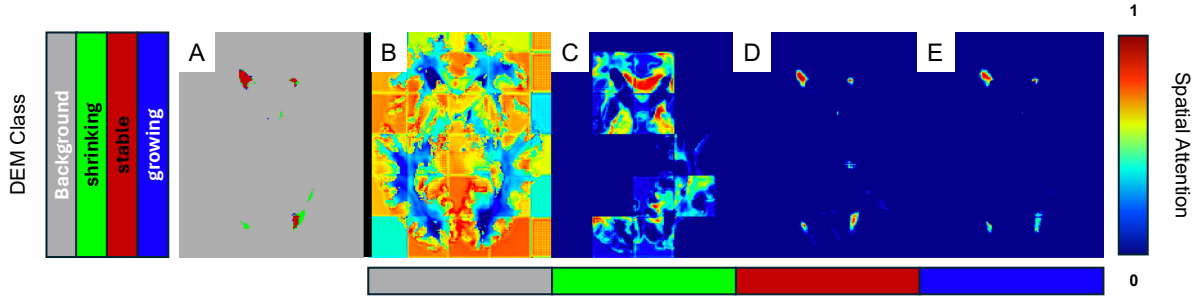
The box background colour indicates the tests were either one-tailed (Green) or two-tailed (Orange).

Figure 6-5: Visual comparison of the ground truth and model predicted DEM.



Ground truth (First column, counting from the left), qMRI (Second column), cMRI (Third column) and combined (Fourth column) output DEMs from three randomly selected MS participants (rows) are displayed here. (gray: background, green: shrinking, maroon: stable and blue: growing)

Figure 6-6: Visualization of the saliency maps of the combined model.



Leftmost (panel A) is the ground truth DEM with each colour representing a DEM class (gray: background, green: shrinking, maroon: stable and blue: growing). The Seg-GradCAM (combined model) saliency maps (panels B-E) for each specific class prediction of a single MS participant is displayed in a normalized scale from 0 to 1. Note: the chosen participant has the median performance out of the 35 iterations of training.

6.5 Discussion

The models learned the spatiotemporal features of the WML evolution from the baseline MRI

All $DSC_{modelled}$ of the models are significantly higher than the $DSC_{baseline}$, and the significance of the main findings is two-fold. First, the moderate-to-high level of DSCs (qMRI, cMRI and combined inputs all having median $DSC_{modelled} > 0.70$) demonstrates the feasibility of predicting spatiotemporal changes in WML tissue evolution in MS over a span of two years using only baseline MRI maps. Second, this indicates that heterogeneity of MRI signals, not just between, but also within MS WMLs and NAWM contains useful information about longitudinal WM transitions very early on (at least 2 years into the future).

Based on region-specific DSCs, the models were relatively more reliable at predicting stable MS WML regions. Besides tissue heterogeneity, this is probably due to the nature of

concentric expansion and contraction of MS WML, which often appear to uniformly grow and shrink from the periphery (Barkhof, 1997; Huh et al., 2014; I. Katz Sand, 2015; I. B. Katz Sand & Lublin, 2013). In other words, the spatial awareness of the whole WML potentially helped the model to identify the central, and therefore more likely stable, regions. However, this hypothesis cannot fully account for the result, because not all WMLs fit within the 32-by-32-by-32 sampling block due to their size. The models performed moderately well at predicting the shrinking MS WML regions, which, by definition, are the complementary regions to the stable regions within the baseline WML masks. Previous ROI-based work from our lab specifically examined both cross-sectional (Chapter 4) and longitudinal (Chapter 5) sensitivity of qMRI in and around WML, which generally showed greater MRI signal heterogeneity within WMLs than in NAWM (with generally higher microstructural measures near the WML boundary) and also consistently higher baseline microstructural measures in contracting (shrinking) portions vs. persisting (stable) portions of MS WMLs. It is also worth emphasizing that the moderate-to-high DSCs observed in stable and shrinking regions are not coincidental. This is because some participants experience drastic changes in tissue evolution (i.e., $DSC_{baseline} < 0.3$, where the shrinking region and growing region are much larger than the stable region). Therefore, the models must learn the features to correctly recognize which voxels within the baseline WML are stable vs. shrinking, as shown in the DSCs. On the other hand, the model struggled to predict the growing regions. Here, the result might be due to NAWM in the baseline not being apparent enough two years in advance and/or to the insensitivity of MRI methods at this level of heterogeneity. On one hand, this is perhaps not unexpected, because the cross-sectional study in the lab has shown that as we move further into the NAWM space (concentrically from MS WML boundaries), qMRI heterogeneity is less apparent than within WMLs (Chapter 4). However, our previous longitudinal study showed

consistent differences across qMRI metrics between stable and transitioning perilesional NAWM, so it is worth noting that the 3D U-Net models in this study were still able to predict at least some growing WML regions. This intriguing result is equally important, because it further underscores that at least some level of heterogeneity exists in the NAWM space, and that this heterogeneity is meaningful for informing the longitudinal growth of WMLs. The statistical significance of the comparisons between $DSCs_{modelled}$ and $DSC_{baseline}$ demonstrated that the growing region is both conceptually and statistically significant within the predicted follow-up WML mask (alongside stable regions). Overall, this finding from the growing regions reveals that heterogeneity in NAWM is perhaps more subtle than the WML and can be hard to capture using a fairly simple U-Net and/or within each 32-by-32-by-32 block.

Training of the combined model did not result in improved model performance

Perhaps somewhat surprisingly, the predictive performance of the qMRI-based and cMRI-based models was not significantly different, and even when cMRI and qMRI inputs were put together to train the combined model, the performance did not significantly improve over the cMRI-based model. From a practical perspective, this finding is especially interesting and highly relevant since T1-w and T2-FLAIR images are routinely acquired for patient diagnosis and treatment monitoring. High-accuracy prediction of MS lesion evolution using these methods, based on baseline MRI characteristics that are routinely available, and without the need for additional advanced imaging, is likely to be much more translatable to clinical practice. With widely available images (at least cMRI), larger sample sizes, and potentially more time frames, this result also bodes well for scalability and generalizability (see Study limitations on how this model can still be considered

generalizable given a low sample size) in future studies. Nonetheless, to determine why cMRI methods are comparable to qMRI in this predictive task requires further investigation.

Where is the model focusing on when predicting a specific class?

Interpreting even a simple neural network architecture like U-Net can be difficult, especially when it is a 3D U-Net with more learnable parameters and connections. Explainable AI (XAI) is the field in which scientists and researchers are developing methods to make model interpretation easier (Dwivedi et al., 2023). In neural networks, the magnitude of the gradient between the output or probability and the logits in the feature layers is often used as an intuitive metric, as the stronger the gradient, the greater its impact on the final prediction. The gradient of each logit is also the fundamental in saliency map computation. Segmentation gradient weighted class activation map (Seg-gradCAM) is a 3D generalized version of gradCAM (Vinogradova et al., 2020). Therefore, Seg-gradCAM is suitable for computing saliency maps for our model to visualize the spatial attention that the model pays toward making specific class predictions. The feature layers of the last convolutional block were chosen to compute the saliency map because deeper features in models generally encode with higher-level information about the inputs and are spatially relevant due to the nature of the U-Net architecture. As shown in Figure 6-6B-E, the Seg-gradCAM-derived saliency map of the feature maps in the deepest convolution block displayed a distinct pattern of spatial importance across different classes. The first thing we noticed is that attention to the location is suppressed when irrelevant information is presented. This is most obvious in predictions with shrinking regions (only blocks containing baseline WML receive attention) and the attention to regions near the core of baseline WML (i.e., bluest/coldest spots) when predicting background. Interestingly enough, the model ‘sees’ a higher level of importance in the NAWM space as

compared to the WML when predicting the shrinking region, for example, the hot spots near the posterior of the brain in Figure 6-6C. In addition, some important structures, for example, regions close to important deep gray matter structures like the caudate nucleus, and putamen and WM tracts like the corpus callosum, received higher level of attention. However, more work is needed to fully tease out the importance and spatial extent of the model's attention to these areas, for example, by using a predefined atlas to more accurately delineate the regions. On the other hand, when predicting stable and growing regions, attention shifted toward the core of the baseline WML regions in a gradient fashion (Figure 6-6D-E). The model relies more on the inner core of the baseline WMLs compared to other regions, suggesting that spatial information is perhaps more important within the WML core as compared to the rim when predicting changes over time (i.e., stable vs. growing). Nonetheless, these attention maps contained strips that was introduced discontinuities at the block boundaries due to the nature of the model architecture (predicting one block at a time). It inevitably created a barrier to fully interpreting the model's spatial attention at the whole-brain level and suggests that if enough GPU memory was available to accommodate all of the data at once, whole brain baseline MRI inputs might yield higher performance (even with the same 3D U-Net architecture and tuning parameters).

Study limitations

Despite being an exciting proof-of-concept for the application of deep learning to predict subject-specific MS lesion evolution using baseline MRI data, this work has several limitations. Perhaps the most obvious limitation is the relatively low participant-level sample size of the current study (N=35). However, although this might lead to questions about the generalizability of the findings to a larger population, several aspects of the model design and training have attempted to

compensate for this limitation by making the model less biased to the dataset and promoting the translatability of the model. These include: 1) block sampling and training, where the model was trained using augmented 4D block samples from the whole brain imaging data (which simultaneously increased the effective training sample size and minimized biases that could have otherwise been introduced into the model by individual differences in brain size, overall lesion loads, specific lesion locations within the brain, and whole-brain microstructural differences/variability throughout NAWM); 2) minibatch and randomized training, where the training dataset for each epoch was split into minibatches with a maximum size of 128 (roughly 0.6% of the whole dataset), and the whole training dataset for each iteration was shuffled into minibatches for each epoch (i.e., meaning the model is learning a very small fraction of the whole dataset each time, and since the data can come from anywhere in the brain across and within participants, the model also has to adjust and adapt rapidly to the randomness in the dataset); 3) use of drop-out layers, since the model used two drop-out layers (i.e., 50% of the parameters are randomly discarded in each iteration of minibatch), which further pushes the model to seek a generalizable set of parameters to compensate for the loss of information and be less sensitive to the model overfitting to the current dataset; 4) leave-one-out cross-validation, where the model was trained on as much data as possible (i.e., coming from all other study participants) and then evaluated using the most extreme case of cross-validation at the single-subject level. Although several of these factors likely served to reduce model performance (especially sub-sampling the images into 4D blocks, rather than using whole-brain data to train the model), they should increase our confidence regarding the model's generalizability given the minimal bias toward participant-level data. In fact, the significant result suggests that the model still learns to generalize and predict portions of the spatiotemporal pattern of WMLs, given the limited FOV and a harsh training

scheme, thereby enhancing the translatability of the concept to other (hopefully larger) external datasets with more diverse populations. Regardless, more work is required to improve the model's predictive performance, particularly in growing WML regions, since accurate prediction of lesion growth at the individual level could set the stage for clinical trials of individualized approaches to optimize treatment strategies for people with MS. Whether achieving this goal can be achieved through the use of much larger data sets and cMRI measures or will require the use of other qMRI methods (or both), remains unanswered. In addition, disease phenotypes and other clinical factors, as well as the status of therapy enrollment can be helpful when interacting with MRI information. For example, the effect of therapies in combination with baseline MRI radiomics has been investigated in specifically looking at predicting patient outcomes (Durso-Finley et al., 2022, 2023) and the detection of new and enlarging T2 lesions (Favero et al., 2025), with the positive results favoring the inclusion of other clinical data beyond MRI to yield a more reliable, realistic and accurate prediction. Although the current study was meant to serve as an early proof-of-principle, more powerful high-performance computing clusters would ideally be used to train future DL models using whole-brain MRI data (to account for expected intensity differences across brain regions) and provide the computational bandwidth to include all the additional features mentioned above to fully exploit the potential of real-world data. Furthermore, although the 3D U-Net employed in the current study demonstrated the feasibility and exciting potential for using MRI-based DL approaches to forecast the evolution of MS WMLs, future work should definitely explore the use of newer and more sophisticated AI approaches, such as generative adversarial networks (GANs) (Goodfellow et al., 2020). Despite a limited sample size (N=5), a study with a 3D GAN-alike model (i.e., Discriminator-based guided attention GAN) employed multi-modal cMRI (i.e., T1-weighted, T2-weighted, proton-density-weighted and FLAIR) with a much larger

receptive field (i.e., 128 x 128 x 128) to predict whole brain FLAIR image (i.e., whole brain image prediction) one year in the future, had reported a superior result in the MS lesion regions over the basic 3D U-Net (peak-signal-to-noise-ratio = 20.05), and also highlighted a comparable overall performance with basic 3D U-Net (structural similarity index = 0.88) (Wang et al., 2022). In addition, a recent work with the same dataset as mentioned above had tried not only to predict whole brain FLAIR image in the follow-up time point (ranging from 1 to 5 years in the future) but also to segment lesion area from the centre 2D slices (i.e., x=0 or y =0 or z=0) of the predicted whole FLAIR brain using another 3D GAN-alike model (i.e., Auxiliary classifier GAN), yielding a mean accuracy of 74.4% (Chan, 2026). Although the workflow (i.e., a two-step prediction-segmentation workflow) of this study was different from our work (i.e., one-step 'segmentation'), and the fact that the resultant segmentation was done in 2D fashion (partially because prediction was done in 3D), and they did not evaluate the whole brain lesion (i.e., only the three centre slices), the finding still highlighted the importance of exploring advanced models/architecture, the value of having multiple time points data, and most importantly the predictive ability of cMRI, which is also especially emphasized by the findings of current study. Lastly, it is important to also point out that the brain can experience varying degrees of atrophy over two years. Despite the straightforwardness of this analysis (by aligning all maps to a common template space and confining lesions to the baseline WM space), the definition of shrinking and growing regions can be affected by both global and local brain atrophy.

6.6 Conclusions

This study shows that a common, relatively simple DL model (3D U-Net), trained on baseline MRI data, is able to predict MS white matter lesion evolution (at least 2 years in advance). The finding

that conventional T1-w and T2-FLAIR images alone were able to forecast white matter tissue transitions bodes well for potential future clinical applicability, since predicting lesion changes in advance could perhaps be used to augment clinical decision-making in terms of patient management and precision medicine. The scalability of the method is also promising, as almost all MS visits include conventional MRI images. This significantly lowers the barrier to data access for a wider population to build potentially more accurate and generalizable models in the future (perhaps in combination with more advanced DL methods and other demographic and clinical input parameters). Future models should also consider using high-performance computing clusters to process whole-brain MRI data, and also feeding the model information from other clinical sources, like clinical tests, lab results and patient-reported outcomes alongside MRI and see how model performance changes. By using XAI methods like seg-gradCAM, we observed that regions beyond WML can also hold information that subjective human analyses by neurologists and radiologists would normally overlook. Therefore, although interesting in its own right, this study points to several exciting lines of future inquiry regarding the use of AI and MRI data to predict MS lesion evolution, and potentially guide proactive personalized medicine approaches in MS prognosis.

6.7 References

Absinta, M., Sati, P., & Reich, D. S. (2016). Advanced MRI and staging of multiple sclerosis lesions. *Nature Reviews Neurology*, *12*(6), 358–368. https://doi.org/10.1038/nrneurol.2016.59

Andravizou, A., Dardiotis, E., Artemiadis, A., Sokratous, M., Siokas, V., Tsouris, Z., Aloizou, A.-M., Nikolaidis, I., Bakirtzis, C., Tsivgoulis, G., Deretzi, G., Grigoriadis, N., Bogdanos,

D. P., & Hadjigeorgiou, G. M. (2019). Brain atrophy in multiple sclerosis: mechanisms, clinical relevance and treatment options. *Autoimmunity Highlights*, *10*(1), 7. https://doi.org/10.1186/s13317-019-0117-5

Barkhof, F. (1997). Comparison of MRI criteria at first presentation to predict conversion to clinically definite multiple sclerosis. *Brain*, *120*(11), 2059–2069. https://doi.org/10.1093/brain/120.11.2059

Bermel, R. A., & Bakshi, R. (2006). The measurement and clinical relevance of brain atrophy in multiple sclerosis. *The Lancet Neurology*, *5*(2), 158–170. https://doi.org/10.1016/S1474-4422(06)70349-0

Boaventura, M., Sastre-Garriga, J., Garcia-Vidal, A., Vidal-Jordana, A., Quartana, D., Carvajal, R., Auger, C., Alberich, M., Tintoré, M., Rovira, À., Montalban, X., & Pareto, D. (2022). T1/T2-weighted ratio in multiple sclerosis: A longitudinal study with clinical associations. *NeuroImage: Clinical*, *34*, 102967. https://doi.org/10.1016/j.nicl.2022.102967

Caba, B., Gafson, A., Lombard, A., Liu, D., Cafaro, A., Battistella, E., Perea, R., Fisher, E., Franchimont, N., Arnold, D. L., Elliott, C., Paragios, N., & Belachew, S. (2022). Machine Learning-Based Prediction of New Multiple Sclerosis Lesion Formation Using Radiomic Features from Pre-Lesion Normal Appearing White Matter (S26.009). *Neurology*, *98*(18_supplement). https://doi.org/10.1212/WNL.98.18_supplement.1687

Chan, C. T. G. (2026). MRI Prediction and Segmentation of Multiple Sclerosis Development using GAN and U-Net Architecture. In *Available at SSRN 6055175*. https://doi.org/10.2139/ssrn.6055175

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In S. Ourselin, L.

Joskowicz, M. R. Sabuncu, G. Unal, & W. Wells (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* (pp. 424–432). Springer International Publishing. https://doi.org/10.1007/978-3-319-46723-8_49

Comabella, M., Sastre-Garriga, J., & Montalban, X. (2016). Precision medicine in multiple sclerosis. *Current Opinion in Neurology*, *29*(3), 254–262. https://doi.org/10.1097/WCO.0000000000000336

Diaz-Hurtado, M., Martínez-Heras, E., Solana, E., Casas-Roma, J., Llufriu, S., Kanber, B., & Prados, F. (2022). Recent advances in the longitudinal segmentation of multiple sclerosis lesions on magnetic resonance imaging: a review. *Neuroradiology*, *64*(11), 2103–2117. https://doi.org/10.1007/s00234-022-03019-3

Disanto, G., Berlanga, A. J., Handel, A. E., Para, A. E., Burrell, A. M., Fries, A., Handunnetthi, L., De Luca, G. C., & Morahan, J. M. (2011). Heterogeneity in Multiple Sclerosis: Scratching the Surface of a Complex Disease. *Autoimmune Diseases*, *2011*, 1–12. https://doi.org/10.4061/2011/932351

Durso-Finley, J., Falet, J.-P., Mehta, R., Arnold, D. L., Pawlowski, N., & Arbel, T. (2023). Improving Image-Based Precision Medicine with Uncertainty-Aware Causal Models. In H. Greenspan, A. Madabhushi, P. Mousavi, S. Salcudean, J. Duncan, T. Syeda-Mahmood, & R. Taylor (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023* (pp. 472–481). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43904-9_46

Durso-Finley, J., Falet, J.-P. R., Nichyporuk, B., Arnold, D. L., & Arbel, T. (2022). Personalized Prediction of Future Lesion Activity and Treatment Effect in Multiple Sclerosis from Baseline MRI. *International Conference on Medical Imaging with Deep Learning*, 387–406. http://arxiv.org/abs/2204.01702

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., & Ranjan, R. (2023). Explainable AI (XAI): Core Ideas, Techniques, and Solutions. *ACM Computing Surveys*, *55*(9), 1–33. https://doi.org/10.1145/3561048

Elliott, C., Momayyezsiahkal, P., Arnold, D. L., Liu, D., Ke, J., Zhu, L., Zhu, B., George, I. C., Bradley, D. P., Fisher, E., Cahir-McFarland, E., Stys, P. K., Geurts, J. J. G., Franchimont, N., Gafson, A., & Belachew, S. (2021). Abnormalities in normal-appearing white matter from which multiple sclerosis lesions arise. *Brain Communications*, *3*(3), fcab176. https://doi.org/10.1093/braincomms/fcab176

Favero, G. M., Luo, G. Y., Fathi, N., Szeto, J., Arnold, D. L., Nichyporuk, B., Pal, C., & Arbel, T. (2025). Spatio-Temporal Conditional Diffusion Models for Forecasting Future Multiple Sclerosis Lesion Masks Conditioned on Treatments. *ArXiv Preprint ArXiv:2508.07006*. http://arxiv.org/abs/2508.07006

Gaser, C., Dahnke, R., Thompson, P. M., Kurth, F., Luders, E., & the Alzheimer's Disease Neuroimaging Initiative. (2024). CAT: a computational anatomy toolbox for the analysis of structural MRI data. *GigaScience*, *13*. https://doi.org/10.1093/gigascience/giae049

Gloor, M., Andelova, M., Gaetano, L., Papadopoulou, A., Burguet Villena, F., Sprenger, T., Radue, E.-W., Kappos, L., Bieri, O., & Garcia, M. (2023). Longitudinal analysis of new multiple sclerosis lesions with magnetization transfer and diffusion tensor imaging. *European Radiology*, *34*(3), 1680–1691. https://doi.org/10.1007/s00330-023-10173-6

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, *63*(11), 139–144. https://doi.org/10.1145/3422622

Gracien, R.-M., Jurcoane, A., Wagner, M., Reitz, S. C., Mayer, C., Volz, S., Hof, S.-M., Fleischer, V., Droby, A., Steinmetz, H., Zipp, F., Hattingen, E., Deichmann, R., & Klein, J. C. (2016). The Relationship between Gray Matter Quantitative MRI and Disability in Secondary Progressive Multiple Sclerosis. *PLOS ONE*, *11*(8), e0161036. https://doi.org/10.1371/journal.pone.0161036

Hoffmann, O., Gold, R., Meuth, S. G., Linker, R. A., Skripuletz, T., Wiendl, H., & Wattjes, M. P. (2024). Prognostic relevance of MRI in early relapsing multiple sclerosis: ready to guide treatment decision making? *Therapeutic Advances in Neurological Disorders*, *17*, 17562864241229324. https://doi.org/10.1177/17562864241229325

Huh, S.-Y., Min, J.-H., Kim, W., Kim, S.-H., Kim, H. J., Kim, B.-J., Kim, B. J., & Lee, K. H. (2014). The usefulness of brain MRI at onset in the differentiation of multiple sclerosis and seropositive neuromyelitis optica spectrum disorders. *Multiple Sclerosis Journal*, *20*(6), 695–704. https://doi.org/10.1177/1352458513506953

Katz Sand, I. (2015). Classification, diagnosis, and differential diagnosis of multiple sclerosis. *Current Opinion in Neurology*, *28*(3), 193–205. https://doi.org/10.1097/WCO.0000000000000206

Katz Sand, I. B., & Lublin, F. D. (2013). Diagnosis and Differential Diagnosis of Multiple Sclerosis. *Continuum*, *19*(4), 922–943. https://doi.org/10.1212/01.CON.0000433290.15468.21

Kelly, B. S., Mathur, P., McGuinness, G., Dillon, H., Lee, E. H., Yeom, K. W., Lawlor, A., & Killeen, R. P. (2024). A Radiomic “Warning Sign” of Progression on Brain MRI in Individuals with MS. *American Journal of Neuroradiology*, *45*(2), 236–243. https://doi.org/10.3174/ajnr.A8104

Korn, T. (2008). Pathophysiology of multiple sclerosis. In *Journal of Neurology* (Vol. 255, Number SUPPL. 6, pp. 2–6). https://doi.org/10.1007/s00415-008-6001-2

Kuhlmann, T., Ludwin, S., Prat, A., Antel, J., Brück, W., & Lassmann, H. (2017). An updated histological classification system for multiple sclerosis lesions. *Acta Neuropathologica*, *133*(1), 13–24. https://doi.org/10.1007/s00401-016-1653-y

Lassmann, H. (2018). Multiple Sclerosis Pathology. *Cold Spring Harbor Perspectives in Medicine*, *8*(3), a028936. https://doi.org/10.1101/cshperspect.a028936

Lassmann, H., Brück, W., & Lucchinetti, C. F. (2007). The Immunopathology of Multiple Sclerosis: An Overview. *Brain Pathology*, *17*(2), 210–218. https://doi.org/10.1111/j.1750-3639.2007.00064.x

Lavrova, E., Lommers, E., Woodruff, H. C., Chatterjee, A., Maquet, P., Salmon, E., Lambin, P., & Phillips, C. (2021). Exploratory Radiomic Analysis of Conventional vs. Quantitative Brain MRI: Toward Automatic Diagnosis of Early Multiple Sclerosis. *Frontiers in Neuroscience*, *15*. https://doi.org/10.3389/fnins.2021.679941

Lema Dopico, A., Choi, S., Hua, J., Li, X., & Harrison, D. M. (2021). Multi-layer analysis of quantitative 7 T magnetic resonance imaging in the cortex of multiple sclerosis patients reveals pathology associated with disability. *Multiple Sclerosis Journal*, *27*(13), 2040–2051. https://doi.org/10.1177/1352458521994556

Liu, Y., Mitchell, P. J., Kilpatrick, T. J., Stein, M. S., Harrison, L. C., Baker, J., Ditchfield, M., Li, K., Egan, G. F., Butzkueven, H., & Kolbe, S. C. (2012). Diffusion tensor imaging of acute inflammatory lesion evolution in multiple sclerosis. *Journal of Clinical Neuroscience*, *19*(12), 1689–1694. https://doi.org/10.1016/j.jocn.2012.03.022

Lomer, N. B., Asalemi, K. A., Saberi, A., & Sarlak, K. (2024). Predictors of multiple sclerosis progression: A systematic review of conventional magnetic resonance imaging studies. *PLOS ONE*, *19*(4), e0300415. https://doi.org/10.1371/journal.pone.0300415

Loshchilov, I., & Hutter, F. (2019). Decoupled Weight Decay Regularization. *ArXiv Preprint ArXiv:1711.05101*. http://arxiv.org/abs/1711.05101

Lucchinetti, C., Brück, W., Parisi, J., Scheithauer, B., Rodriguez, M., & Lassmann, H. (2000). Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. *Annals of Neurology*, *47*(6), 707–717. https://doi.org/10.1002/1531-8249(200006)47:6<707::AID-ANA3>3.0.CO;2-Q

Ma, Y., Zhang, C., Cabezas, M., Song, Y., Tang, Z., Liu, D., Cai, W., Barnett, M., & Wang, C. (2022). Multiple Sclerosis Lesion Analysis in Brain Magnetic Resonance Images: Techniques and Clinical Applications. *IEEE Journal of Biomedical and Health Informatics*, *26*(6), 2680–2692. https://doi.org/10.1109/JBHI.2022.3151741

McGinley, M. P., Goldschmidt, C. H., & Rae-Grant, A. D. (2021). Diagnosis and Treatment of Multiple Sclerosis. *JAMA*, *325*(8), 765. https://doi.org/10.1001/jama.2020.26858

Miller, D. H., Johnson, G., Tofts, P. S., Macmanus, D., & McDonald, W. I. (1989). Precise relaxation time measurements of normal-appearing white matter in inflammatory central nervous system disease. *Magnetic Resonance in Medicine*, *11*(3), 331–336. https://doi.org/10.1002/mrm.1910110307

Moll, N. M., Rietsch, A. M., Thomas, S., Ransohoff, A. J., Lee, J., Fox, R., Chang, A., Ransohoff, R. M., & Fisher, E. (2011). Multiple sclerosis normal-appearing white matter: Pathology–imaging correlations. *Annals of Neurology*, *70*(5), 764–773. https://doi.org/10.1002/ana.22521

Montalban, X., Lebrun-Frénay, C., Oh, J., Arrambide, G., Moccia, M., Pia Amato, M., Amezcua, L., Banwell, B., Bar-Or, A., Barkhof, F., Butzkueven, H., Ciccarelli, O., Chataway, J., Cohen, J. A., Comi, G., Correale, J., Deisenhammer, F., Filippi, M., Fiol, J., … Thompson, A. J. (2025). Diagnosis of multiple sclerosis: 2024 revisions of the McDonald criteria. *The Lancet Neurology*, *24*(10), 850–865. https://doi.org/10.1016/S1474-4422(25)00270-4

Pathak, L. (2023). Personalized Treatment for Multiple Sclerosis: The Role of Precision Medicine. *Neurology Letters*, *2*(1), 30–34. https://doi.org/10.52547/nl.2.1.30

Polman, C. H., Reingold, S. C., Banwell, B., Clanet, M., Cohen, J. A., Filippi, M., Fujihara, K., Havrdova, E., Hutchinson, M., Kappos, L., Lublin, F. D., Montalban, X., O'Connor, P., Sandberg-Wollheim, M., Thompson, A. J., Waubant, E., Weinshenker, B., & Wolinsky, J. S. (2011). Diagnostic criteria for multiple sclerosis: 2010 Revisions to the McDonald criteria. *Annals of Neurology*, *69*(2), 292–302. https://doi.org/10.1002/ana.22366

Rachmadi, M. F., del C. Valdés-Hernández, M., Makin, S., Wardlaw, J. M., & Komura, T. (2019). *Predicting the Evolution of White Matter Hyperintensities in Brain MRI Using Generative Adversarial Networks and Irregularity Map* (pp. 146–154). https://doi.org/10.1007/978-3-030-32248-9_17

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer International Publishing. https://doi.org/10.1007/978-3-319-24574-4_28

Rotstein, D., & Montalban, X. (2019). Reaching an evidence-based prognosis for personalized treatment of multiple sclerosis. *Nature Reviews Neurology*, *15*(5), 287–300. https://doi.org/10.1038/s41582-019-0170-8

Saltarelli, G., Di Cerbo, G., Innocenzi, A., De Felici, C., Splendiani, A., & Di Cesare, E. (2025). Quantitative MRI in Neuroimaging: A Review of Techniques, Biomarkers, and Emerging Clinical Applications. *Brain Sciences*, *15*(10), 1088. https://doi.org/10.3390/brainsci15101088

Seccia, R., Romano, S., Salvetti, M., Crisanti, A., Palagi, L., & Grassi, F. (2021). Machine Learning Use for Prognostic Purposes in Multiple Sclerosis. *Life*, *11*(2), 122. https://doi.org/10.3390/life11020122

Straub, S., El-Sanosy, E., Emmerich, J., Sandig, F. L., Ladd, M. E., & Schlemmer, H. (2023). Quantitative magnetic resonance imaging biomarkers for cortical pathology in multiple sclerosis at 7 T. *NMR in Biomedicine*, *36*(3). https://doi.org/10.1002/nbm.4847

Tagge, I. J., Leppert, I. R., Fetco, D., Campbell, J. S., Rudko, D. A., Brown, R. A., Stikov, N., Pike, G. B., Giacomini, P. S., Arnold, D. L., & Narayanan, S. (2022). Permanent tissue damage in multiple sclerosis lesions is associated with reduced pre-lesion myelin and axon volume fractions. *Multiple Sclerosis Journal*, *28*(13), 2027–2037. https://doi.org/10.1177/13524585221110585

Tardif, C. L., Bedell, B. J., Eskildsen, S. F., Collins, D. L., & Pike, G. B. (2012). Quantitative Magnetic Resonance Imaging of Cortical Multiple Sclerosis Pathology. *Multiple Sclerosis International*, *2012*, 1–13. https://doi.org/10.1155/2012/742018

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., Correale, J., Fazekas, F., Filippi, M., Freedman, M. S., Fujihara, K., Galetta, S. L., Hartung, H. P., Kappos, L., Lublin, F. D., Marrie, R. A., Miller, A. E., Miller, D. H., Montalban, X., … Cohen, J. A. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. In *The Lancet Neurology* (Vol. 17, Number 2, pp. 162–173). Lancet Publishing Group. https://doi.org/10.1016/S1474-4422(17)30470-2

Tranfa, M., Pontillo, G., Petracca, M., Brunetti, A., Tedeschi, E., Palma, G., & Cocozza, S. (2022). Quantitative MRI in Multiple Sclerosis: From Theory to Application. *American Journal of Neuroradiology*, *43*(12), 1688–1695. https://doi.org/10.3174/ajnr.A7536

Uddin, M. N., Figley, T. D., Kornelsen, J., Mazerolle, E. L., Helmick, C. A., O'Grady, C. B., Pirzada, S., Patel, R., Carter, S., Wong, K., Essig, M. R., Graff, L. A., Bolton, J. M., Marriott, J. J., Bernstein, C. N., Fisk, J. D., Marrie, R. A., & Figley, C. R. (2022). The comorbidity and cognition in multiple sclerosis (CCOMS) neuroimaging protocol: Study rationale, MRI acquisition, and minimal image processing pipelines. *Frontiers in Neuroimaging*, *1*. https://doi.org/10.3389/fnimg.2022.970385

van Wijnen, A., Petrov, F., Maiworm, M., Frisch, S., Foerch, C., Hattingen, E., Steinmetz, H., Klein, J. C., Deichmann, R., Wagner, M., & Gracien, R.-M. (2020). Cortical quantitative MRI parameters are related to the cognitive status in patients with relapsing-remitting multiple sclerosis. *European Radiology*, *30*(2), 1045–1053. https://doi.org/10.1007/s00330-019-06437-9

Vázquez-Marrufo, M., Sarrias-Arrabal, E., García-Torres, M., Martín-Clemente, R., & Izquierdo, G. (2023). A systematic review of the application of machine-learning algorithms in multiple sclerosis. *Neurología (English Edition)*, *38*(8), 577–590. https://doi.org/10.1016/j.nrleng.2020.10.013

Vinogradova, K., Dibrov, A., & Myers, G. (2020). Towards Interpretable Semantic Segmentation via Gradient-Weighted Class Activation Mapping (Student Abstract). *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(10), 13943–13944. https://doi.org/10.1609/aaai.v34i10.7244

Wang, J., Berger, D., Mazerolle, E., Delamer, J.-A., & Levman, J. (2022). Lesion-Specific Prediction with Discriminator-Based Supervised Guided Attention Module Enabled GANs in Multiple Sclerosis. *ArXiv Preprint ArXiv:2208.14533*. http://arxiv.org/abs/2208.14533

West, J., Aalto, A., Tisell, A., Leinhard, O. D., Landtblom, A.-M., Smedby, Ö., & Lundberg, P. (2014). Normal Appearing and Diffusely Abnormal White Matter in Patients with Multiple Sclerosis Assessed with Quantitative MR. *PLoS ONE*, *9*(4), e95161. https://doi.org/10.1371/journal.pone.0095161

Zeng, C., Gu, L., Liu, Z., & Zhao, S. (2020). Review of Deep Learning Approaches for the Segmentation of Multiple Sclerosis Lesions on Brain MRI. *Frontiers in Neuroinformatics*, *14*. https://doi.org/10.3389/fninf.2020.610967

Chapter 7

General summary

7.1 Research findings

This thesis work has demonstrated the potential of multimodal quantitative MRI (qMRI) (and conventional MRI (cMRI) to a certain extent) to characterize heterogeneity within and between MS white matter lesions (WMLs) and normal appearing white matter (NAWM). Most importantly, it shows early evidence that the MS WML spatial evolution can be predicted years into the future using only baseline MRI features.

The exploratory analysis of the qMRI metrics has shown that they are highly correlated with volumetric measures of brain features (e.g., lesions, whole brain, deep gray matter), which are widely recruited as outcome measures to inform decision-making in clinical trials (De Stefano et al., 2014). They show the most consistent association with disease progression among early clinical features (Tintore et al., 2015), and brain volume loss correlates with disease progression (Radue et al., 2015) and cognitive outcomes (T. Vollmer et al., 2016). Specific lesion (e.g., slowly expanding lesion (SEL) and central vein sign (CVS)) volume measures have demonstrated the ability to predict disabilities (Rocca et al., 2024). The findings from the first study provide a stepping stone for future qMRI implementation to possibly complement routinely collected cMRI scans to enable higher-dimensional analyses and better disease characterization. On the flip side of the same study, it is interesting to observe how much different qMRI metrics share information with each other. Although metrics within the same imaging framework are likely to exhibit information redundancy, this has only a minor effect on data collection and offline pre-processing, but it increases the computational burden when performing high-dimensional analyses. However,

there are certain metrics, such as quantitative susceptibility mapping (QSM) (myelin water fraction (MWF) in WML and T1-weighted-by-T2-weighted (T1w/T2w) ratio in NAWM to some extent), that appeared to contain relatively novel/unique information to offer other qMRI methods. Therefore, especially now that QSM demonstrates greater clinical potential – as evidenced by the inclusion of susceptibility-weighted imaging in the latest McDonald diagnostic criteria for MS (Montalban et al., 2025), this can and should be further explored.

The results from the first phase (Chapters 2 and 3) motivated a more detailed investigation of the sensitivity of qMRI metrics, specifically in and around MS WML regions, following the concentric evolution pattern observed in typical MS cases. The primary goal of the second phase (Chapter 4) was to visualize qMRI metric readings from the core of the WML to the outer NAWM and characterize them on each WML. This was done by evaluating how sensitive different qMRI metrics are to distinguishing the adjacent layers across the two tissue classes, and how these finer regions relate to one another. The statistical findings revealed that both WML and NAWM are heterogeneous (with different degrees) and showed that the signal transitions rapidly within WMLs, slowly extends outward into the NAWM, and eventually reaches a plateau (again, with similar general trends observed across different qMRI metrics). Furthermore, most correlation statistics for these qMRI signals from the concentric shells showed stronger, denser correlations within tissue classes than between-tissue-class correlations. However, although the inter-metric correlation analysis within the same shells revealed distinct patterns across different pairs of qMRI metrics, correlations in regions around the WML-perilesional NAWM (pNAWM) boundary were less likely to reach statistical significance. Those findings encourage a closer look at the WML-pNAWM regions, especially pNAWM, to better understand the entire white matter (WM) space alongside WML in MS.

As WML spatially evolves, the third phase (Chapter 5) of the thesis work moves from a cross-sectional to a longitudinal study design. The goal was to further stratify MS WM regions using longitudinal (baseline and 2-year follow-up) WML masks to examine how sensitive different qMRI metrics were at baseline in terms of their ability to potentially inform different WM changes over time. This time, a relatively naïve set of definitions of the regions was established based on the evolution pattern of the individual WML from two time points. The result of the pairwise comparisons between these ROIs is two-fold. First, many of these pairwise comparisons across different ROIs for the same metric are statistically significant, which again shows the remarkable sensitivity of qMRI metrics and highlights heterogeneity in the finer WM space from a different perspective (compared to the second phase of the thesis work). Second, and more importantly, the directionalities of these comparisons (examined after the two-tailed comparison) further strengthen the significance of these findings by matching the transition/stable WML regions and their degree of repair and neurodegeneration, as measured by baseline qMRI metrics. Together, these findings not only provide direct evidence that heterogeneity in the baseline WM space is longitudinally relevant for WML evolution (up to two years in advance) but also demonstrate the ability of qMRI metrics to potentially capture spatiotemporal information about the fate of various WM transitions (i.e., WML evolution) quite early on.

By showing that baseline MRI signals contain information (that could potentially be learned) about MS WML transitions, the previous phase provided the momentum that has translated into the final model-building stage (Chapter 6), where we assessed how well even a relatively simple 3D U-Net performs in predicting WML evolution using only information from baseline MRI scans. Both qMRI and cMRI were recruited in this study to facilitate model training and comparison across different input configurations, alongside WML and NAWM masks. The

output DEM follows a similar definition to the previous phase, which enhances the overall interpretation of the study. The comparison between the $DSC_{baseline}$ and $DSC_{modelled}$ metrics, although imperfect, shows that all models learned significant insights into spatiotemporal changes in MS WMLs. This result directly demonstrates again that qMRI metrics throughout the WM space encode information about future changes in WMLs. It extends this to show that cMRI (i.e., T1-weighted and T2-FLAIR image intensities) also demonstrated significant (comparable to qMRI) potential to forecast MS WML evolution. This is an important finding, as previously highlighted, as cMRIs included here are much more accessible than qMRIs for routine clinical workups and monitoring, not to mention that some qMRIs are not yet available in clinical settings. At the same time, since cMRI included in this study has only two images, the comparable result of the models' performance (especially with the one with combined inputs with 14 features) has raised an open question regarding whether the plateau in predictive accuracy has to do with limitations of the relatively simplistic 3D U-Net model architecture or whether it is due to limitations within the training data (in this case the baseline MRI data) in terms of what can be learned to inform WML evolution over time? Nevertheless, the positive result from this phase of the thesis work has underscored an important avenue for future research.

By combining the results throughout different phases, and with the ROI selection in the linearly connected phases of the research, shrunken from large and anatomically defined spaces (i.e., NAWM, WML and DGM) slowly to WML evolution-driven space (i.e., concentric shells, shrinking vs growing and voxel-level receptive field in 3D U-Net), this thesis has showcased that the WM spatial heterogeneity can be sensitized by different qMRI (and cMRI) metrics, creating a multi-dimensional feature space that is relevant to MS WML (or WM, in general) evolution, even looking at changes two years in advance.

7.2 Study strengths and limitations

In addition to the limitations outlined for each phase, there is room for improvement in the overall design of the work. One limitation is that the study samples are relatively small, especially in the last phase when we are building the deep learning (DL) model. In this case, the limited sample size likely reduces both the statistical significance and the overall generalizability of the result. Future work in this area should endeavour to use a larger sample and validate the findings using an external dataset. However, it is also worth noting that the main findings of the studies remain significant despite the sample size limitations. Also, MS is known to disproportionally affect women (~4 times higher than men) and young adults (mean age of diagnosis: 32 years old) in the general population (Walton et al., 2020b), with an unbalanced distribution of phenotypes highly bias to relapsing-remitting forms (85%) (vs. progressive forms) (McKay et al., 2015). The Comorbidity and Cognition in Multiple Sclerosis (CCOMS) Study was particularly designed to reflect this demographic characteristic in its recruitment of participants (i.e., without strict screening criteria regarding age, sex, and phenotype of MS), as briefly outlined previously (Section 1.8). Therefore, we acknowledge that the interpretation and the outcome of the work from this dissertation may not be directly translatable and applicable to a more focused, especially the underrepresenting, population in MS. Eventually, detailed studies should be done to investigate and validate the findings (potentially answering a totally different research question) in other MS subpopulations, especially but not limited to the progressive forms of MS. And this could even be done by using the data subsets in the CCOMS Study, despite the caveat of having a more limited data sample size. Nonetheless, as the spectrum of MS phenotypes continues to expand with discoveries in the field, the population in this dataset may be considered moderately diverse and

heterogeneous. By pooling data regardless of demographic background and phenotypes, and allowing for some levels of latent heterogeneity that are not captured in the dataset, the study's overall findings indicate a significant degree of generalizability of the disease within MRI, alongside personalization.

Nevertheless, with greater availability of the resource and the growing initiative in multi-institutional data collection protocols (e.g., The Canadian prospective cohort study to understand progression in multiple sclerosis (CanProCo): rationale, aims, and study design (J. Oh et al., 2021)) in research nowadays, it may be easier to acquire a dataset with a similar demographic profile, a much larger sample size, and more advanced qMRI sequences. Unfortunately, this dataset does not have any biospecimens (e.g., lumbar puncture samples or lab charts) collected. And those would be instrumental in understanding the relationship between MRI features, not only limited to the WM landscape, with MS pathophysiology.

Another consideration is that analyses can be done in template space or subject space. The former addresses differences in brain size between participants, and controls for within-participant changes (e.g., global and/or regional atrophy) over time. However, this puts spatial/distance measures into a relative scale – as opposed to subject-space analyses, which would preserve absolute spatial/distance measures. It can be argued that, in the end, forward and backward deformation can be applied to transform between the spaces, but the statistical interpretation differs because the tests are run in different locations, with different ROI sizes and spatial distances. The differences between the spaces of analysis are not necessarily a limitation but rather a choice that was made, ultimately leading to the final phase (Chapter 6) where we felt that the separation of DEM class required longitudinal maps/masks from each participant to be somehow put into a common space, and for maps/masks from different participants to then be put into a common space

prior to input into the 3D U-Net model and facilitate training. Given also the fact that the method used to normalize MRI images (that facilitates data analyses in all chapters in this dissertation) has been systematically validated in the context of MS (Pirzada et al., 2020), the decision was made to effectively do both at once by non-linearly warping all MRI images and masks to the MNI-152 template.

This thesis work focuses exclusively on MS WMLs, lesions in the WM space of the brain. While this approach helped us narrow our scope, our analyses did not consider the interplay with other parts of the central nervous system, such as the gray matter (GM) (e.g., cortical surface), the brain stem, or the spinal cord. MS lesions in other regions of the brain, particularly cortical GM lesions, have been shown to have a high specificity and clinical relevance (Absinta et al., 2012; Kolb et al., 2022). However, cortical lesions are generally harder to image due to the need for high resolution (Bagnato et al., 2024; Kolb et al., 2022; Rocca et al., 2024). Nonetheless, future studies could use our studies on MS WMLs as a template to investigate cortical GM lesions, and to determine whether baseline MRI metrics can be used along with machine learning methods to predict patterns of MS cortical lesion evolution.

WML masks may differ when different imaging sequences are used because not all T1-weighted WML are visible on T2-weighted images, and around 50% of T2 lesions appear as T1 lesions (Valcarcel et al., 2018). Because WML masks were segmented from T2-FLAIR images, the results and discussion here are confined to T2 WMLs. But given the fact that T1 WMLs (or T1 hypointensities) may still be structurally different from T2 WMLs (e.g., experiencing more extensive axonal loss and chronic activities) (Sahraian et al., 2009b), this introduces a caveat with an additional degree of heterogeneity when interpreting the findings throughout this thesis.

But this does not necessarily mean that the T1 WML regions are virtually meaningless at different phases of the research, as phase 2 (Chapter 4) has revealed not only a relatively greater degree of heterogeneity in WML regions (vs. NAWM), but also the subtle heterogeneity in NAWM regions (both of which might be due to the co-existence of and the differences between T2 WML and T1 WML, albeit this was not explicitly tested or discussed in the Chapter 4). However, the interplay between the two types of WML in this context cannot be examined until T1 WML is segmented. While one study demonstrates the transferability of an automatic segmentation tool between T1 and T2 modalities (Valcarcel et al., 2018), lesion segmentation remains an active field of research, especially with the growing accessibility and rapid advancement of artificial intelligence (AI) (Diaz-Hurtado et al., 2022; Thabet et al., 2025). However, inconsistent use of the algorithm creates a similar issue of compatibility and changes the statistical findings across studies when the ROI definition differs. Nevertheless, the robustness of the T2 lesion segmentation for this work was briefly highlighted in Section 1.8. This has a strong implication for the reliability of the result, even though the sample sizes are small in some phases.

The last highlighted aspect is that this thesis work, paired with the dataset (phenotype-agnostic participant recruitment), does not distinguish between relapsing-remitting, secondary progressive, or primary progressive phenotypes within the cohort or between WML types (when individual WMLs were separated). However, MS is a heterogeneous disease, in which phenotypes are treated differently within the conventional disease framework. In addition, perhaps contradicting with the co-varying structural findings in healthy brains, for example in white matter tracts, cortical gray matter, and cortical thickness (Alexander-Bloch et al., 2013; Wahl et al., 2010), WMLs can be heterogeneous in morphologies, stages and evolutionary pathways, as suggested by MRI findings (Harrison et al., 2016; Thaler et al., 2021, Chapters 4 and 5) and histological findings

(Lucchinetti et al., 2000). On one hand, the positive results across different MS phenotypes and WML types indicate that MRI data contains information that (at least to some extent) may generalize both across and within individuals. On the other hand, this could limit our findings, since pooling heterogeneous data may reduce statistical power when participants or each type of WML exhibit lower similarities. Therefore, while pooling WMLs (even within a participant) can be a valid approach for statistical analysis, interesting research questions such as how different WMLs relate to each other at the participant level, how WML heterogeneity alters the original WM network, and how the above relate to different clinical phenotypes and disability remain to be addressed (ideally with a dataset that has a more equally weighted distribution of disease phenotypes and more finely annotated WML subtypes). But acquiring such a dataset can be laborious and resource-intensive, as we need to consider characterizing more phenotypes, such as radiologically and clinically isolated syndromes, and growing evidence have shown that MS can be a spectrum rather than a discrete classification (Giovannoni, 2018; Lublin et al., 2014; T. L. Vollmer et al., 2021). Second, the recognition of WML types has also advanced with recent MRI methods, highlighting the requirement of a more rigorous framework to identify all possible WML types. This may include recruiting domain experts and acquiring additional imaging features in the short term; however, with future research, this could be another application where AI could become quite useful. Moreover, MS progression over time, which allows these types and categories to evolve dynamically, makes longitudinal analysis and the development of a research framework much more complex. Nevertheless, these are ultimately important questions to understand before formulating a comprehensive radiological association with the disease, thereby guiding better clinical decisions and having a greater impact on patient outcomes.

7.3 Future outlook

WM is heterogeneous, as shown by different qMRI metrics in this dissertation, particularly in and around MS WMLs. As the clinical use of MRI remains largely qualitative and is limited in its scope for binarized comparisons between WML and NAWM, this thesis used a series of analyses to highlight the clinical potential of extracting finer features across the entire WM space using a multimodal qMRI (and cMRI) approach. Given the significance and limitations as listed in previous sections, here are some suggested avenues that could be explored to further elucidate and improve upon our findings:

This dissertation has showcased two distinct patterns of WM heterogeneity in MS: 1) WML morphology-driven (Chapter 4), 2) WM tissue transition-driven (Chapter 5). Both scenarios can be detected using qMRI metrics. Nonetheless, WM heterogeneity can be attributed to many other inherent factors, such as anatomical preference – i.e., since WML are more likely to be found near the juxtacortical and periventricular regions (Filippi et al., 2019), and other recent work from our lab has also shown WMLs disproportionately affect white matter tracts in certain functionally-connected brain networks (Figley et al., 2026), and physiological variability – as well as the activities and evolutionary changes in different WML subtypes (Absinta et al., 2016). Given the demonstrated high sensitivity of qMRI in previous chapters in this dissertation, future work should consider employing qMRI to continue the exploratory journey toward better understanding and mapping WM heterogeneity both cross-sectionally and longitudinally. Potential directions include, but are not limited to: 1) the refinement of ROIs, since the results depend on how the ROIs are defined and computed, and what is the voxel inclusion/exclusion criteria; 2) the spatial distribution of MWLs, and whether their anatomical location (e.g., juxtacortical, periventricular, within specific white matter tracts) affects the ability to accurately measure and/or predict their

evolutionary patterns; and 3) the stratification of WML subtypes, and the extent to which findings differ between chronically active, inactive and acute WMLs (and whether, for example, lesion evolution prediction accuracy differs for paramagnetic rim lesions, WMLs with central vein sign, etc.). The exploration will most likely be significant and impactful, leading to a more well-rounded understanding of WM heterogeneity in MS.

The lack of feature inclusion beyond qMRI (and cMRI) is something that we did not have time to explore, but is likely to have tremendous value. For example, this dissertation did not investigate how MS clinical presentations, types of symptoms, severity and progression status are related to qMRI metrics across different tissue classes (Chapters 2 and 3), the heterogeneity in WM that is sensitized by qMRI metrics (Chapters 4 and 5), or the WML evolution pattern (Chapters 5 and 6). Since the clinic-radiological paradox is not yet resolved (Barkhof, 2002; Chard & Trip, 2017; Ojha et al., 2025), future work should incorporate the relevant clinical measures to broaden and deepen the understanding of the findings from this thesis. Although MRI findings such as lesion volume, number of lesions and brain volume are often used as treatment response measures for MS (Ananthavarathan et al., 2024; Gasperini et al., 2019) and imaging outcomes in clinical trials for drug development (Tur et al., 2018), our findings and those from others (Elahi et al., 2025, Chapters 2-5) suggest that cross-sectional and longitudinal effects of interventional procedures and disease modifying therapies (DMTs) should be monitored with a suitable combination of qMRI measures that are able to more sensitively and/or specifically elucidate changes in the underlying WM pathologies (McGinley et al., 2021).

It is worth noting that the CCOMS dataset includes a comprehensive battery of MRI scans (e.g., perfusion and functional MRI) and a variety of demographic, clinical and lab measures. Historically, clinical tests, lab results, and patient demographic information have been collected

alongside MRI to compile a comprehensive profile of an individual with MS and to help clinicians and researchers better understand the disease and make better clinical decisions. With dedication and advances, more refinements have been implemented in the diagnostic workup and the monitoring protocol for MS (Montalban et al., 2025; Voigt et al., 2024). In particular, more and more subcellular-level markers and measures (e.g., biomarker and -omic data) have been discovered, adding unique and detailed perspectives to neuroscience research (O'Connor et al., 2023), and shedding light on the field of MS (Lorefice et al., 2023). In addition, as the world moves toward full digitalization of data collection and storage, data accessibility may be greatly enhanced. Although it facilitates a more robust multidimensional analysis (by adding more MRI data or pairing with other data sources) to achieve specific (and potentially more challenging) goals such as lesion evolution prediction, the findings will not always be easy to interpret, whether or not the results are positive. Therefore, rigorous study designs are important. And data analysis should be carefully planned to systematically characterize the relationships among data sources, particularly when aiming for clinical implementation (e.g., generalizability and interpretability).

cMRI can already be quite powerful, given its validity and popularity in clinical settings for MS. Near the end of the thesis, the significance of cMRI findings motivates ongoing research to enhance the quantitative use of cMRI alongside qMRI. However, cMRI images require processing to normalize image signal intensities, especially for diseases like MS, where widespread abnormalities are found across many brain regions. Beyond the naïve whole-brain intensity normalization method used in this thesis, other sophisticated normalization methods have been proposed in the literature for MS (Brier et al., 2021; Roy et al., 2013; Shah et al., 2011). However, regardless of MRI modalities, differences in the sequence parameters and post-processing pipeline are always significant sources for the variability in the image quality and

information therefore impacting analytical results and interpretation (Biberacher et al., 2016; Shinohara et al., 2017). This is also extremely relevant when translating the study to other centers either to independently validate the work and enhance the quality of the results or to collaboratively engage in multicenter studies. The proper implementation of the methodological framework of data collection plays an important role, yet it is a non-trivial task. Recommendations on the selection of MRI sequences/modalities, including their associated imaging parameters, radiological features relevant to MS, analytical methods and data storage in large multi-center settings (De Stefano et al., 2022), and recommendations on MRI usage in MS diagnosis, prognosis and monitoring applications have been previously proposed by international groups of researchers and clinicians (Wattjes et al., 2021). Nonetheless, harmonized data acquisition is only part of multicenter study designs (Ashton, 2010), so two other equally important aspects must be considered in multimodal MRI studies at different centers (and even longitudinally within the same center): 1) Site selection, both in terms of the availability/compatibility of different MRI hardware across sites (and their ability to acquire the same kinds of data, ideally with the same or similar pulse sequence parameters and similar image quality) and; 2) Image processing and data analysis pipelines should be centralized or carefully harmonized to ensure consistent post-processing using the same software (and software versions) with the same settings. Then, if any changes are necessary to the data acquisition or analysis protocols, images should be systematically compared before and after the modification, so that an informed decision can be made regarding the continuation of the study, the change(s) should be explicitly noted in any resulting publications, and the details of any mixed datasets (pooling the new and old data) should be thoroughly described. Beyond the imaging-related factors, some other practical considerations in terms of logistics and human resources (e.g., staff training, data management, finances, communication)

have been previously described (Das, 2022). In spite of the challenges, translating this study design to a multi-center setting would increase the robustness of the results and strengthen the findings.

In addition, with a relatively small sample size, the potential complementarity between cMRI and qMRI remains underexplored, and the amount of information shared between them is yet to be determined. Future work can deepen this understanding by delving into the biophysics underlying the modalities, by employing empirical designs to examine them at the surface level, or by using machine learning approaches to determine whether maps can be generated from their counterparts. The latter ties back to the need for feature selection, considering the challenges in multi-center studies and real-world data acquisition constraints (e.g., time and clinical workload). Even if the qMRI and cMRI metrics are all deemed important and exhibit a low level of redundancy, collecting all images in the real-world clinical setting would be costly, could potentially cause patient discomfort, and lead to even longer MRI wait times due to increased scan time. In parallel with technological advancements and the emergence of newly developed qMRI maps, we as data scientists should also focus on refining sensitivity and specificity through interpretation of the metrics, with a continually updated understanding of MS, and therefore compile a hierarchy system to help the field identify metrics with high clinical practicality and relevance in the real-world setting more quickly.

In conclusion, even though more work remains to interpret the clinical value of quantifying heterogeneity in the WM space in MS, this thesis work, using a multi-modal approach and qMRI metrics, has already shown that heterogeneity in the WM space in MS is not completely arbitrary, but rather spatiotemporally relevant, which even showed an early promise in lesion evolution prediction when pairing with AI and DL. These findings may all eventually help advance understanding of the disease, enabling a finer stratification of pwMS beyond conventional

phenotyping, guiding better clinical trials, improving prognosis to enable precision medicine approaches, and ultimately improving the quality of life for pwMS.

7.4 References

Absinta, M., Rocca, M. A., Colombo, B., Copetti, M., Feo, D., Falini, A., Comi, G., & Filippi, M. (2012). Patients with migraine do not have MRI-visible cortical lesions. *Journal of Neurology*, *259*(12), 2695–2698. https://doi.org/10.1007/s00415-012-6571-x

Absinta, M., Sati, P., & Reich, D. S. (2016). Advanced MRI and staging of multiple sclerosis lesions. *Nature Reviews Neurology*, *12*(6), 358–368. https://doi.org/10.1038/nrneurol.2016.59

Alexander-Bloch, A., Giedd, J. N., & Bullmore, E. (2013). Imaging structural co-variance between human brain regions. *Nature Reviews Neuroscience*, *14*(5), 322–336. https://doi.org/10.1038/nrn3465

Ananthavarathan, P., Sahi, N., & Chard, D. T. (2024). An update on the role of magnetic resonance imaging in predicting and monitoring multiple sclerosis progression. *Expert Review of Neurotherapeutics*, *24*(2), 201–216. https://doi.org/10.1080/14737175.2024.2304116

Ashton, E. (2010). Quantitative MR in multi-center clinical trials. *Journal of Magnetic Resonance Imaging*, *31*(2), 279–288. https://doi.org/10.1002/jmri.22022

Bagnato, F., Sati, P., Hemond, C. C., Elliott, C., Gauthier, S. A., Harrison, D. M., Mainero, C., Oh, J., Pitt, D., Shinohara, R. T., Smith, S. A., Trapp, B., Azevedo, C. J., Calabresi, P. A., Henry, R. G., Laule, C., Ontaneda, D., Rooney, W. D., Sicotte, N. L., … Absinta, M. (2024). Imaging chronic active lesions in multiple sclerosis: a consensus statement. *Brain*, *147*(9), 2913–2933. https://doi.org/10.1093/brain/awae013

Barkhof, F. (2002). The clinico-radiological paradox in multiple sclerosis revisited. *Current Opinion in Neurology*, *15*(3), 239–245. https://doi.org/10.1097/00019052-200206000-00003

Biberacher, V., Schmidt, P., Keshavan, A., Boucard, C. C., Righart, R., Sämann, P., Preibisch, C., Fröbel, D., Aly, L., Hemmer, B., Zimmer, C., Henry, R. G., & Mühlau, M. (2016). Intra- and interscanner variability of magnetic resonance imaging based volumetry in multiple sclerosis. *NeuroImage*, *142*, 188–197. https://doi.org/10.1016/j.neuroimage.2016.07.035

Brier, M. R., Xiang, B., Li, Z., Naismith, R. T., Yablonskiy, D., Cross, A. H., Benzinger, T. L. S., & Snyder, A. Z. (2021). Signal Intensity Normalization of Multispectral Magnetic Resonance Images Obtained with Clinical Sequences. In *medRxiv* (pp. 2010–2021). Cold Spring Harbor Laboratory Press. https://doi.org/10.1101/2021.10.27.21265570

Cercignani, M., & Bouyagoub, S. (2018). Brain microstructure by multi-modal MRI: Is the whole greater than the sum of its parts? *NeuroImage*, *182*, 117–127. https://doi.org/10.1016/j.neuroimage.2017.10.052

Chard, D., & Trip, S. A. (2017). Resolving the clinico-radiological paradox in multiple sclerosis. *F1000Research*, *6*, 1828. https://doi.org/10.12688/f1000research.11932.1

Das, M. K. (2022). Multicenter Studies: Relevance, Design and Implementation. *Indian Pediatrics*, *59*(7), 571–579. https://doi.org/10.1007/s13312-022-2561-y

De Stefano, N., Airas, L., Grigoriadis, N., Mattle, H. P., O'Riordan, J., Oreja-Guevara, C., Sellebjerg, F., Stankoff, B., Walczak, A., Wiendl, H., & Kieseier, B. C. (2014). Clinical Relevance of Brain Volume Measures in Multiple Sclerosis. *CNS Drugs*, *28*(2), 147–156. https://doi.org/10.1007/s40263-014-0140-z

De Stefano, N., Battaglini, M., Pareto, D., Cortese, R., Zhang, J., Oesingmann, N., Prados, F., Rocca, M. A., Valsasina, P., Vrenken, H., Gandini Wheeler-Kingshott, C. A. M., Filippi, M., Barkhof, F., & Rovira, À. (2022). MAGNIMS recommendations for harmonization of MRI data in MS multicenter studies. *NeuroImage: Clinical*, *34*, 102972. https://doi.org/10.1016/j.nicl.2022.102972

Diaz-Hurtado, M., Martínez-Heras, E., Solana, E., Casas-Roma, J., Llufriu, S., Kanber, B., & Prados, F. (2022). Recent advances in the longitudinal segmentation of multiple sclerosis lesions on magnetic resonance imaging: a review. *Neuroradiology*, *64*(11), 2103–2117. https://doi.org/10.1007/s00234-022-03019-3

Elahi, R., Taremi, S., Najafi, A., Karimi, H., Asadollahzadeh, E., Sajedi, S. A., Rad, H. S., & Sahraian, M. A. (2025). Advanced MRI Methods for Diagnosis and Monitoring of Multiple Sclerosis (MS). *Journal of Magnetic Resonance Imaging*, *62*(6), 1546–1578. https://doi.org/10.1002/jmri.29817

Figley, T. D., Kornelsen, J., Uddin, M. N., Wong, K., Pirzada, S., Carter, S., Helmick, C. A., O'Grady, C. B., Mazerolle, E. L., Patel, R., Bernstein, C. N., Fisk, J. D., Marrie, R. A., & Figley, C. R. (2026). White Matter Damage in Multiple Sclerosis Disproportionately Targets Default Mode, Executive Control, and Salience Networks. *The Journal of Neuroscience*, *46*(22), e0278252026. https://doi.org/10.1523/JNEUROSCI.0278-25.2026

Filippi, M., Preziosa, P., Banwell, B. L., Barkhof, F., Ciccarelli, O., De Stefano, N., Geurts, J. J. G., Paul, F., Reich, D. S., Toosy, A. T., Traboulsee, A., Wattjes, M. P., Yousry, T. A., Gass, A., Lubetzki, C., Weinshenker, B. G., & Rocca, M. A. (2019). Assessment of lesions on magnetic resonance imaging in multiple sclerosis: practical guidelines. In *Brain* (Vol. 142, Number 7, pp. 1858–1875). Oxford University Press. https://doi.org/10.1093/brain/awz144

Gasperini, C., Prosperini, L., Tintoré, M., Sormani, M. P., Filippi, M., Rio, J., Palace, J., Rocca, M. A., Ciccarelli, O., Barkhof, F., Sastre-Garriga, J., Vrenken, H., Frederiksen, J. L., Yousry, T. A., Enzinger, C., Rovira, A., Kappos, L., Pozzilli, C., Montalban, X., & De Stefano, N. (2019). Unraveling treatment response in multiple sclerosis. *Neurology*, *92*(4), 180–192. https://doi.org/10.1212/WNL.0000000000006810

Giovannoni, G. (2018). Disease-modifying treatments for early and advanced multiple sclerosis: a new treatment paradigm. *Current Opinion in Neurology*, *31*(3), 233–243. https://doi.org/10.1097/WCO.0000000000000561

Harrison, D. M., Li, X., Liu, H., Jones, C. K., Caffo, B., Calabresi, P. A., & van Zijl, P. (2016). Lesion Heterogeneity on High-Field Susceptibility MRI Is Associated with Multiple Sclerosis Severity. *American Journal of Neuroradiology*, *37*(8), 1447–1453. https://doi.org/10.3174/ajnr.A4726

Kolb, H., Al-Louzi, O., Beck, E. S., Sati, P., Absinta, M., & Reich, D. S. (2022). From pathology to MRI and back: Clinically relevant biomarkers of multiple sclerosis lesions. *NeuroImage: Clinical*, *36*, 103194. https://doi.org/10.1016/j.nicl.2022.103194

Lorefice, L., Pitzalis, M., Murgia, F., Fenu, G., Atzori, L., & Cocco, E. (2023). Omics approaches to understanding the efficacy and safety of disease-modifying treatments in multiple sclerosis. *Frontiers in Genetics*, *14*. https://doi.org/10.3389/fgene.2023.1076421

Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sørensen, P. S., Thompson, A. J., Wolinsky, J. S., Balcer, L. J., Banwell, B., Barkhof, F., Bebo, B., Calabresi, P. A., Clanet, M., Comi, G., Fox, R. J., Freedman, M. S., Goodman, A. D., Inglese, M., Kappos, L., … Polman, C. H. (2014). Defining the clinical course of multiple sclerosis. *Neurology*, *83*(3), 278–286. https://doi.org/10.1212/WNL.0000000000000560

Lucchinetti, C., Brück, W., Parisi, J., Scheithauer, B., Rodriguez, M., & Lassmann, H. (2000). Heterogeneity of multiple sclerosis lesions: Implications for the pathogenesis of demyelination. *Annals of Neurology*, *47*(6), 707–717. https://doi.org/10.1002/1531-8249(200006)47:6<707::AID-ANA3>3.0.CO;2-Q

McGinley, M. P., Goldschmidt, C. H., & Rae-Grant, A. D. (2021). Diagnosis and Treatment of Multiple Sclerosis. *JAMA*, *325*(8), 765. https://doi.org/10.1001/jama.2020.26858

McKay, K. A., Kwan, V., Duggan, T., & Tremlett, H. (2015). Risk Factors Associated with the Onset of Relapsing-Remitting and Primary Progressive Multiple Sclerosis: A Systematic Review. *BioMed Research International*, *2015*(1), 1–11. https://doi.org/10.1155/2015/817238

Montalban, X., Lebrun-Frénay, C., Oh, J., Arrambide, G., Moccia, M., Pia Amato, M., Amezcua, L., Banwell, B., Bar-Or, A., Barkhof, F., Butzkueven, H., Ciccarelli, O., Chataway, J., Cohen, J. A., Comi, G., Correale, J., Deisenhammer, F., Filippi, M., Fiol, J., … Thompson, A. J. (2025). Diagnosis of multiple sclerosis: 2024 revisions of the McDonald criteria. *The Lancet Neurology*, *24*(10), 850–865. https://doi.org/10.1016/S1474-4422(25)00270-4

O'Connor, L. M., O'Connor, B. A., Lim, S. Bin, Zeng, J., & Lo, C. H. (2023). Integrative multi-omics and systems bioinformatics in translational neuroscience: A data mining perspective. *Journal of Pharmaceutical Analysis*, *13*(8), 836–850. https://doi.org/10.1016/j.jpha.2023.06.011

Oh, J., Arbour, N., Giuliani, F., Guenette, M., Kolind, S., Lynd, L., Marrie, R. A., Metz, L. M., Patten, S. B., Prat, A., Schabas, A., Smyth, P., Tam, R., Traboulsee, A., & Yong, V. W. (2021). The Canadian prospective cohort study to understand progression in multiple sclerosis (CanProCo): rationale, aims, and study design. *BMC Neurology*, *21*(1), 418. https://doi.org/10.1186/s12883-021-02447-7

Ojha, A., Tommasin, S., Piervincenzi, C., Baione, V., Gangemi, E., Gallo, A., d'Ambrosio, A., Altieri, M., De Stefano, N., Cortese, R., Valsasina, P., Tedone, N., Pozzilli, C., Rocca, M. A., Filippi, M., Pantano, P., Petsas, N., Giannì, C., Storelli, L., … Esposito, F. (2025). Clinical and MRI features contributing to the clinico-radiological dissociation in a large cohort of people with multiple sclerosis. *Journal of Neurology*, *272*(5), 327. https://doi.org/10.1007/s00415-025-12977-6

Pirzada, S., Uddin, M. N., Figley, T. D., Kornelsen, J., Puig, J., Marrie, R. A., Mazerolle, E. L., Fisk, J. D., Helmick, C. A., O'Grady, C. B., Patel, R., & Figley, C. R. (2020). Spatial normalization of multiple sclerosis brain MRI data depends on analysis method and software package. *Magnetic Resonance Imaging*, *68*, 83–94. https://doi.org/10.1016/j.mri.2020.01.016

Radue, E.-W., Barkhof, F., Kappos, L., Sprenger, T., Häring, D. A., de Vera, A., von Rosenstiel, P., Bright, J. R., Francis, G., & Cohen, J. A. (2015). Correlation between brain volume loss and clinical and MRI outcomes in multiple sclerosis. *Neurology*, *84*(8), 784–793. https://doi.org/10.1212/WNL.0000000000001281

Rocca, M. A., Preziosa, P., Barkhof, F., Brownlee, W., Calabrese, M., De Stefano, N., Granziera, C., Ropele, S., Toosy, A. T., Vidal-Jordana, À., Di Filippo, M., & Filippi, M. (2024). Current and future role of MRI in the diagnosis and prognosis of multiple sclerosis. *The Lancet Regional Health - Europe*, *44*, 100978. https://doi.org/10.1016/j.lanepe.2024.100978

Roy, S., Carass, A., Shiee, N., Pham, D. L., Calabresi, P., Reich, D., & Prince, J. L. (2013). Longitudinal intensity normalization in the presence of multiple sclerosis lesions. *2013 IEEE 10th International Symposium on Biomedical Imaging*, 1384–1387. https://doi.org/10.1109/ISBI.2013.6556791

Sahraian, M. A., Radue, E.-W., Haller, S., & Kappos, L. (2009). Black holes in multiple sclerosis: definition, evolution, and clinical correlations. *Acta Neurologica Scandinavica*, *122*(1), 1–8. https://doi.org/10.1111/j.1600-0404.2009.01221.x

Shah, M., Xiao, Y., Subbanna, N., Francis, S., Arnold, D. L., Collins, D. L., & Arbel, T. (2011). Evaluating intensity normalization on MRIs of human brain with multiple sclerosis. *Medical Image Analysis*, *15*(2), 267–282. https://doi.org/10.1016/j.media.2010.12.003

Shinohara, R. T., Oh, J., Nair, G., Calabresi, P. A., Davatzikos, C., Doshi, J., Henry, R. G., Kim, G., Linn, K. A., Papinutto, N., Pelletier, D., Pham, D. L., Reich, D. S., Rooney, W., Roy, S., Stern, W., Tummala, S., Yousuf, F., Zhu, A., … Bakshi, R. (2017). Volumetric Analysis from a Harmonized Multisite Brain MRI Study of a Single Subject with Multiple Sclerosis. *American Journal of Neuroradiology*, *38*(8), 1501–1509. https://doi.org/10.3174/ajnr.A5254

Thabet, R. M., Shedeed, H. A., Al-Berry, M., & Khattab, D. (2025). Multiple sclerosis classification and segmentation in neuroimaging MRI using different machine and deep learning techniques: a review. *Artificial Intelligence Review*, *58*(8), 255. https://doi.org/10.1007/s10462-025-11143-8

Thaler, C., Kyselyova, A. A., Faizy, T. D., Nawka, M. T., Jespersen, S., Hansen, B., Stellmann, J.-P., Heesen, C., Stürner, K. H., Stark, M., Fiehler, J., Bester, M., & Gellißen, S. (2021). Heterogeneity of multiple sclerosis lesions in fast diffusional kurtosis imaging. *PLOS ONE*, *16*(2), e0245844. https://doi.org/10.1371/journal.pone.0245844

Tintore, M., Rovira, À., Río, J., Otero-Romero, S., Arrambide, G., Tur, C., Comabella, M., Nos, C., Arévalo, M. J., Negrotto, L., Galán, I., Vidal-Jordana, A., Castilló, J., Palavra, F., Simon, E., Mitjana, R., Auger, C., Sastre-Garriga, J., & Montalban, X. (2015). Defining high, medium and

low impact prognostic factors for developing multiple sclerosis. *Brain*, *138*(7), 1863–1874. https://doi.org/10.1093/brain/awv105

Tur, C., Moccia, M., Barkhof, F., Chataway, J., Sastre-Garriga, J., Thompson, A. J., & Ciccarelli, O. (2018). Assessing treatment outcomes in multiple sclerosis trials and in the clinical setting. *Nature Reviews Neurology*, *14*(2), 75–93. https://doi.org/10.1038/nrneurol.2017.171

Valcarcel, A. M., Linn, K. A., Khalid, F., Vandekar, S. N., Tauhid, S., Satterthwaite, T. D., Muschelli, J., Martin, M. L., Bakshi, R., & Shinohara, R. T. (2018). A dual modeling approach to automatic segmentation of cerebral T2 hyperintensities and T1 black holes in multiple sclerosis. *NeuroImage: Clinical*, *20*, 1211–1221. https://doi.org/10.1016/j.nicl.2018.10.013

Voigt, I., Fischer, S., Proschmann, U., Konofalska, U., Richter, P., Schlieter, H., Berger, T., Meuth, S. G., Hartung, H.-P., Akgün, K., & Ziemssen, T. (2024). Consensus quality indicators for monitoring multiple sclerosis. *The Lancet Regional Health - Europe*, *40*, 100891. https://doi.org/10.1016/j.lanepe.2024.100891

Vollmer, T., Huynh, L., Kelley, C., Galebach, P., Signorovitch, J., DiBernardo, A., & Sasane, R. (2016). Relationship between brain volume loss and cognitive outcomes among patients with multiple sclerosis: a systematic literature review. *Neurological Sciences*, *37*(2), 165–179. https://doi.org/10.1007/s10072-015-2400-1

Vollmer, T. L., Nair, K. V., Williams, I. M., & Alvarez, E. (2021). Multiple Sclerosis Phenotypes as a Continuum. *Neurology Clinical Practice*, *11*(4), 342–351. https://doi.org/10.1212/CPJ.0000000000001045

Wahl, M., Li, Y.-O., Ng, J., LaHue, S. C., Cooper, S. R., Sherr, E. H., & Mukherjee, P. (2010). Microstructural correlations of white matter tracts in the human brain. *NeuroImage*, *51*(2), 531–541. https://doi.org/10.1016/j.neuroimage.2010.02.072

Walton, C., King, R., Rechtman, L., Kaye, W., Leray, E., Marrie, R. A., Robertson, N., La Rocca, N., Uitdehaag, B., van der Mei, I., Wallin, M., Helme, A., Angood Napier, C., Rijke, N., & Baneke, P. (2020). Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Multiple Sclerosis Journal*, *26*(14), 1816–1821. https://doi.org/10.1177/1352458520970841

Wattjes, M. P., Ciccarelli, O., Reich, D. S., Banwell, B., de Stefano, N., Enzinger, C., Fazekas, F., Filippi, M., Frederiksen, J., Gasperini, C., Hacohen, Y., Kappos, L., Li, D. K. B., Mankad, K., Montalban, X., Newsome, S. D., Oh, J., Palace, J., Rocca, M. A., … Oh, J. (2021). 2021 MAGNIMS–CMSC–NAIMS consensus recommendations on the use of MRI in patients with multiple sclerosis. *The Lancet Neurology*, *20*(8), 653–670. https://doi.org/10.1016/S1474-4422(21)00095-8

Appendix. Contributions of Authors

For Chapters 2 and 3, CRF (Chase R. Figley) and KW (Kaihim Wong) conceived the idea of the study, designed the bulk structure of the analysis, and defined the scope of the work.

For Chapters 4, 5 and 6 KW conceived the idea of the study and designed the bulk structure of the analysis. CRF and KW defined the scope of the work.

For all Chapters, KW conducted the statistical analyses. KW drafted the manuscripts. KW, Md Nasir Uddin, Teresa D. Figley, Jennifer Kornelsen, Carl A. Helmick, Christopher B. O'Grady, Erin L. Mazerolle, Ronak Patel, Charles N. Bernstein, John D. Fisk, Ruth Ann Marrie and CRF will contribute to the revision, and everyone will review the revisions and approve the manuscript before journal submission for publication.