USING MAGNETIC RESONANCE IMAGING FOR THE IN VIVO DETECTION AND CHARACTERIZATION OF CEREBRAL GREY MATTER PATHOLOGY IN PATIENTS WITH MULTIPLE SCLEROSIS

Mishkin Derakhshan

Department of Biomedical Engineering
Faculty of Medicine
McGill University
Montreal, Canada

December 2012

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

To Chobo: Feels like the end of a lifetime.

I suppose it is tempting, if the only tool you have is a hammer, to treat everything as if it were a nail.

—The Law of the Instrument, Abraham Maslow

CONTENTS

ILLUSTRATIONS

FIGURES

TABLES

ACKNOWLEDGEMENTS

First and foremost I would like to thank my doctoral advisors, D. Louis Collins and Douglas L. Arnold. Doug, for affording me opportunities and presenting me with challenges that no one else could have, and Louis, for helping me find the solutions to all the challenges I faced along the way. I am grateful for the wisdom on science, medicine, and life that you have both imparted to me; it has helped shape me into the person I am today.

I had the benefit of having a supportive and affable supervisory committee. In addition to my supervisors, I would like to thank Sridar Narayanan, Kaleem Siddiqi, and Bruce Pike for their insightful discussions, guidance, and encouragement. Bruce once said the average Ph.D. student thinks they are one year away from submitting their thesis for the last three years of their degree. I was no exception.

A great many other people have also helped make my time at the BIC so enjoyable. To old man Simon and Arieh (and their wives): it has been most entertaining. To Jacqueline: for the tutelage, mentorship, and in general for being a role model for the scientific process. To Alex and Sergey: for accommodating my many technical needs. To Jennifer and Ellie: for always taking care of me. To Aki: for being Aki and doing the many things Aki does. Every research unit needs an Aki. To David, Josefina, and Paul: for the invaluable perspectives they lent to the conversations. To Gabriel, Line, and Rozie: for all the help with the CLMS project. To Kelvin, Pierre, Nicolas, and Illana: for great friendships from the start. To Mallar and Ives: my older brothers in science to whom I will surely turn to in the future.

I am greatly indebted to the members of the MINC mailing list who have displayed much patience with me over the years: Andrew Janke, Claude Lepage, Jason Lerch, and, in particular, Vladimir Fonov. Thank you for taking the time.

I am also indebted to Ives Levesque for providing the French translation of the abstract of this thesis.

I would like to acknowledge financial support received from the Neuroinflammation Training Program, the endMS Training Network, and the NSERC/MITACS joint Industrial Postgraduate Scholarship with NeuroRx Research.

I wish to also thank my family, who encouraged me to pursue my education to its fullest and have supported me along the way.

Finally, this thesis is dedicated to Amanda for seeing me through to the seemingly interminable end. Thank you for the unwavering support, love, and editing.

Thanks!

Mishkin Derakhshan
December 8, 2012

COMMON ABBREVIATIONS

CNS	Central Nervous System
CSF	Cerebrospinal Fluid
cGM	Cortical Grey Matter
dGM	Deep Grey Matter
DIR	Double Inversion Recovery
FLAIR	Fluid Attenuated Inversion Recovery
GM	Grey Matter
MR	Magnetic Resonance
MT	Magnetization Transfer
MTR	Magnetization Transfer Ratio
MS	Multiple Sclerosis
PDw	Proton Density Weighted
RRMS	Relapsing-Remitting Multiple Sclerosis
SPMS	Secondary-Progressive Multiple Sclerosis
T	Tesla
T1w	T_1 Weighted
T2w	T_2 Weighted
WM	White Matter

ABSTRACT

Magnetic resonance imaging has become an invaluable tool in clinical neurology because of its ability to provide excellent contrast between different soft tissues of the brain, including pathologically altered tissues. Image processing methods can be used to further the detection and quantification of such tissues, as has been the case for the well-studied white matter (WM) lesions in the brains of patients with multiple sclerosis (MS). Considered an immune-mediated disease of the central nervous system, MS has long been believed to be a disease of the WM, though recent studies have shifted the focus to the pathologically altered grey matter (GM) tissue, which is difficult to visualize in vivo. As the appreciation grows for cortical GM pathology in patients with MS, so does the need for imaging methods that capture said pathology and for understanding the relationship of those methods to the clinical outcome of the patient. Hence, the main objective of this thesis was to evaluate and develop image processing methods for the in vivo quantification of GM pathology in patients with MS. We reveal the strengths and weaknesses of the currently available techniques for measuring GM pathology, highlighting the need for separate cortical and deep GM methodologies. We also, introduced a novel surface-based technique involving magnetization transfer ratio (MTR) images for imaging putative areas of subpial demyelination, the most common form of cortical lesion seen on postmortem analysis that has yet to be captured via in vivo imaging techniques. Finally, we showed that, of the imaging metrics available for measuring cortical pathology, widespread cortical damage detected with surface-based measures of MTR have the strongest relationship to cognitive performance in cognitively impaired patients with MS. Importantly, all cross-sectional studies were performed using images typically obtained in the clinical setting and at the accessible field strengths of 1.5 T and 3 T, giving the work instant clinical feasibility and relevance.

ABRÉGÉ

L'imagerie par résonance magnétique est un outil précieux en neurologie grâce à sa capacité à fournir un excellent contraste entre différents tissus du cerveau, y compris les tissus altérés pathologiquement. Les méthodes de traitement d'image peuvent être utilisées pour favoriser la détection et la quantification de ces tissus, comme dans le cas des lésions de la substance blanche (SB) du cerveau des patients atteints de sclérose en plaques (SEP). Considérée comme une maladie auto-immune du système nerveux central, la SEP a longtemps été étudiée comme une maladie de la substance blanche. Des études récentes ont déplacé l'attention vers les pathologies de matière grise (SG) du tissu qui sont difficiles à visualiser in vivo. L'appréciation se développe pour la pathologie de la SG corticale chez les patients atteints de SEP, tout comme le besoin de méthodes d'imagerie qui capturent ladite pathologie et la compréhension de la relation de ces méthodes à l'évolution clinique du patient. Par conséquent, l'objectif principal de cette thèse est l'évaluation et le développement de méthodes de traitement d'image pour la quantification in vivo de la pathologie de SG chez les patients avec la SEP. Nous révélons les forces et les faiblesses des techniques disponibles pour mesurer la pathologie de SG, soulignant la nécessité de séparer les méthodes pour le SG corticale et profonde. Ensuite, on présente une nouvelle technique de surface impliquant des images de rapport de transfert de magnétisation (MTR) pour étudier les zones possibles de démyélinisation sous-pie-mère, la forme la plus commune de lésion corticale observée en analyse post-mortem qui n'ait pas encore été vue par l'intermédiaire d'imagerie in vivo. Enfin, nous montrons que, parmis les indicateurs disponibles pour mesurer pathologie corticale en imagerie, les lésions corticales détectées à l'aire de mesures de MTR sur des surfaces ont la corrélation la plus forte avec la performance cognitive chez des patients atteints de SEP avec des facultés cognitives affaiblies. Fait important, toutes les études transversales ont été réalisées à partir d'images obtenues typiquement en milieu clinique et à des niveaux de champs de 1,5 T et 3 T, conférant à cet ouvrage faisabilité et pertinence clinique.

ORIGINAL CONTRIBUTIONS

The following original scientific contributions of this Ph.D. thesis are claimed:

1. The development of a surface-based framework tailored for quantifying cortical abnormalities in the brains of patients with a neurological disease;

2. The application of the developed surface-based methodology for the in vivo detection of areas of decreased MTR in patients with MS;

3. The evaluation of the sensitivity of surface-based measures of MTR;

4. The demonstration that widespread cortical damage being a better predictor of cognitive performance than focal or regional cortical damage;

5. The demonstration that surface-based measures of MTR are more sensitive than surface-based (or volumetric) measures of cortical atrophy to detecting cognitive impairment in patients with MS;

6. The validation of the accuracy of the resultant grey matter segmentations of several automated methods commonly used for the detection of grey matter atrophy in patients with MS; and more specifically,

    a. The quantification of the variability and severe shortcomings in the deep grey matter segmentations of said methods and

    b. The quantification of the variability and the establishment of an acceptable degree of accuracy in the cortical grey matter segmentation of said methods

CONTRIBUTION OF AUTHORS

As the first author of all three manuscripts included in this thesis, I was responsible for the design, implementation, and validation of all methodological developments. I was involved with the experimental design, performed all the data and statistical analysis, interpreted the results, and drafted the manuscript for each study. The following list summarizes the various contributions of each author for each of the manuscripts presented in Chapters 3 to 5.

EVALUATION OF AUTOMATED TECHNIQUES FOR THE QUANTIFICATION OF GREY MATTER ATROPHY IN PATIENTS WITH MULTIPLE SCLEROSIS

MISHKIN DERAKHSHAN[1,2], ZOGRAFOS CARAMANOS[1,2], PAUL S. GIACOMINI[1], SRIDAR NARAYANAN[1,2], JOSEFINA MARANZANO[2], SIMON J. FRANCIS[1,2], DOUGLAS L. ARNOLD[1,2], AND D. LOUIS COLLINS[1]

[1]McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, Montreal, Canada H3A 2B4
[2]NeuroRx Research, 3605 University, Suite #4, Montreal, Canada H3A 2B3

NeuroImage 52 (2010) 1261-1267

Guarantors of the study:	All authors
Study concept and design:	All authors
Manual segmentations:	PSG, JM
Multispectral Bayesian classifier:	SJF
Data analysis:	MD
Statistical guidance:	ZC
General guidance and supervision:	DLA, DLC
Manuscript preparation:	MD
Manuscript revision:	All authors
Final editing:	MD, SN, ZC, DLA, DLC

SURFACE-BASED ANALYSIS REVEALS REGIONS OF REDUCED CORTICAL MAGNETIZATION TRANSFER RATIO IN PATIENTS WITH MULTIPLE SCLEROSIS: A PROPOSED METHOD FOR IMAGING SUBPIAL DEMYELINATION

MISHKIN DERAKHSHAN[1], ZOGRAFOS CARAMANOS[1], SRIDAR NARAYANAN[1], DOUGLAS L. ARNOLD[1,2], AND D. LOUIS COLLINS[1]

[1]McConnell Brain Imaging Centre, Montreal Neurological Institute, Montreal, Canada H3A 2B4
[2]NeuroRx Research, 3575 Parc Avenue, Suite #5322, Montreal, Canada H2X 4B3

Human Brain Mapping (In Submission: HBM-12-0884)

Guarantors of the study:	All authors
Study concept and design:	MD, SN, DLA, DLC
Methodological development:	MD
Data analysis:	MD
Statistical guidance:	ZC
General guidance and supervision:	DLA, DLC
Manuscript preparation:	MD
Manuscript revision, final editing:	All authors

WIDESPREAD CORTICAL PATHOLOGY MEASURED WITH MAGNETIZATION TRANSFER RATIO BEST PREDICTS COGNITIVE IMPAIRMENT IN PATIENTS WITH MULTIPLE SCLEROSIS

MISHKIN DERAKHSHAN[1], GABRIEL LEONARD[1], SRIDAR NARAYANAN[1], ZOGRAFOS CARAMANOS[1], D. LOUIS COLLINS[1] AND DOUGLAS L. ARNOLD[1,2],

[1]McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, Montreal, Canada, H3A 2B4
[2]NeuroRx Research, 3575 Parc Avenue, Suite #5322, Montreal, Canada H2X 4B3

NeuroImage (in preparation)

Guarantors of the study:	All authors
Study concept and design:	All authors
Data analysis:	MD
Statistical guidance:	ZC
General guidance and supervision:	DLC, DLA
Neuropsychological data:	GL
Manuscript preparation:	MD
Manuscript revision, final editing:	All authors

OTHER PUBLICATIONS

The following peer-reviewed conference articles and proceedings were also produced during the course of this doctoral thesis:

Derakhshan, M., G. Leonard, D. Araujo, D. L. Collins, D. L. Arnold, and S. Narayanan. "Hippocampal Magentization Transfer Ratio and Not Hippocampal Atrophy Best Explains Memory Dysfunction in Patients with Multiple Sclerosis." Paper submitted to the International Society for Magnetic Resonance in Medicine (ISMRM), Salt Lake City, USA, April 20-26, 2013

Derakhshan, M., S. Narayanan, D. L. Collins, and D. L. Arnold. "Combining Siena and Sienax for Improved Quantification of Grey and White Matter Atrophy." Paper presented at the International Society for Magnetic Resonance in Medicine (ISMRM), Melbourne, Australia, May 5-11, 2012

Derakhshan, M., C. Griffiths, D. Araujo, S. Narayanan, D. L. Collins, G. Leonard, and D. L. Arnold. "Global and Regional Cortical Thickness Correlates of Cognitive Function in Cognitively Impaired Patients with Multiple Sclerosis." Paper presented at the American Academy of Neurology (AAN), Honolulu, USA, April 9-16, 2011

Giacomini, P., D. Araujo, **M. Derakhshan**, Z. Caramanos, S. Narayanan, and D. L. Arnold. "Brain Volume Change Following Inititation of Ifn-B1a Therapy Is Not Fully Explained by 'Pseudoatrophy'." Paper presented at the American Academy of Neurology (AAN), Honolulu, USA, April 9-16, 2011

Miller, D.H., C. Dalton, S. Narayanan, **M. Derakhshan**, D. Tozer, K. Hunter, D. G. MacManus, C. Wheeler-Kingshott, R.J. Fox, R. Gold, M. Yang, K. Dawson, and D. L. Arnold. "Feasibility of Magnetization Transfer Ratio Imaging in Multicenter Trials of Patients with Multiple Sclerosis." Paper presented at the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS), Gothenburg, Sweden, October 13-16, 2010

Derakhshan, M., S. Narayanan, J.T. Chen, P.S. Giacomini, D. L. Arnold, and D. L. Collins. "'Pseudo-T2' Relaxation Times Using a Dual Turbo Spin-Echo Sequence: Methodology and Validation." Paper presented at the Organization for Human Brain Mapping (OHBM), Barcelona, Spain, June 6-10, 2010

Arnold, D. L., C. Dalton, S. Narayanan, **M. Derakhshan**, D. Tozer, K. Hunter, D. G. MacManus, C. Wheeler-Kingshott, R.J. Fox, R. Gold, M. Yang, K. Dawson, and D.H. Miller. "Magnetization Transfer Ratio Imaging Is Feasible in Multicenter Multiple Sclerosis Trials." Paper presented at the The consortium of multiple sclerosis centers, San Antonio, USA, June 2-5, 2010

***Derakhshan, M.**, Z. Caramanos, S. Narayanan, D. L. Arnold, and D. L. Collins. "Surface-Based Techniques Reveal Regions of Reduced Cortical Magnetization Transfer Ratio in Patients with Ms." Paper presented at the International Society for Magnetic Resonance in Medicine (ISMRM), Stockholm, Sweeden, May 1-7, 2010

Arnold, D. L., C. Dalton, S. Narayanan, **M. Derakhshan**, D. Tozer, K. Hunter, and D.H. Miller. "Magnetization Transfer Ratio Imaging Is Feasible in Large Multicenter Ms Trials." Paper presented at the American Academy of Neurology (AAN), Toronto, Canada, April 11-16, 2010

Derakhshan, M., Z. Caramanos, S. Narayanan, P.S. Giacomini, J. Maranzano, T. Li, S. J. Francis, D. L. Arnold, and D. L. Collins. "An Evaluation of Automatic Techniques for the Quantification of Grey Matter Atrophy in Multiple Sclerosis." Paper presented at the Medical Image Analysis and Acquisition in Multiple Sclerosis (MIAMS) workshop at Medical Image Computing and Computer-Assisted Interventation (MICCAI), London, England, September 20-24, 2009

***Derakhshan, M.**, Z. Caramanos, S. Narayanan, D. L. Collins, and D. L. Arnold. "Regions of Reduced Cortical Magnetization Transfer Ratio Detected in Ms Patients Using Surface-Based Techniques." Paper presented at the International Society for Magnetic Resonance in Medicine (ISMRM), Honolulu, USA, April 18-24, 2009

Caramanos, Z., V.S. Fonov, **M. Derakhshan**, D. L. Collins, S. Narayanan, and D. L. Arnold. "Effects of Actual and Simulated Z-Shifts on Freesurfer-Segmented Whole-Brain-Tissue, Hippocampal, and Lateral-Ventricle Volumes." *Alzheimer's & Dementia* 4, no. 4: T276-T77, 2008

Caramanos, Z., V.S. Fonov, **M. Derakhshan**, D. L. Collins, S. Narayanan, and D. L. Arnold. "Z-Shift Effects on Segmented Brain-Tissue Volumes." *Mult Scler* 14, no. Suppl. 1: S99, 2008

Giacomini, P.S., **M. Derakhshan**, S. Narayanan, J. G. Sled, A. D. Sadovnick, A. Bar-Or, B. Banwell, and D. L. Arnold. "Patterns of Regional Brain Atrophy and Growth in Children with Clinically Isolated Syndromes at High and Low Risk of Conversion to Multiple Sclerosis." *Mult Scler* 14, no. Suppl. 1: S105, 2008

Panet-Raymond, V., P.S. Giacomini, **M. Derakhshan**, A. Fiocco, D. L. Collins, L. Souhami, D. L. Arnold, and G. Shenouda. "Differential Regional Brain Atrophy Following Prophylactic Cranial Irradiation in Small Cell Lung Cancer Patients." Paper presented at the 50th Annual Meeting of the American Society for Therapeutic Radiology and Oncology, Boston, USA, Sept 21-25, 2008

Panet-Raymond, V., P.S. Giacomini, R. DelCarpio, **M. Derakhshan**, A. Fiocco, D. L. Collins, D. L. Arnold, L. Souhami, and G. Shenouda. "Ventricular Enlargement and Brain Colume Changes Post Prophylactic Cranial Irradiation in Small Cell Lung Cancer Patients." Paper presented at the 22nd Annual Meeting of the Canadian Association of Radiation Oncology, Montreal, Canada, Sept 10-13, 2008

Chen, J. T., D. L. Collins, **M. Derakhshan**, and D. L. Arnold. "Quantifying MRI with Increased Specificity for Ms Pathology: A Longitudinal Method for Obtaining Whole-Brain Metrics of 3d Maps Derived from Non-Conventional MRI." Paper presented at the International Society for Magnetic Resonance in Medicine (ISMRM), Toronto, Canada, May 3-9, 2008

*selected for oral presentation

Chapter 1

INTRODUCTION

Multiple sclerosis (MS), an immune-mediated disease of the central nervous system (CNS), has been brought to the attention of the general public to the point that almost everyone has at least heard of it. Multiple sclerosis is the most common neurological disease affecting young adults in Canada. Though MS is a complex disease, most people associate it only with paralysis. Some will know that the disease affects the myelin in the brain (and/or spinal cord), while others will know that MS is most often characterized by lesions in the white matter (WM). Few will have knowledge about the involvement of grey matter (GM) in the disease, and worst of all, no one will be able to state the cause or cure for MS as they are still unknown.

Pathologists have been studying the disease for nearly two centuries, but it was not until the advent of magnetic resonance (MR) imaging in the 1970s that research into MS exploded. This noninvasive in vivo imaging technique provided an otherwise unseen window into the behaviour of the disease, allowing us to learn much about WM and, to a lesser extent, GM pathology. The reasons for our limited familiarity with GM

pathology are numerous, yet, with the evolution of both postmortem and in vivo imaging techniques, we are gaining more insight into this aspect of the disease.

The in vivo detection and quantification of GM pathology in patients with MS is necessary for a complete characterization of the disease. Understanding the role of GM damage in MS is critical not only to our understanding of MS, but also to the development of therapeutic treatments that can potentially delay or even reverse disease progression.

The main goal of this dissertation was to evaluate and develop image processing methods for the in vivo quantification of GM pathology in patients with MS.

Given the multidisciplinary nature of this thesis, the reader is required to have a basic understanding of various topics to gain a full appreciation of the content herein. From the clinical and pathological descriptions of MS to the physics of MR imaging and its applications in MS, from an assortment of image processing techniques to the various tests used for assessing cognitive impairment; all the required background knowledge can be found in detail in Chapter 2.

The body of the dissertation is based on three manuscripts presented in Chapters 3 to 5. Published in *NeuroImage*, the study in Chapter 3 is an evaluation of the most common image processing methods that have been used to draw conclusions about GM atrophy in patients with MS. The main finding of the study—that there are severe shortcomings in the resultant deep GM (dGM) segmentations—along with the other identified strengths and weaknesses will help shape how studies of GM atrophy will be conducted and interpreted in the future.

Chapter 4 introduces a novel surface-based image processing method for imaging putative areas of subpial demyelination, a more focused type of GM pathology that is the most common form seen on postmortem examination and had yet to be well captured via in vivo imaging techniques. The method revealed areas of decreased

cortical magnetization transfer ratio (MTR), a more specific marker for myelin than other conventional imaging modalities, with a predilection for those patients suffering from progressive forms of MS. In assessing the sensitivity of the newly developed technique, we found that, while the results have not been directly confirmed with postmortem histology, they are still in line with what pathologists have reported. The manuscript is currently in submission to *Human Brain Mapping*.

Chapter 5 combined the knowledge gleaned from Chapter 3 and the new methodology introduced in Chapter 4 to investigate multiple forms of GM pathology in patients with MS, and then looked at the relationship between these imaging characteristics and the patients' cognitive performance. The study found that widespread cortical damage, in particular demyelination (as assessed via surface-based measures of MTR), and not cortical atrophy is primarily responsible for the cognitive deficiencies present in patients with MS. The manuscript is currently in preparation for submission to *NeuroImage*.

Finally, the novelty, strengths and weaknesses, scientific importance, and clinical relevance of each of the manuscripts and of the thesis as a whole are discussed in Chapter 6, along with suggestions for future work.

Chapter 2

It is now well recognized that the cortex is affected in disseminated sclerosis.

—James Walker Dawson, **c.1914**

BACKGROUND

The following background chapter is cleaved into four broad categories. The first section is on multiple sclerosis, and presents a clinical and a pathological description of the disease with an emphasis on the neuropathological descriptions of cGM lesions. The second is on magnetic resonance imaging, beginning with a discussion of the basic physics behind it, and ending with the role of MR imaging in MS. The third portion of the chapter is on image processing in MS, presenting some fundamentals of image processing before reviewing many of the automated methods specific to the work in this thesis. The chapter concludes with a review of cognitive dysfunction in patients with MS and their MR imaging correlates.

2.1 MULTIPLE SCLEROSIS

Ask anyone in Canada if they've heard of MS, and chances are they'll say yes. After all, there are television commercials for MS awareness and highly successful nationwide fundraising events such as the MS Walk and the MS Bike Tour, and in the past couple of years, potential and controversial therapies have been widely featured in the media [1]. Despite this high level of awareness, if you were to ask people to describe the disease, you would get many vague and varied responses—and for good reason. Multiple sclerosis is a difficult disease to characterize; there is no known cause and no known cure. Symptoms vary between individuals and may include vision problems, fatigue, motor-sensory deficits, cognitive impairments, and, in patients with severe disease, paralysis that may eventually lead to death.

What follows, then, is simply our best guess at characterizing this multifaceted disease, gleaned from various studies to date.

2.1.1 CLINICAL DESCRIPTIONS

What is MS?

Multiple sclerosis is an immune-mediated disease of the human CNS, which comprises the brain and spinal cord. The disease attacks the protective covering (myelin) wrapped around the nerves of the CNS. Multiple sclerosis is a chronic neurological disease and is characterized pathologically by inflammation, demyelination, axonal degeneration, and gliosis.

The term sclerosis indicates an abnormal hardening of a tissue (i.e., a scar), and it was Dr. Charcot[i] who, in 1868, detailed his clinical and pathological observations of a disease

[i] Though Charcot is credited with giving the disease a name in 1868, he wasn't the first to describe it, nor was it actually referred to as multiple sclerosis at that time. Several individuals prior to Charcot had documented cases of the disease but failed to give it a name (e.g., Carswell in 1838 and Cruveilhier in 1841 illustrated cases of MS in their atlases, and perhaps the earliest recorded description of the disease is of a Dutch nun c. 1308). The English translation of Charcot's name for the disease is "disseminated sclerosis", and it was only in Germany where the term *multiple Sklerose* was used from the onset. It wasn't until the publication of McAlpine's monograph in 1955 that a consensus was reached and the disease was universally referred to as multiple sclerosis. See [2] and [3] for a detailed history of MS.

that he named *la sclérose en plaques disséminées,* bringing the disease to the attention of the medical world.

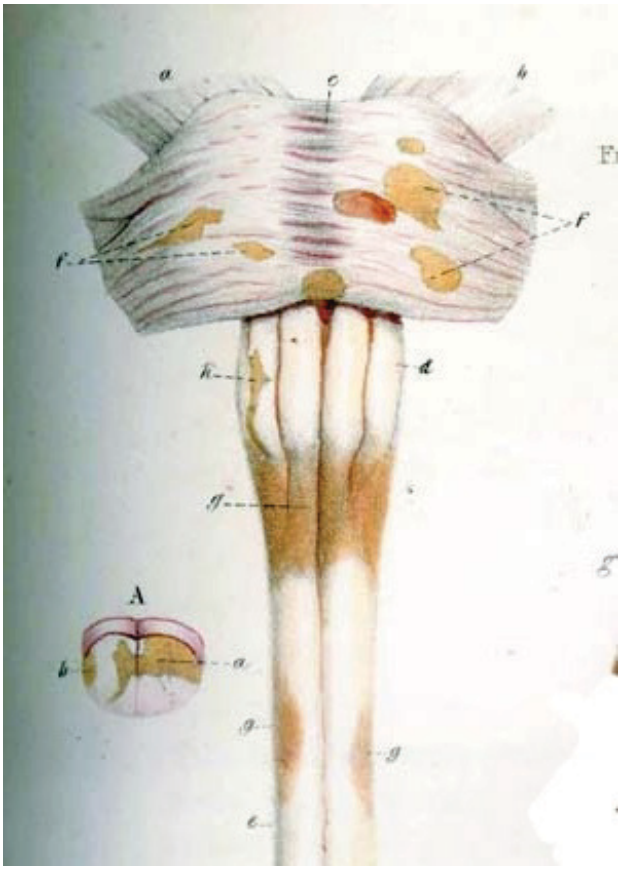

Figure 2.1 – Drawing from Carswell's 1838 atlas depicting lesions in the brainstem and spinal cord [4].

Who is affected by MS?

Multiple sclerosis affects approximately 2.5 million people worldwide [5], and the Multiple Sclerosis Society of Canada estimates that between 55,000 and 75,000 Canadians have multiple sclerosis [6]. Just as prevalence rates vary by country, the Canadian MS prevalence rate differs by region, with a national average of 240 per 100,000 (95%CI: 210-280), making it one of the highest in the world [7]. Multiple sclerosis has been shown to affect roughly three[ii] times as many women as men [8] and is considered the most common neurological disease affecting young adults [9], with most

[ii] The sex ratio in multiple sclerosis was first reported to affect more men than women in the early 1900s. It then moved to close to unity (1:1) until the 1950s and 1960s, when females were reportedly more affected by a ratio of 1.4:1. The current Canadian estimate now exceeds 3.2:1. The interested reader is referred to the excellent discussion in Orton et al. 2006 [8].

individuals having their clinical onset between the ages of 20 and 40 years (peak age of onset is 24 years in females, 25 in males [10]). Still, well-documented pathological cases are known from the first decade of life [11] to the ninth [12].

What causes MS?

The etiology of MS is yet unknown, although several factors have been implicated in its development. Genetics appear to play a role (genetically identical monozygotic twins are more likely to both develop MS than are dizygotic twins or siblings [13]), but importantly, MS is not a purely genetic disease as even genetically identical monozygotic twins have a concordance rate of only approximately 30% [14].

Environmental factors (e.g., high latitudes, low sun exposure, nutritional factors, etc.) also seem to have an effect, though their exact roles are unknown. There are believed to be high-, medium-, and low-risk zones of MS worldwide [15], but again, the geographic distinctions for each are changing. Early life exposure to some environmental factors is suggested by epidemiological data [16]. A maternal parent-of-origin effect has been reported in half-siblings [17], and a month-of-birth effect (highest risk in May, lowest in November) has also been reported [18]. Dietary components, in particular vitamin D, have also been found to play a role in the development of and/or protection from MS [19].

Many theories of what causes MS have been proposed, and many are being investigated. Nonetheless, as there is still no clear answer, the consensus seems to be that the trigger is complex, likely originating from an interaction of genetic and environmental factors.

What are the subtypes of MS?

Multiple Sclerosis has four clinical course definitions [20], with the relapsing-remitting subtype being generally less severe than the other, progressive subtypes.

1. Relapsing-remitting (RRMS): unpredictable relapses (i.e., attacks with a clinical manifestation) followed by periods of full or partial recovery.

2. Primary-progressive (PPMS): a gradual and continual progression of disability from onset with little or no remission or improvement.

3. Secondary-progressive (SPMS): a progressive phase that may follow in those initially with a RRMS course, characterized by progressive neurological decline between acute attacks without any definite period of remission. About 50% of people with RRMS will develop SPMS within 10 years of diagnosis [6].

4. Progressive-relapsing (PRMS): the least common of all subtypes, characterized by steady neurological decline from onset, but with clear superimposed attacks.

2.1.2 PATHOLOGICAL DESCRIPTIONS

Primer

In order to grasp the neuropathology and pathophysiology of the MS lesion, one needs a general picture and a fundamental understanding of neurobiology. Electrical signals, or messages, are transmitted throughout the CNS by specialized nerve cells called neurons. The human brain has approximately 85 billion neurons [21], which are considered the main functional unit of the CNS. A neuron consists of three main parts: a cell body (or soma), dendrites, and an axon (Figure 2.2). Dendrites bring information to the cell body, while the axon takes the information away, passing it along to the next nerve cell via the presynaptic terminals at the other end of the neuron. This entire process, the transmission of a message from one neuron to another—formally termed the propagation of an action potential—is facilitated by myelin. Commonly likened to the insulation around an electrical wire, myelin is composed of roughly 40% water and, in

terms of dry weight, 75% lipids and 25% proteins [2]. The lipids in particular are what insulate the axons from electrically charged atoms and ions, allowing electrical signals to propagate along the axon at increased speeds.

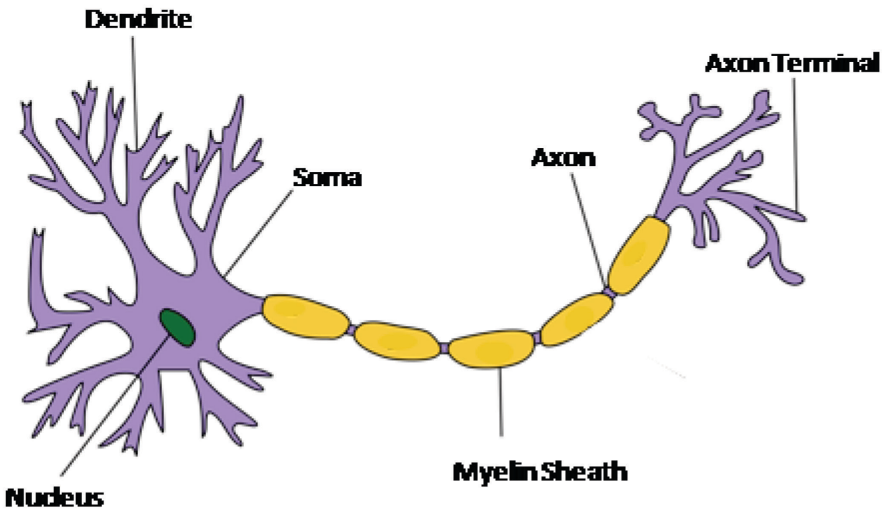

Figure 2.2 – A neuron. The dendrites, axon, and soma form the three main parts of the nerve cell. Adapted from [22].

In addition to nerve cells, and found in at least equal numbers in the brain if not more [21], are neuroglia. Neuroglia are the support cells that ensure the proper functioning of the neurons and are also involved in the immune response system. The immunologic lexicon is rich, containing specific terminology to describe the same cell depending on its stage of development and its environment. The following does not provide as exhaustive a list as one would require to appreciate hematopoiesis in the human body, but it should help the reader define the scope and hierarchy of the major cells involved in the immunology of MS.

One glial cell in particular, the oligodendrocyte, plays a major role in the support of neurons and in MS. Oligodendrocytes are responsible for the myelin sheath around an axon (Figure 2.3). In fact, a single oligodendrocyte extends multiple processes that myelinate multiple axons. In MS, it is not entirely clear whether the primary event is an attack on the oligodendrocyte, the myelin sheath, or both; however, evidence suggests that the death of the oligodendrocyte is secondary to the loss of myelin [23].

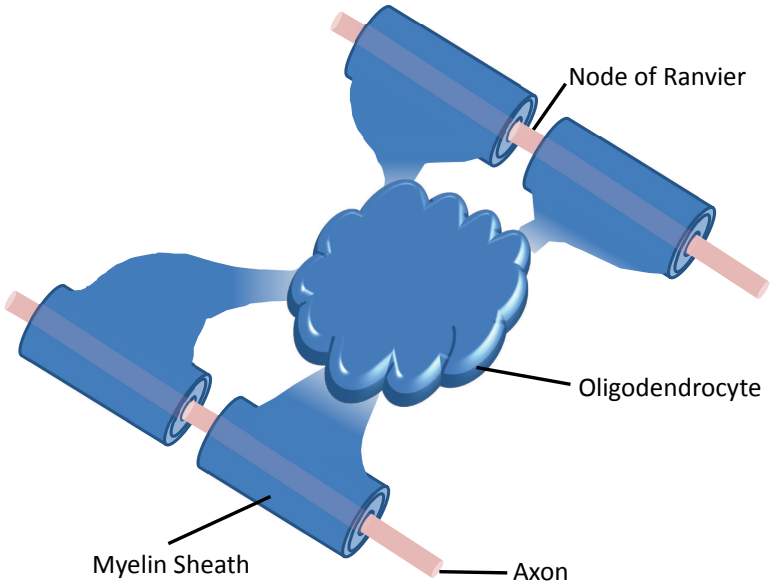

Figure 2.3 – Oligodendrocyte. A single oligodendrocyte (blue) is responsible for synthesizing the myelin sheath of multiple axons (beige). The space in between the oligodendrocytes' processes make up the nodes of Ranvier.

Astrocytes, the star-shaped glial cells in the CNS, are probably the most abundant cells in the human brain and are another major component of MS immunology. They have multiple functions, including but not limited to maintenance and repair, and are central to the scarring processes in the brain following traumatic injuries.

Lymphocytes, the broad class of cells that derive from the lymphoid progenitor stem cell, play a major role in the innate (natural killer cells) and adaptive (T-cells and B-cells) immune response systems.

Finally, microglia are the resident macrophages of the brain. Recall that a macrophage is simply a monocyte that has left the bloodstream and differentiated within a specific tissue. Microglia can be thought of as the clean-up crew of the CNS as they are responsible for destroying and cleaning up any infectious agents.

White Matter Lesions

The traditional pathological view of MS centred on focal WM plaques, or lesions, in the brain and spinal cord, even though the earliest descriptions of the disease mentioned

both extralesional pathology and lesions in the GM[iii]. Recent advances in immunohistochemical staining techniques have led to a boom in the investigation of the GM pathology in MS, and it is now believed that MS is much more than merely a disease of the WM. Nevertheless, we will begin with a description of the WM pathology, as it is a prerequisite for understanding the subtle differences observed in the GM pathology in MS.

The basic WM plaque consists of a sharply demarcated lesion that varies in size, location, and volume. It is generally ovoid in shape but may show finger-like extensions that follow the path of small vasculature[iv] [24]. White matter lesions are heterogeneous, and while they have been classified into subtypes (e.g., early vs. chronic, active vs. inactive), a discussion of each is beyond the scope of this work. The defining histological features that are present to some degree or another in WM lesions are inflammation, gliosis, demyelination, and axonal degeneration.

Inflammation is a complex system and is part of the body's attempt to remove harmful stimuli. In the case of the WM lesion in MS, the inflammatory infiltrate comprises lymphocytes and monocytes (T-cells, B-cells, microglia, etc.) and is often multifocal. Inflammation can also be seen in the leptomeninges (thin membranes that surround the brain), but is most prominent in the perivascular spaces (around or surrounding a vein, artery, or capillary).

Gliosis refers to the formation of a glial scar and involves the production of a dense, fibrous network of glial cells that replaces the normal tissue after an injury. Astrocytes are the main cellular component of the glial scar, though microglia are also present.

Demyelination is a degenerative process that, as its name implies, involves the erosion of the myelin sheath that normally protects nerve fibres. Though not specific to MS, it is

[iii] Although the literature and indeed the content of this chapter focus on the lesions of the brain, it should be noted that MS lesions occur in all aspects of the CNS: cerebrum, cerebellum, brain stem, spinal cord, and optic nerve.
[iv] James Walker Dawson was the first to describe this condition c.1914, but it was Charles E. Lumsden (1913-1974) who named them Dawson's fingers, a term still commonly used today.

considered one of the trademark characteristics of the disease. Demyelination slows the conduction of signals in the affected neurons and can manifest as impairment of sensation, movement, cognition, or other functions, depending on which neurons are involved and to what degree. Complementary to demyelination is the remyelination of a denuded axon, a once thought impossible feat that we now believe happens continuously in the brain and has even been demonstrated in MS.

Axonal pathological changes were mentioned by Charcot; yet for a long time, it was assumed that axons were spared from the destructive process. Seminal work by Trapp et al. in 1998 [25] quantified the extent of axonal degeneration and loss in demyelinating MS lesions, and evidence of axonal transection can clearly be seen in Figure 2.4.

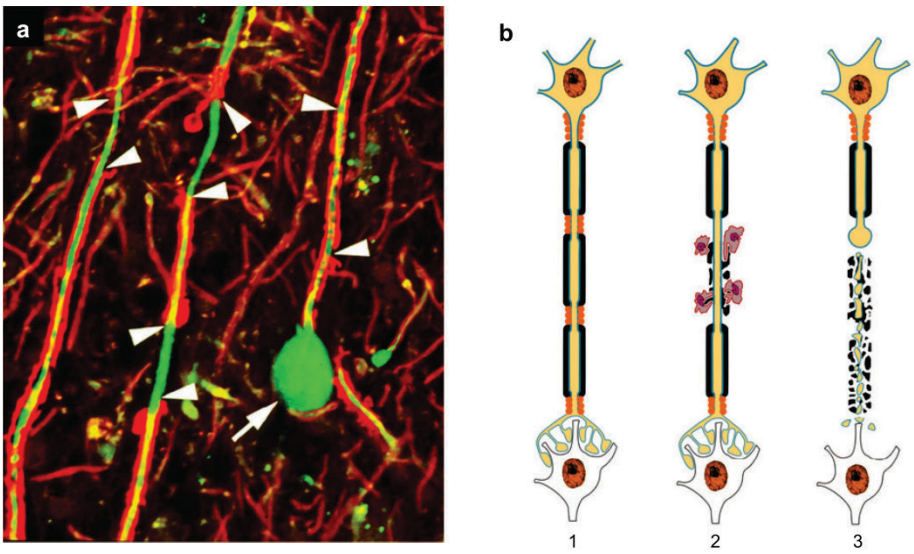

Figure 2.4 – Axons are transected during demyelination. (a) Confocal image of an actively demyelinating MS lesion stained for myelin protein (red) and axons (green). The three vertically oriented axons have areas of demyelination (arrowheads). The axon on the right ends in a large swelling (arrowhead), or axonal retraction bulb, which is the hallmark of the proximal end of a transected axon. (b) Schematic summary of axonal response during and following transection. 1. Normal-appearing myelinated axon. 2. Demyelination. 3. Transection and the formation of a transectional bulb, as well as the degeneration of the distal end of the transected axon. Panel and modified caption from [26].

Cortical Grey Matter Lesions

As recently as 15 years ago, the focus of MS pathology was almost exclusively on demyelinating WM, with little or no mention of GM and/or cortical lesions—an inconsistency to say the least, considering that neuropathologists of the past identified GM involvement in the disease over a century ago.

Accounts from the late 19th and early 20th century are predominantly from the German school of neuropathologists. In 1894, Edward Wyllys Taylor observed three chronic MS cases and concluded that demyelination affects WM and GM equally in all parts of the CNS [27]. Sander in 1898 also documented a single chronic MS case with severe cortical disease involvement. He concurred that the cortex and WM were equally affected by demyelination, but added that the extent of reactive astrogliosis (i.e., the increase in astrocytes that leads to the formation of a glial scar) and axonal loss were also equal in GM lesions [28]. A few years later, in 1904, Dinkler documented another chronic MS case with cortical lesions, but went so far as to say, upon histopathological examination, that demyelinating lesions were mainly located in the cerebral cortex [29].

A decade later, the final 250 pages of the third volume of the 1913-1914 Transactions of the Royal Society of Edinburgh contain the humbling work on the histology of disseminated sclerosis by James Walker Dawson, and of course, there is ample mention of cortical pathology [24]. Though his quantitative descriptions were limited to saying that "cortical and subcortical [lesional] areas were numerous in six of the cases, few in two, and in the ninth case not evident", it is astonishing how the pathological descriptions and the figures presented are on par with the work produced today, after a hundred years of technological advancement. Of cortical lesions, Dawson noted that the essential change is a demyelination and he observed it in mixed WM and GM lesions, in small lesions that lie wholly within the cortex, and in larger areas where "the demyelination may reach from the surface of the convolution to varying depths, even to the border between cortex and white matter."

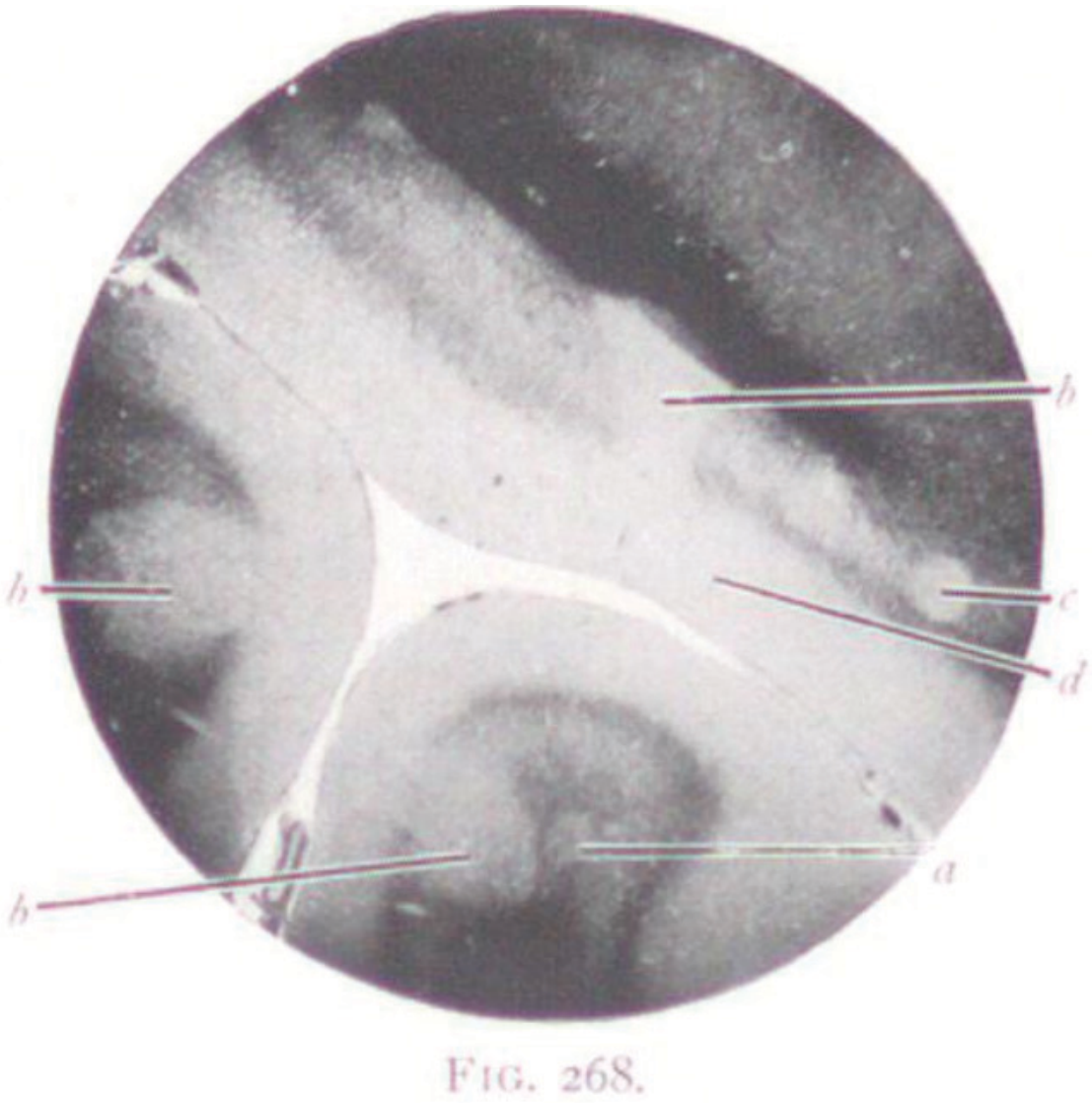

FIG. 268.

Figure 2.5 – Early histological pictures of cortical lesions. Unfortunately, the caption for Fig. 268 in the original work appears to be mismatched with the figure. The correct caption, erroneously labeled for Fig. 270 reads as follows: "a = small area confined to a medullary ray; b = areas involving both medullary ray and radiations; c = confined to the deep cortex; d = extensive demyelination of the superficial cortex." [24]

After these first pivotal case reports of GM damage in MS, a larger study by Brownell and Hughes in 1962 looked at 1,500 plaques in 22 cases of MS. Regarding distribution, the authors found that 17% of the lesions were leukocortical (i.e., cortical GM with some subcortical WM involvement), but that only approximately 5% were limited solely to the cortex. Unlike the histopathological studies of the past, Brownell and Hughes identified these plaques macroscopically and remarked that "cortical plaques were difficult to see

macroscopically and it is possible that histological examination might have shown more in this situation" [30].

Charles E. Lumsden's monumental work on the neuropathology of MS from the 1970s also examined the macroscopic distribution of plaques, but he came to a profoundly different conclusion than that of Brownell and Hughes. Of the first 60 consecutively obtained specimens, 93.5% of them exhibited plaques located in the cerebral gyri (16% had 3 or fewer plaques; 87.5% had 4 or more). He noted that the gyral plaques (which included both leukocortical and purely intracortical lesions) were smaller than their WM counterparts, usually between 2 mm and 8 mm, and in some individuals, the number of these plaques ran into the hundreds (e.g., 465 in one case, 302 in another, and 103 in a third). In a subset of 10 early and severe cases of MS (482 plaques total), Lumsden found that 59% of the plaques were distributed in the cerebral gyri. An examination of the distribution of the lesions revealed that no particular gyrus was exempt, that perhaps the greatest frequency was for the superior frontal gyrus (though it is by far the most voluminous gyrus he examined), and surprisingly, that the lesions were equally distributed bilaterally, even in individual cases. Whatever their distribution, though, he noted that lesions in the cortex are easily overlooked (due to staining methodology and erroneous preconceptions) and that the restriction of MS to WM is a fiction [31].

Despite these and other studies that highlighted GM demyelination in MS, the topic was largely disregarded until the 21st century, mostly for the following two reasons:

1. Visualization of postmortem cortical GM (cGM) lesions was difficult because of the limitations of conventional histochemical staining procedures.

2. Attention was predominantly focused on the generally more conspicuous process of inflammatory WM demyelination (exacerbated by the advent of nuclear magnetic imaging in the 1970s).

In histochemistry, Luxol fast blue (LFB) is a stain commonly used to observe myelin under light microscopy. The stain works via an acid-base reaction wherein the base of the lipoprotein in myelin replaces that of the dye, causing a change in colour. Unfortunately, this acid-base reaction is not limited to the lipoproteins in myelin; that is, while LFB may be sensitive to myelin, it is not specific to it. Furthermore, the sensitivity of LFB to myelin depends on the protocol being followed. The standard protocol involves dyeing the entire section until the dye is apparent everywhere, and then rinsing or washing off the excess. The amount of washing done depends on the operator, but the guidelines are to rinse the section until the GM appears colorless [32]. Thus, while LFB is useful for visualizing myelin density in WM, it is almost fated to inadequacy for such detection in GM. Attempts have been made to improve the sensitivity of LFB for the detection of cGM lesions by reducing the destaining time, but they proved unsuccessful because of higher levels of (nonspecific) background staining [33].

A commonly held fallacy asserts that there is no myelin in GM. Indeed, the classically presented cytoarchitectonic view of the cortical layers makes it easy to overlook the presence of myelin in cortex; however, an alternative view focusing on the fibre structures within the cortex, termed "myeloarchitectonics," clearly illustrates the presence of myelin in the cortex.[v] German pathologist Oskar Vogt's schematic representations of the cortex (first published in 1910 and reproduced in Figure 2.6) are clear depictions of myelin in the cortex [35].

[v] The stratification of the cortex is not limited to cyto- and myeloarchitectonics. Pigmentoarchitectonics is yet another view wherein pigment granules are the focus of the cortical lamination. See [34], pg. 66, Figure 4 for an example.

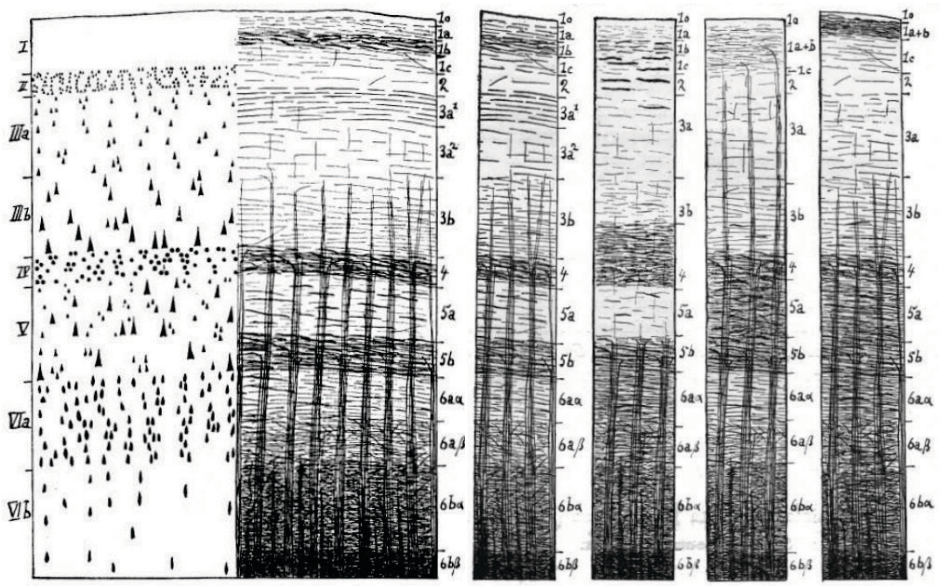

Figure 2.6 – Cyto- and myeloarchitectonic schematics of the human cerebral cortex. Left: Cytoarchitectonic (numerals I-VIb) and the corresponding myeloarchitectonic layers (numbered 1o-6bβ). Right: Four other variants of the myeloarchitectonic layers of the cortex. Reproduced from [35].

As for the second reason the investigation of GM lesions in MS was abandoned for so long—namely, the concentrated focus on WM lesions—paradoxically, that focus later sparked the reinvestigation of the role of GM in the disease. As mentioned before, the work by Trapp et al. at the end of the 20th century brought into centre stage that there is significant axonal and neuronal loss in the WM lesions of patients with MS [25, 36], a finding that caused a paradigm shift in the views on the pathogenesis of tissue injury in MS. The MS research community now suspected that neuronal destruction might play a more central role in the disease and thus began their search for answers where neuronal density is best examined: the cGM.

Perhaps the first modern study to investigate cGM lesions in MS was that of Kidd et al. in 1999 [37]. This frequently cited and highly influential paper characterizes seven different lesion types based on the pathological assessment[vi] of the extent of infiltration

[vi] Though this study did not use the more specific immunohistochemical staining techniques for myelin of the studies that would follow, it improved upon previous staining techniques by refining the LFB staining protocol and by combining it with a hematoxylin and eosin stain (for nuclei and cytoplasm), a cresyl violet stain (for neurons), Heidenhain's myelin stain, and a phosphotungstic acid hematoxylin stain (for glial cells).

of mononuclear inflammatory cells, and the degree of cellularity both within and at the edge of each lesion with WM involvement.

Though this classification system for cortical lesions has today been abandoned in favour of a simpler one, the seven lesion types first enumerated by Kidd et al. were as follows:

Type 1: Located in the deeper cortical layers and include adjacent WM (42% of the 478 observed lesions).

Type 2: Involve all cortical layers, but not the underlying WM (4%).

Type 3: Extensive and located only in the superficial cortical layers (5%).

Type 4: Involve only the subcortical U-fibres (2%).

Type 5: Large, affecting all cortical layers as well as subcortical WM (14%).

Type 6: Small and occur in any part of the cortical ribbon (12%).

Type 7: Largest, affecting both banks of a gyrus with or without involvement of the subcortical WM (21%).

Of the lesions that involved both GM and WM (Types 1, 5, and 7), the WM component was found to be no different than purely WM lesions (i.e., in terms of evidence of inflammation, cellular infiltrates, and edge activity). For the portions of the lesions that involved the cortex, and for the 23% of lesions that were restricted to the cortex (Types 2, 3, 4, and 6), the cortex tended not to be inflamed. Overall, cortical lesions were described as "sharply demarcated demyelinated areas with relative preservation of both axons and neurons with or without an accompanying inflammatory cell infiltration," though the authors remarked that it was difficult to assess this accurately without the use of modern immunohistochemical preparations.

In 2001, Peterson et al. used myelin protein immunocytochemistry[vii] to reliably identify cortical lesions [32]. Through their assessment of 112 cortical lesions from 50 patients with MS, they reduced the seven-type cortical lesion classification schema proposed by Kidd et al. to just three types of cortical demyelination[viii]:

Type I: Juxtacortical lesions that involved both subcortical WM and cortex (representing 34% of the 112 total cortical lesions).

Type II: Purely intracortical lesions, often small and containing a vessel at their centre (16%).

Type III: Subpial demyelination extending from the pial surface into the cortex (mostly to cytoarchitectonic layer III or IV) and often comprising entire or multiple gyri (50%).

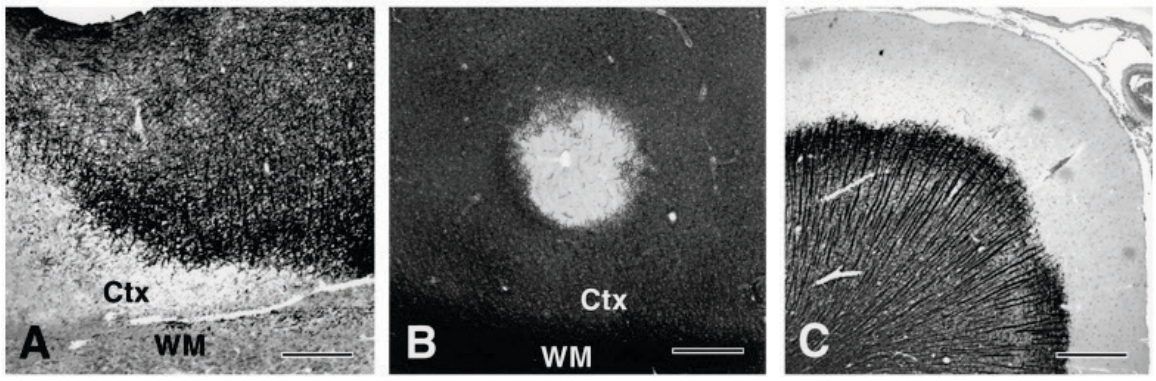

Figure 2.7 – Immunocytochemical distribution of myelin in the cerebral cortices of multiple sclerosis brains identified three patterns of demyelination. A. Type I lesions. B. Type II lesions. C. Type III lesions. WM: white matter. Ctx: cortex. Bars = 200 μm. Reproduced from [32].

[vii] Immunocytochemistry differs from immunohistochemistry in that the former is performed on samples of intact cells where most, if not all, of the surrounding extracellular matrix has been removed. In both cases, antibodies are used that target specific protein antigens, yielding a stain with a very high degree of specificity. In the case of staining for myelin, myelin basic protein (MBP) and proteolipid protein (PLP) are commonly used as they are both expressed on the myelin membrane produced by oligodendrocytes.

[viii] Just two years later, Bo et al. [33] introduced yet another taxonomy for cGM lesions. They divided Kidd et al.'s Type III lesions into those that were subpial but only extended to cortical layer III or IV and those that demyelinated the entirety of the cortex, stopping at the WM. The authors did indicate that their descriptive system should not be considered definitive as classification may change with further knowledge of cortical lesion pathogenesis. Indeed, Peterson et al.'s system is preferred by the MS community; thus, in the remainder of this dissertation, we refer to the three-type classification system of cGM lesions.

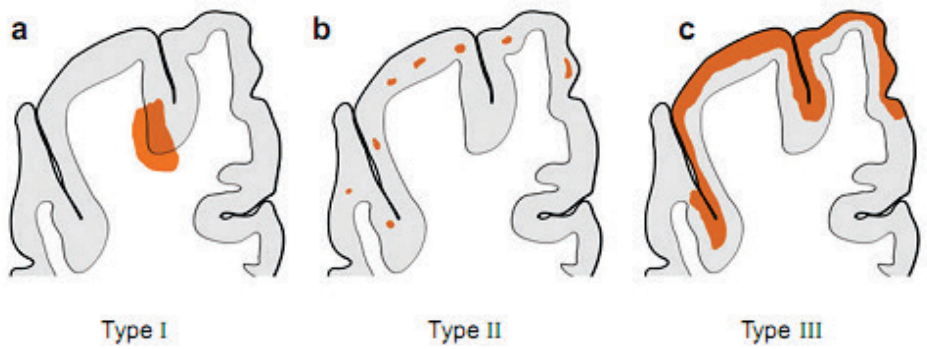

Figure 2.8 – Schematic representation of the three types of cortical lesions. Reproduced from [26].

In addition, they also made several key discoveries:

i. Cortical lesions differed from WM lesions in their inflammatory cell profiles (cortical lesions had 13 times fewer lymphocytes and 6 times fewer microglia/macrophages).

ii. Cortical lesions rarely contained perivascular cuffs (i.e., the area surrounding an inflamed blood vessel), which are common in WM lesions.

iii. Cortical lesions contained transected neurites (axons or dendrites), and their numbers correlated with the inflammatory activity of the lesion.

iv. Compared with myelinated cortex, cortical lesions showed a significant increase in neuronal apoptosis (programmed cell death).

v. Demyelination can occur in MS brains in the absence of extensive inflammatory cell infiltration.

vi. Activated microglia were found in the cortex and they extended processes to and ensheathed neurites and neuronal cell bodies.

In 2003, Bo and colleagues went on to characterize the degree of inflammation present in the different types of cortical lesions [33, 38]. Demyelinated intracortical MS lesions (Types II and III) were not associated with significant lymphocyte (T- or B-cell)

infiltration when compared with adjacent myelinated cortical areas, or with cortical areas in control brains. In contrast to their findings that the cortex in Type I lesions did have significantly more T-cells than in nondemyelinated GM and that the distribution of Type I lesions is similar to other WM lesions that follow the small cerebral veins, this finding suggests that, while the pathogenesis of lesion development may be the same for WM and Type I cortical lesions, it appears to be different for intracortical lesions. To further that hypothesis, they also noted that Type I lesions extended radially from a central vessel without any apparent respect for the WM/cortex border, whereas intracortical lesions showed a "striking and enigmatic" respect for said border [38].

In a separate study, the same authors reported on the distribution they observed of 109 cortical MS lesions that they obtained using immunohistochemistry on 78 blocks from 20 patients with MS [33]. They found that, of the four locations they investigated (cingulate gyrus, frontal cortex, parietal cortex, temporal cortex), the cingulate gyrus contained the most prominent demyelination of both GM and WM. As for distribution, they found that, of the total demyelinated area, Type I lesions accounted for 14.4%, Type II for only 1.2%, and Type III for the majority of the lesions at 84.3%.[ix]

The histopathological studies reviewed to date focused on small block tissue samples, and—although their results were illuminating—high-powered magnification of immunohistochemically stained sections place the onus on the reader to visualize the extent of pathology present in the entire brain. The large hemispheric or double hemispheric brain sections presented by Kutzelnigg et al. in 2005 allowed the MS community to appreciate the full extent of cortical demyelination [39]. In their study, LFB was used for basic classification of demyelination, but immunocytochemistry was performed for a more detailed examination of the pathology, which is illustrated in the *camera lucida*[x] illustration in Figure 2.9. They found that, while patients with RRMS had

[ix] Subpial lesions that only involved the outer layers of the cortex accounted for 67%, while the lesions that extended throughout the full width of the cortex accounted for 17.3%.

[x] Latin for "light room", the *camera lucida* is an artist's aid dating from at least the early 19th century. In its simplest form, the artist looks through an eyepiece that allows them to view a projection of the object on their canvas. The simultaneous view of their illustration with a superimposed virtual image of the object, both being in focus, results in a verbatim tracing of the object.

relatively little cortical demyelination, the PPMS and SPMS cohorts they inspected at autopsy had extensive cortical demyelination, in some cases accounting for up to 68% of the total area of forebrain examined. Furthermore, though only observed qualitatively, the most prevalent type of cortical lesion was the Type III subpial demyelinating lesion that often extended over multiple gyri.

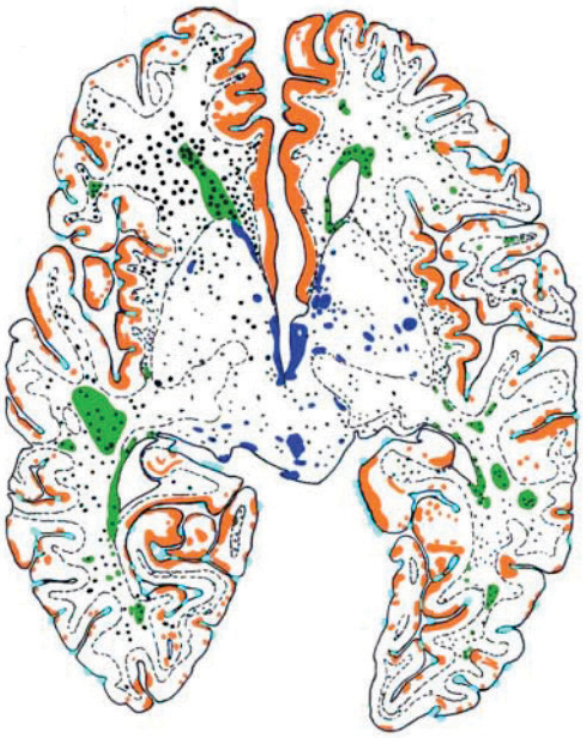

Figure 2.9 – Schematic map of the pathology in the brain from a patient with PPMS. Orange: cortical demyelination. Green: focal demyelinated plaques in the WM. Blue: lesions in the deep GM. Dark blue dots: inflammatory infiltrates in the brain. Cyan: inflammatory infiltrates in the meninges. Reproduced from [39].

Though difficult to make out, the meningeal inflammation depicted in cyan in Figure 2.9 was found, in their study, to be present at all stages of the disease. A separate case study did report conflicting results, noting meningeal inflammation only in their patients with SPMS [40]; however, the subsequent findings of a much larger study (N = 138) on cortical demyelination in early MS that was published in the New England Journal of Medicine, corroborate that meningeal inflammation is indeed present at all stages of the disease [41]. Not only is this discovery novel, revealing previously unknown facts about

inflammation outside of the parenchyma in MS, but that these inflammatory infiltrates in the meninges have been found to be topographically associated with Type III subpial lesions [39-44] hints at a unique pathogenesis for this particular type of cortical lesion. Indeed, ectopic[xi] B-cell follicles (i.e., one of the inflammatory infiltrates mentioned above) were detected in the meninges of SPMS patients (with a predilection for the depths of sulci) and, interestingly, were all found adjacent to large Type III cortical lesions, suggesting that soluble factors diffusing from these structures have a pathogenic role[xii]. It is also worth noting that, clinically, in comparison with subjects without the ectopic B-cells, subjects presenting with them had a younger age of MS onset, a younger age of irreversible disability onset, and a younger age at death [42].

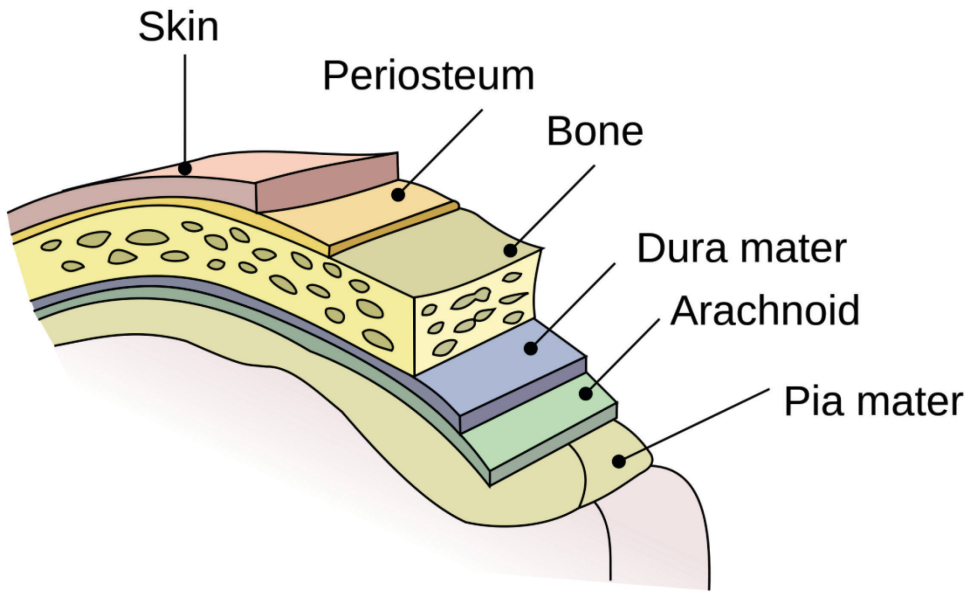

Figure 2.10 – The meninges. The meninges consist of three layers: the dura mater, the arachnoid mater, and the pia mater [47].

[xi] In an abnormal place or position

[xii] A purported explanation for these abnormally placed B-cells has been the discovery of the presence of dendritic cells in the inflamed meninges [44] as well as in nonlesional GM in patients with MS [45]. Dendritic cells are key players in both the induction and maintenance of autoimmunity and are thought to activate T- and B-cells as well as to organize ectopic lymphoid tissue [46].

Further to Peterson et al., several others have also characterized the cellularity of cortical lesions [48, 49]. Wegner and colleagues focused on Type I cortical lesions, but they did report substantial glial, neuronal, and synaptic loss within lesional cortex (36%, 10%, and a striking 47%, respectively), with no such changes in nonlesional cortex [49]. Vercellino et al. focused on Type III subpial lesions, and despite finding a significantly reduced neuronal density in cortical lesions compared with normal cortex (~20% reduction), they found no significant differences in synaptic densities [48]. Wegner et al. also tried to quantify the extent of neuronal loss by a histological measurement of cortical thickness. They reported an overall relative neocortical thinning of 10% for patients with MS, and while one might be tempted to presume that the hypocellularity of the cortical lesion accounts for this thinning, no significant correlation was seen between the mean cortical thickness and the mean extent of cortical demyelination in the areas measured in the patient group [49].

Thus far, the picture presented for cortical demyelination in MS has been one where the inflammatory cell content is much lower than in WM plaques [32, 38], where the few inflammatory cells that are present are mostly activated microglia [50] rather than infiltrating lymphocytes and macrophages [32, 33, 37, 38], and where the demyelination is driven in part by meningeal inflammatory infiltrates [40-44, 51]. Combined with the evidence for neuronal degeneration and glial loss [32, 33, 48, 49, 52], this has led to the theory that there is a pathogenesis in MS where neurodegeneration proceeds independently of inflammation, though this theory is not widely accepted[xiii]. It should also be noted that all of these observations were from studies that relied on postmortem tissue analysis from patients with longstanding, end-stage MS.

As one would expect, obtaining postmortem tissue samples for pathological examination from patients with early MS is rare, given that death occurs much less

[xiii] While there is no lymphocyte and macrophage infiltration, microglial activation is still a form of inflammation and could in fact lead to oxidative stress, and thus neurodegeneration. Moreover, other possible infiltrates such as the diffusion of soluble factors into the cortex from the lymphoid follicle-like structures in the arachnoid could be part of the complex inflammatory process and would fit well with the pathology observed.

frequently than at the end-stages of the disease. Nevertheless, the pathology of cortical MS lesions at the early stages of the disease has been characterized. Popescu in 2011 reported a single case study wherein biopsy was being performed to exclude neoplasm and, incidentally, an inflammatory subpial demyelinating lesion was found [53]. The same group went on to publish a much more comprehensive study of inflammatory cortical demyelination in early MS, wherein they retrospectively examined the biopsy tissue from 138 patients (563 patients were screened, but only 138 patients had sufficient cortex). They found cortical demyelination in 53 of those patients, interestingly, with T-cell infiltrates and myelin-laden macrophages present in all lesion types [41]. This finding is in stark contrast to the previously supposed theory that cortical demyelination was devoid of inflammation. Their study did, however, corroborate the finding that meningeal inflammation was not only present, but also topographically associated with subpial lesions.

Table 2.1 – Postmortem pathological cortical GM lesion prevalence numbers.

Study	MS cohort	# subjects	# lesions	Type I	Type II	Type III
Kidd et al. 1999 [37]	end stage	12	478	56%[a]	12%	32%
Peterson et al. 2001 [32]	end stage	7	48	27%	25%	48%
Bo et al. 2003 [38]	5 SP, 5 PP	10	60	25%	7%	68%[b]
Bo et al. 2003 [33]	10 SP, 7 PP, 3 RR	20	109	16%	17%	67%[b]
Geurts et al. 2005 [54]	5 SP, 2PP, 2 unknown	9	90	30%	13%	57%
Kutzelnigg et al. 2005 [39]	11 AMS, 6 RR, 14 PP, 20 SP	52				++[c]
Vercellino et al. 2005 [48]	3 RR, 3SP	6				>90%[b]
Wegner et al. 2006 [49]	end stage	22	161	38%	18%	44%
Gray et al. 2008 [50]	PP, SP, unknown	30	34	56%	15%	29%
Lucchinetti et al. 2011 [41]	early MS	138	104	50%	16%	34%

[a]Kidd et al. used a seven-type system which were combined into the three type system as follows: Type I = types 1,5; Type II = type 6; Type III = type 2,3,4,7.
[b]This study used a four-type system, thus, Type IV (affecting the entire width of the cortex) was included as part of Type III for calculations.
[c]Quantitative numbers were not provided, but the authors noted "demyelination mainly affected the subpial layers of the cerebral cortex…"
AMS = Marburg's acute multiple sclerosis, **RR** = relapsing remitting, **SP** = secondary progressive, **PP** = primary progressive

Extralesional Pathology

To think that MS is strictly a lesional disease would be as nearsighted as the MS-ologists of the past who said that MS was merely a disease of the WM. From at least as early as 1914, the extralesional pathology of MS has been well-recognized. Dawson noted diffuse alterations of the myelin sheath and glial cells [24]. Allen and McKeown further characterized the histology of the nonlesional WM in MS, remarking that, while the WM is macroscopically "normal-appearing," at the microscopic level, there is mild inflammation, gliosis, and foci of demyelination [55, 56]. Still more histopathological studies have found diffuse axonal injury [39, 57, 58] as well as diffuse inflammation (T-cells, macrophages, and profound microglial activation) in the normal-appearing WM [39].

As already mentioned above, the pathology present in the GM is not restricted to focal plaques in the cortex. Widespread microglial activation [50] and dendritic cells [45] are found in the nonlesional cerebral cortex of patients with MS, as is pathological evidence for neocortical thinning [49]. Also mentioned previously is the widespread meningeal inflammation [39-44], which was even noted a century ago as being diffuse whenever present [24]. Widespread demyelination has also been reported in the cerebellar cortex and spinal cord [59, 60]. The deep GM nuclei, specifically the thalamus, where an approximate 20% reduction in volume and neuronal density has been reported in patients with MS [61], are also involved.

2.2 MAGNETIC RESONANCE IMAGING

In 1895, Wilhelm Conrad Rontgen's discovery of x-rays introduced a novel and most powerful capability to the practice of medicine: the ability to visualize the interior of the human body without surgical intervention. Today, numerous noninvasive methods exist to help visualize the vast array of components that make up the human body. The following section focuses on MR imaging, a flexible and robust imaging modality that has proven particularly useful for the in vivo study of soft tissue.

Magnetic resonance imaging relies on the premise that atoms containing an uneven number of protons and/or neutrons (e.g., ^{1}H, ^{13}C, ^{23}Na) have a small magnetic charge. When placed in a large external magnetic field, these nuclei align themselves along that field[xiv] and, with the help of a stimulus, create a signal that can be detected by specialized equipment. This signal can then be used to construct an image. The entire procedure is noninvasive and, though a bit noisy and claustrophobic, presents no known health risks under normal operating conditions [62].

Magnetic resonance imaging can provide high-resolution images with excellent contrast between different soft tissues of the body, including pathologically altered tissues. It has become an extremely important tool in clinical medicine, especially neurology, and its value has been recognized by the awarding of multiple Nobel Prizes to those involved in its development.

2.2.1 HISTORICAL PERSPECTIVE

The research leading up to the development of MR imaging was simultaneously pioneered by two different groups in the United States in the 1940s. Felix Bloch and Edward Mills Purcell shared the 1952 Nobel Prize in Physics for their independent contributions to the discovery of nuclear magnetic resonance (NMR) in condensed

[xiv] The nuclei can arrange themselves either parallel or antiparallel to the magnetic field. The difference between the number of nuclei aligned with the field and those against the field is very small (~7 ppm) [62], yet that is still enough to create a net magnetization along the axis of the external magnetic field.

matter.[xv] Nuclear magnetic resonance went on to be, and is still today, used in chemistry and physics for studying molecular structure and diffusion. It wasn't until the 1970s that its applications to human imaging were shrewdly pursued by American chemist Paul C. Lauterbur [63]. In fact, the machine he used was the NMR spectrometer in the chemistry department at the State University of New York at Stony Brook. His insight into using magnetic field gradients to determine spatial localization allowed him to go beyond Herman Carr's one-dimensional (1D) NMR images of the 1950s and publish the first 2D (and subsequently 3D) NMR image in 1973 (Figure 2.11).

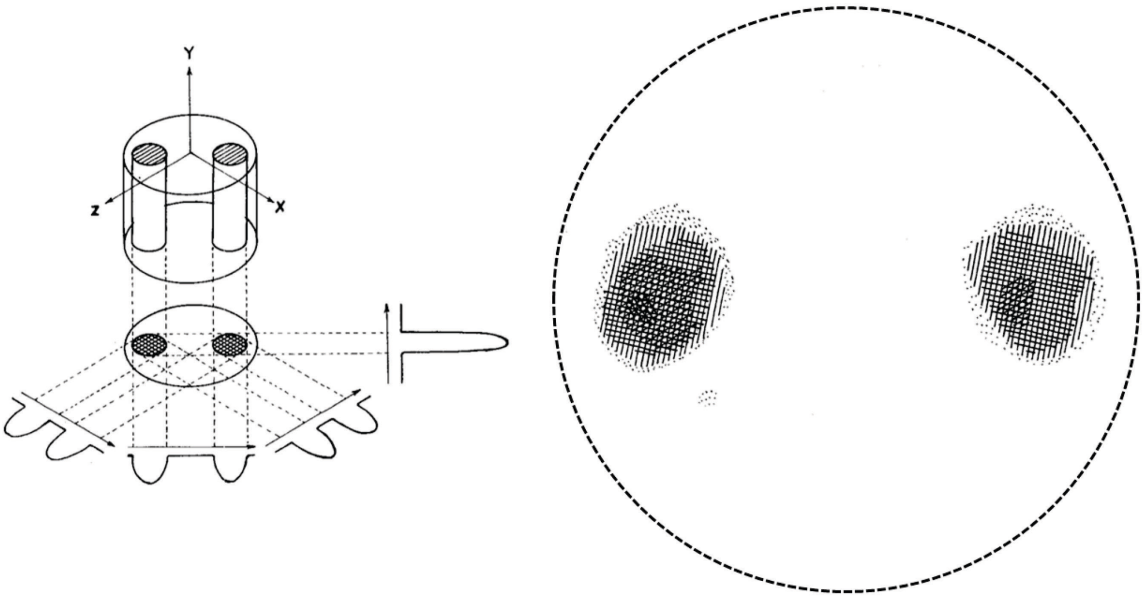

Figure 2.11 – First 2D NMR images published in Nature, 1973. Left: The relationship between a 3D object, its 2D projection along the y-axis, and four 1D projections at 45° intervals in the XZ-plane. Right: The proton NMR zeugmatogram[xvi] of the test object. The object consisted of a glass tube (4.2 mm inside diameter) filled with D_2O, or "heavy water," and within that, two thin-walled glass capillaries of H_2O (1 mm inside diameter). The dashed circle was not part of the original publication but represents the larger glass tube, while the two hashed ellipses depicted within are the reconstructions of the two D_2O-containing capillaries. No other imaging technique in existence at that time could distinguish between two different kinds of water, a particularly important achievement given that the human body consists mostly of water. Figures modified from [63].

[xv] Bloch and Purcell's work on NMR in liquids and solids (1946) was preceded a decade earlier (1938) by that of Isidor Rabi, who was the first to describe and measure NMR in molecular beams, for which he was awarded the Nobel Prize in Physics in 1944.
[xvi] Lauterbur proposed the technique be known as zeugmatography, from the Greek *zeugma*, meaning "that which is used for joining." The rationale was that the object was constructed from the coupling of the main field and the gradient in the second field.

During the 1970s, most of the work on MR imaging took place in academia, particularly in the United Kingdom. It was British physicist Sir Peter Mansfield who developed a mathematical technique that would allow scans to take seconds rather than hours and produce higher quality images than Lauterbur had. For their research on nuclear magnetic resonance imaging[xvii], and perhaps because it took a while for the clinical value of MR imaging to be appreciated, Lauterbur and Mansfield were awarded the Nobel Prize in Physiology or Medicine some 30 years later in 2003.

2.2.2 PHYSICS

Atoms, the once believed indivisible particles of the universe, spin, or precess, around their own axes and in so doing create a small magnetic field termed a magnetic moment. In reality, the magnetic moment of an atom is the vector sum of the moments of the elementary particles it comprises. The magnetic moment, or dipole, is often described using the spin quantum number, or spin. Atoms with magnetic spins of +/- ½ are unique in that they can exhibit magnetic resonance; that is, they are able to absorb and then re-emit electromagnetic energy. Because the ^{1}H protons are by far the most abundant nuclei in the human body from which NMR signals can be generated (mostly from H_2O), imaging based on the proton is the most practised case and will be the subject of further discussion [2, 62].

In order for these hydrogen protons to display NMR, they must first be placed in an external static magnetic field (B_0). In the presence of this field, two very important phenomena happen [62]:

1. These atoms precess at a well-characterized frequency known as the Larmor frequency, which is proportional to the gyromagnetic ratio of the atom in question and the field strength of B_0. For a hydrogen proton, the Larmor frequency is 42.575 MHz per Tesla (T).

[xvii] Although the chemistry and physics communities tend to refer to nuclear magnetic resonance (NMR), in the imaging community, the word "nuclear" is usually omitted because of the associated stigma—hence, the common abbreviation MRI rather than NMRI.

2. The precessing atoms align themselves to be either parallel (spin-up) or antiparallel (spin-down) to the main axis of B_0. The slight tendency of the spins to align themselves with the direction of B_0 creates a net, or bulk, magnetization (M_0) in the direction of B_0.

Since M_0 is in the same direction and several orders of magnitude lower than B_0, it is an extremely difficult signal to measure. It is here where the significance of the first phenomena can be seen. By using a magnetic pulse ($B_\perp$) that is tuned to the Larmor frequency and applied perpendicularly to the main axis of B_0, protons will begin to spin along the main axis of $B_\perp$ instead of B_0, causing M_0 to have a component that spirals into the $B_\perp$ plane ($M_\perp$). This component is often referred to as transverse magnetization, while the component along the M_0 direction is called longitudinal magnetization (M_Z). A receiver also tuned to the Larmor frequency and in the $B_\perp$ plane would now detect the signal of interest.

Once the $B_\perp$ pulse[xviii] is removed, the precessing protons, and by extension the macroscopic magnetization of a group of these spins, will return to equilibrium via an exchange of energy with its environment (i.e., in the presence of B_0, $M_\perp$ and M_Z will return to M_0). This phenomenon is termed relaxation and is governed by two distinct processes:

1. the recovery of longitudinal magnetization M_Z and

2. the decay of transverse magnetization $M_\perp$.

The recovery of longitudinal magnetization refers to the component of magnetization oriented along the main static field B_0 and is the result of thermal interaction between the nuclear spins and their surrounding environment. It is termed spin-lattice relaxation, or more commonly T_1 relaxation, as the time it takes for the longitudinal magnetization

[xviii] Also called a radiofrequency pulse as the Larmor frequency is within the 3 kHz to 300 GHz range of radio waves. Recall that the Larmor frequency for hydrogen is 42.575 MHz/T; thus, for a common 3.0 T MR imaging system, it would be 3 x 42.575 MHz = 127.725 MHz.

to return to equilibrium (M_0) is governed by an exponential decay with a characteristic time constant, T_1 [62].

$$M_z(t) = M_0 + [M_Z(t_0) - M_0]e^{-\frac{t}{T_1}}$$

The decay of transverse magnetization $M_\perp$ refers only to the component of magnetization that is perpendicular to the static magnetic field B_0. Interactions of one spin with another cause those spins in the $B_\perp$ plane to lose phase coherence, or to dephase. Thus, this type of relaxation is termed spin-spin relaxation, or more commonly T_2 relaxation, as the exponential decay that governs this process has a characteristic time constant, T_2 [62].

$$M_\perp(t) = M_\perp(t_0)e^{-\frac{t-t_0}{T_2}}$$

Magnetic resonance imaging generates tissue contrast largely as a result of variations in both the amount of water in the tissue (i.e., proton density) and the T_1 and T_2 relaxation times, which are influenced by the macromolecular environment of the water molecules [2].

T_1-Weighted Imaging

T_1 relaxation times are generally longer in pure liquids than in biological tissues. This is as expected as in a liquid there is rapid motion, rotation, and vibration of molecules, which results in short spin-lattice interaction times and thus longer relaxation times (i.e., pure liquids have the smallest signal and appear hypointense on a T_1-weighted (T1w) MR image; see Figure 2.12). In tissue, there are other pathways for relaxation present that result in shorter T_1 relaxation times, and of course the composition of the tissue also affects the time. These differences in T_1 relaxation times for different tissues are the bases for contrast in T1w imaging. In terms of T_1 contrast, in general, cerebrospinal fluid (CSF), which has a long T_1, will be the darkest, followed by GM and then WM, which have progressively shorter T_1 relaxation times.

T_2-Weighted Imaging

The spin-spin interaction is much greater for spins with restricted motion (e.g., in solids, gels, semi-solids) than for spins that are relatively free (e.g., in liquids). This increased interaction in semi-solids results in an extremely fast decay (short T_2 time on the order of microseconds) which generally precludes direct observation in MR imaging. In liquids, the spins experience less spin-spin interaction (i.e., dephasing) resulting in T_2 relaxation times that are long enough (>10 ms) to yield a coherent signal for imaging. A T_2-weighted (T2w) image of the brain will have bright CSF and often little contrast between the comparatively hypointense WM and GM tissues.

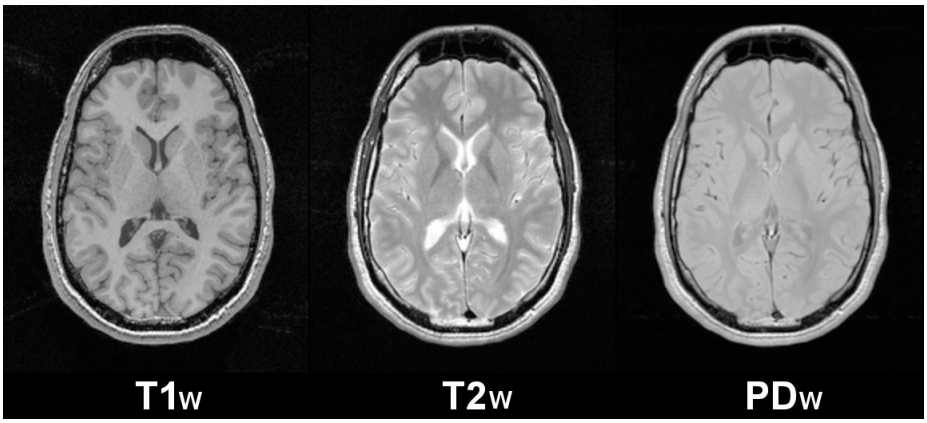

Figure 2.12 – T1w, T2w, and proton density–weighted (PDw) images of the author's brain. PDw imaging is used to image all the MR imaging-visible nuclei of hydrogen atoms available in a sample and can be particularly useful in detecting pathology, as will be shown later.

FLAIR and DIR

In the examples described above, the focus was on the commonly used 90° radiofrequency pulse[xix] that has the effect of converting longitudinal magnetization M_Z into transverse magnetization $M_\perp$. In an inversion recovery sequence, a 180° pulse is first used to invert all the longitudinal magnetization M_Z to $-M_Z$; then, after a specified period of time, the 90° pulse is administered with the consequence that whatever longitudinal magnetization (-Mz) has recovered to exactly zero will produce zero

[xix] A 90° pulse is obtained by applying the radiofrequency pulse for a specific amount of time such that enough energy is absorbed by the precessing protons that, instead of being aligned with the static field B_0, they are aligned perpendicular to it (hence, 90°). By leaving the radiofrequency pulse on for a longer period of time, the precessing protons could absorb enough energy to be aligned antiparallel to B_0; this would be called a 180° pulse.

transverse magnetization. One of the advantages of such a technique is that it can be used to suppress the signal from a tissue of interest (e.g., the signal from fat), as tissues have unique T_1 relaxation times that vary with field strength. One practical sequence that attempts to suppress the signal from the CSF is the FLuid Attenuated Inversion Recovery sequence, or FLAIR. By waiting until the longitudinal magnetization of CSF protons has recovered just enough from the 180° pulse so that it is zero, the effect of the 90° pulse will be that none of the longitudinal magnetization M_Z from the CSF will be converted into $M_\perp$, and consequently, almost none of that signal will be recorded [62].

Another application of the inversion recovery sequence would be to suppress the signal from two different tissues. Using the same principles described above, two 180° pulses would be applied before the 90° pulse is administered. The time between the first and second 180° pulse would determine the first tissue signal to suppress, and the time between the second 180° pulse and the 90° pulse can be tuned to suppress the signal from a second tissue.[xx] The Double Inversion Recovery sequence, or DIR, is just such a sequence where the signals from the WM and the CSF are both suppressed. As DIR has been used to investigate GM pathology in MS with varying degrees of success, it will be discussed later.

Magnetization Transfer Imaging

As already mentioned above, the most practised form of MR imaging is that based on the hydrogen proton, more specifically, the hydrogen protons in water. Biological tissue, however, comprises a multitude of different proton environments, and though they can be modelled in different ways, we will focus on the two-pool model of magnetization exchange [64].

The majority of hydrogen protons in the body are found in small mobile molecules (mostly water). The protons are relatively unrestricted in their movement and have

[xx] Technically the inversion times are not independent but rather depend on both the species that are trying to be suppressed.

sufficiently long T_2 relaxation times (>10 ms) to facilitate imaging.[xxi] This pool of relatively "free" spins is what most MR imaging sequences are able to image directly.

Hydrogen protons are also found in larger semi-solid macromolecules, though in smaller proportion. In contrast with the free protons, these "restricted" protons are less mobile, and T_2 decay happens too quickly for direct imaging (i.e., a couple of μs). However, coupling between the protons in the restricted pool and those in the free pool allows the spin state of the former to influence the latter (Figure 2.13a). Recall that the basis of NMR is that spins are able to absorb and emit energy. In the case of magnetization transfer (MT), the restricted protons can be selectively excited (i.e., absorb energy) and then allowed to relax (i.e., release energy). As they relax, the energy they emit can in turn be absorbed, or transferred, to the protons in the free pool. The result of this transfer of magnetization is that, when we go to image these free pool protons, they do not start with the same longitudinal magnetization M_Z as in a system where we did not selectively excite the restricted pool a priori; thus, there is a decrease in the signal observable with MR imaging.

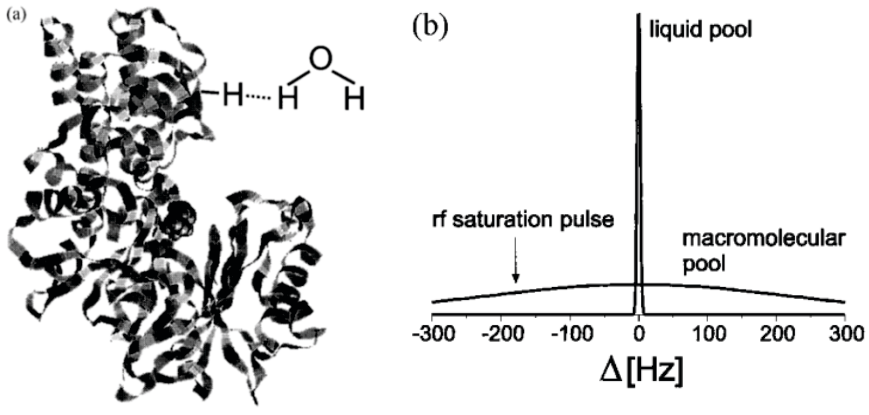

Figure 2.13 – Magnetization transfer. (a) The exchange (dotted line) between the hydrogen proton attached to a macromolecule (i.e., the restricted, or semi-solid, pool) and water (i.e., the free, or liquid, pool). (b) The Larmor absorption spectrum for the liquid pool (thin peak of only a couple of Hz wide and centred on the Larmor frequency (Δ0 Hz) and for the macromolecular pool (broad absorption line). Reproduced from [64].

[xxi] The T_2 relaxation time is long enough so that spatial encoding gradients can be played out between excitation and acquisition before the signal has completely decayed. These gradients are what allow the localization of the signal from different points in space and what allowed Lauterbur in 1973 to obtain the first 2D NMR images.

In MT experiments, the intent is to indirectly manipulate the free pool by first saturating the restricted pool. This saturation can be achieved in a number of ways [64], but for the purpose of this review, we will focus on the off- and on-resonance pulsed waves. The classically presented method for proton saturation is as was previously described: A radiofrequency pulse is applied at the Larmor frequency (i.e., on-resonance). Special binomial on-resonance pulses (e.g., $1\bar{1}$, $1\bar{2}1$, $1\bar{3}3\bar{1}$) are used to ensure that the magnetization of the free protons remains unchanged (i.e., they are self-compensating—total angle of 0°), while restricted protons are saturated due to the much faster dephasing and T_2 decay of these protons [65].

Pulsed saturation can also be applied off-resonance as brief (i.e. 10 to 30 ms), shaped radiofrequency pulses not at the Larmor frequency (Figure 2.13b). Again, these pulses leave the free protons unchanged because they are unable to absorb energy at the off-resonance frequency (i.e., they have a narrow Larmor absorption spectrum with a peak width of about 20 Hz), while the restricted protons are up to 10^6 times more sensitive to an appropriately placed pulse (i.e., they have a much broader Larmor spectrum with a peak width of tens of kHz) [64]. Off-resonance pulses are usually offset by a couple of kHz from the water resonance frequency, and the effective coupling between the restricted pool spins ensures the spread of saturation throughout the entire pool.

A complete characterization of the MT effect can be obtained by quantitative MT imaging where enough data is gathered to fit the parameters of a two- (or more) pool model [66, 67]. Unfortunately, this method is not yet feasible for large-scale whole-brain imaging studies mostly due to time constraints. Thus, the MT effect is most widely reported as the percent difference between images obtained with[xxii] and without MT pulses, known as the magnetization transfer ratio (MTR), and is a single, semi-quantitative index:

$$MTR = 100\% * \left[\frac{NoSat - Sat}{NoSat}\right]$$

[xxii] An image acquired with MT saturation alone does not produce very interesting contrast.

where Sat refers to the image acquired with a saturation pulse and NoSat is the image without the pulse. Though strictly speaking unitless, MTR is often reported as having percent units (p.u.).

In the brain, the two-pool model commonly refers to the free pool (e.g., intra- and extracellular water) and the restricted pool (e.g., cellular membranes, myelin, proteins, etc.). Magnetization transfer ratio, and more generally MT imaging, is focused on providing insight into the restricted proton pool, and the amount of MT occurring in macromolecular solutions is known to increase with higher concentrations of macromolecules [68, 69]. Though perhaps not an exclusive marker for myelin, MTR is still sensitive to it—the predominant element of the restricted pool in WM is the myelin sheath, comprising about 50% of its total dry weight [70].[xxiii] The value of an imaging method that can characterize myelin is obvious when investigating pathology that affects the myelin sheath, such as in MS. Still, MT imaging is not the only method that attempts to image myelin.

Myelin Water Fraction Imaging

Much of the insight for the development of the MT models came from looking at T_2 relaxation curves. As already mentioned, the T_2 process (i.e., transverse magnetization) is characterized by exponential decay and, importantly, is theorized as a single decreasing exponential. Indeed, in samples of pure water, a single exponential decay is observed; yet, it is possible to decompose the transverse magnetization decay from tissue (e.g., human, cat, or bovine WM, muscle tissue, etc.) into multiple reproducible exponential components [67, 71-74]. The most accepted view for this multiexponential observation is that there are multiple anatomically separate water components present

[xxiii] Myelin in situ has a water content of about 40%, and the 60% that is dry weight is mostly lipids (approximately 70% lipids and 30% proteins). The high lipid content is what gives myelin its insulating properties and its characteristic pale colour. For reference, most other membranes are protein dominant. In WM, where the myelin content is high (~50%), the lipid-to-protein ratio by dry weight is 55:40 (5% other), while in GM, where there is less myelin and an abundance of other cellular membranes, the lipid-to-protein ratio drops to 35:55 (10% other). In terms of the composition of the lipids and proteins in myelin, cholesterol is probably the single largest lipid component (~25%), while PLP and MPB account for the majority of the proteins in myelin (60-80%) [70].

in the tissue sampled, and several studies have investigated these components in normal human brain tissue [72-76].

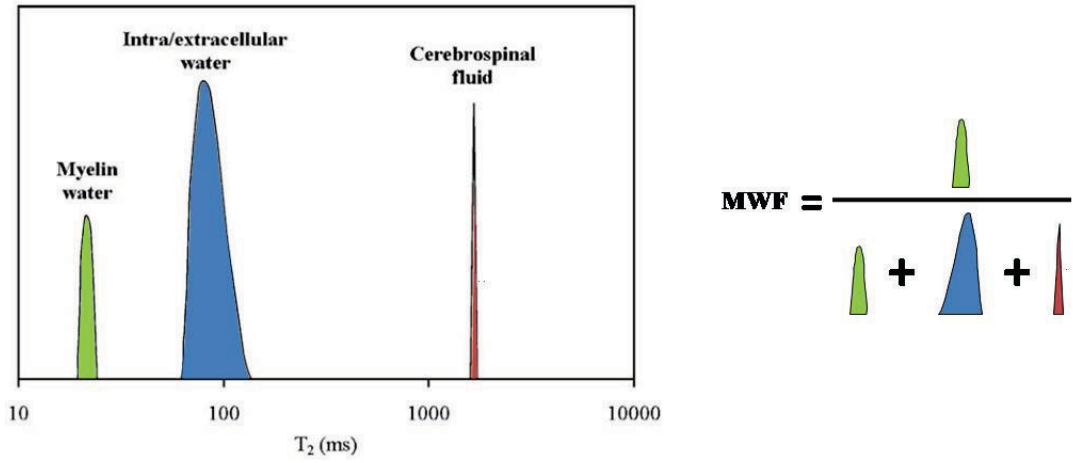

Figure 2.14 – Myelin water fraction (MWF). Left: An example of a T_2 distribution of human WM in vivo with labelled regions corresponding to compartments of WM. Note the logarithmic scale for T_2 relaxation time. Right: Schematic representation of the MWF as the ratio of the signal from the myelin water component to the total observable T_2 signal (all components). Modified from [72].

Figure 2.14 shows three distinct T_2 components in normal human brain. Their interpretation is as follows: The component seen at T_2 > 2,000 ms is associated with CSF. Recall that protons in a less restricted environment will experience fewer spin-spin interactions due to their rapid tumbling and increased motion and thus will have longer T_2 relaxation times. Correspondingly, the largest component with a peak centred between 70 ms to 90 ms is that for intra-/extracellular water (e.g., axonal water).[xxiv] Finally, water believed to be trapped within the myelin bilayer itself is conceivably the most restricted—and thus experiences the shortest T_2 relaxation time—and is believed to be responsible for the myelin water component seen between 10 ms and 50 ms.

The myelin water fraction (MWF) is defined as the ratio of the T_2 signal in the myelin water component (formally defined as being between 0 ms and 50 ms) to the total signal in the T_2 distribution (Figure 2.14). If water trapped in the myelin bilayer is proportional

[xxiv] In studies of peripheral nerves, there is a differentiation between axonal water and extracellular water with peaks at 80 ms and 250 ms respectively [77, 78]. It is likely that signal-to-noise limitations available for in vivo T_2 relaxation studies are responsible for the difficulty in separating these components in human brain [72].

to myelin content, then one can see how the MWF is a potential candidate for the in vivo quantification of myelin.

2.2.3 IMAGING THE PATHOLOGY OF MS

After its advent in clinical practice in the 1980s, the role of MR imaging in MS has been extensive, ranging from use in diagnosis to monitoring progression and treatment to therapy development in large-scale clinical trials.

Prior to MR imaging, the diagnosis of MS was based on clinical features alone. In 1965, Schumacher et al. defined the basis of what would become the gold standard for the diagnosis of MS [79]: two or more attacks affecting two or more separate sites within the CNS (this wording has evolved to be more commonly presented as a dissemination in time and a dissemination in space). Later criteria [80, 81] accepted that laboratory tests (e.g., imaging, CSF examination, etc.) could be used to supplement evidence for diagnosis. Even today, there is still ongoing refinement and debate[xxv] about diagnostic criteria (some even argue that autopsy is the only proved diagnosis), but the most accepted are referred to as the McDonald criteria [81] (latest modifications made in 2010 [82]), in which MR imaging is heavily involved.

White Matter Lesions

Signal abnormalities confirmed as lesions were first observed in MR images in the 1980s, when researchers found close similarities between the distribution of lesions at postmortem examination of a formalin-fixed brain and the distribution of aberrant signal on an MR image of the same sample [83]. The most common form of signal abnormality seen on MR images is what is commonly referred to as a T_2-weighted, T_2 hyperintense, or simply, T_2 lesion of the cerebral WM. As the names imply, T_2 contrast is the key feature that provides the high sensitivity for depicting focal WM plaques. Recall that T_2 decay is driven by spin-spin interactions and tissues that have higher liquid content will thus appear hyperintense on T2w MR images. Edema, the collection of an

[xxv] For a history of the evolution of diagnostic criteria, the reader is referred to [9] and [82].

excess of watery fluid and part of the process of acute inflammation, is believed to be the pathological component driving the T_2 hyperintensities.

Several types of WM lesions have been described pathologically, and although clinical MR imaging is highly sensitive to detecting pathological abnormalities in soft tissue (recall that the dominant mechanism for the contrast between healthy and pathological tissue is the concentration and macromolecular environment of the water protons), its weakness is its limited pathological specificity. Still, some degree of specificity can be obtained by exploiting specific MR image contrast mechanisms or, more practically, by using different imaging sequences (i.e., modalities). Not every lesion will appear the same on each modality; thus, each modality has its own advantages for imaging lesions at a particular stage of pathological development.

Considered the most destructive because of a breakdown in blood-brain barrier (BBB) is the active inflammatory lesion[xxvi]. In addition to such a breakdown, these lesions are characterized pathologically by a heavy infiltrate of macrophages with myelin debris, lymphocytes, and large astrocytes [84]. While MR imaging lacks the specificity to image the inflammatory infiltrates directly, it is possible to identify breaches in the BBB. These compromised areas can be best imaged with the help of a contrast agent, specifically, gadolinium-containing chemical compounds (i.e., chelates). Because the chelates contain unpaired electrons, water protons that come within the vicinity of the gadolinium will experience a marked decrease in their T_1 relaxation time. Under normal circumstances, gadolinium chelates do not enter the CNS because of the tight BBB, but in MS, the breakdown of the BBB allows the agent to pass through, thus decreasing the T_1 relaxation time of protons in its proximity, which translates to a hyperintense signal on a T1w image [2]. Though beyond the scope of this review, there is much interest in quantifying the amount of these gadolinium-enhancing lesions as they are considered a marker of active and ongoing inflammation in an MS brain. As a result, many clinical

[xxvi] Acute active lesions are defined as having macrophages that contain early and late degradation products throughout the entire lesion. Chronic active lesions, on the other hand, have macrophage accumulation at the plaque edge, but diminishing concentrations towards the inactive centre [84]. Chronic active lesions often appear as ring (or arc) gadolinium-enhancing lesions on MR images.

trials have used this quantification as an endpoint to evaluate the efficacy of their therapies [85, 86].

After the transient breakdown of the BBB, and after the inflammation process has had a chance to evolve, the lesion will tend to appear hyperintense on a T2w scan. The border between lesion and ventricular CSF can be difficult to distinguish on a regular T2w scan; however, the suppression of CSF that results from FLAIR imaging allows for these lesions to be more conspicuous, which is why many clinical centres add the sequence as part of their standard diagnostic protocol [2]. As these T2w lesions are perhaps the easiest to image, they have been well studied. The total number or volume of these T2w lesions is referred to as the T2w lesion load, and is often used as a marker of the burden of accumulated focal WM disease. The distribution of T_2 hyperintense WM lesions seen on MR images [87, 88] closely follows what is described pathologically; the majority are small (<5 mm) ovoid plaques that can be widespread but have a predilection for the periventricular areas.

The other modality that is traditionally included as part of the diagnostic protocol is a proton density–weighted (PDw) image. A PDw image is commonly obtained at the same time as a T2w image as the first echo in a dual-echo sequence. The signal from a desirable PDw image for lesion identification in MS is not, as the name implies, proportional to the total number of protons, but rather, is an intermediate weighting between T_1 and T_2 such that MS lesions appear bright but CSF appears dark. In combination with T1w and T2w images, it can be used to better discern WM lesions adjacent to CSF (specifically T_2 hyperintense lesions). Proton density-weighted imaging is used in many automated classification methods because of the complementary nature of the information it provides; however, it is increasingly being substituted or added to by FLAIR imaging, which can provide even more unique information.

Finally, hypointense lesions on T1w images are also commonly[xxvii] seen in the WM of patients with MS. Recall that T1w contrast is based on the spin-lattice interactions. In areas where the extracellular matrix has broken down, the T_1 signal will be quite low. These T_1 hypointense lesions, then, are thought to represent lesions where the inflammatory process has subsided and extensive axonal destruction has resulted in cell loss and an accumulation of extracellular water [89]. For these reasons, such lesions, when chronic, are also colloquially (and misleadingly) referred to as "black holes."

T_1 hypointense, T_2 hyperintense, and gadolinium-enhancing lesions represent the three most conventionally imaged WM lesions in the brains of patients with MS; yet, none of these imaging markers is specific to the hallmark of MS—demyelination. Magnetization transfer imaging, and specifically MTR, has been shown to be sensitive to changes in myelin content in WM [90-97]. Magnetization transfer ratio has been found to correlate well with both WM myelin content ($r = 0.84$, $p < .001$) and axonal count ($r = 0.66$, $p < .001$), but importantly, when MTR was regressed on axonal count and myelin content simultaneously, the association between MTR and axonal count disappeared, while MTR and myelin content remained significantly correlated (partial $r = 0.73$, $p < .001$) [95]. Furthermore, significant decreases and increases in MTR have been shown to be associated with de- and remyelination, respectively, of WM lesions on postmortem histology [90, 91, 95]. In another postmortem histology study, MTR was found to be the best predictor of myelin content in unfixed MS brain tissue amongst several MR imaging markers (i.e., MTR, T_1 and T_2 relaxation times, macromolecular proton fraction, and several diffusion weighted imaging parameters) [97]. Thus, although MTR may not be a purely specific marker for myelin since it may also be correlated with axonal loss and, to a lesser degree, inflammation and edema [98, 99], it does not appear to be related to the extent of gliosis [95] and does appear to be weighted towards myelination [92, 95, 99, 100].

[xxvii] T_1 hypointense lesions account for 20-30% of all T_2 visible lesions.

The Clinico-Radiological Paradox

The clinico-radiological paradox refers to the weaker than expected association between the conventional neuroradiological markers of MS (e.g., T2w lesion load, gadolinium-enhancing lesions, etc.) and clinical markers (e.g., disability, relapses, etc.).

On average, patients have clinical relapses every one to two years during the relapsing-remitting phase of the disease. High-frequency serial MR imaging studies have shown that lesions develop about 5 to 20 times more frequently than clinical relapses depending on the cohort studied and the MR imaging protocol used [101-104]. Regardless of study design, the conclusion is that, although patients with (RR) MS appear to have clinically active and quiescent periods, inflammatory lesions are developing and evolving almost continuously. In other words, more disease activity is taking place than is clinically apparent (Figure 2.15).

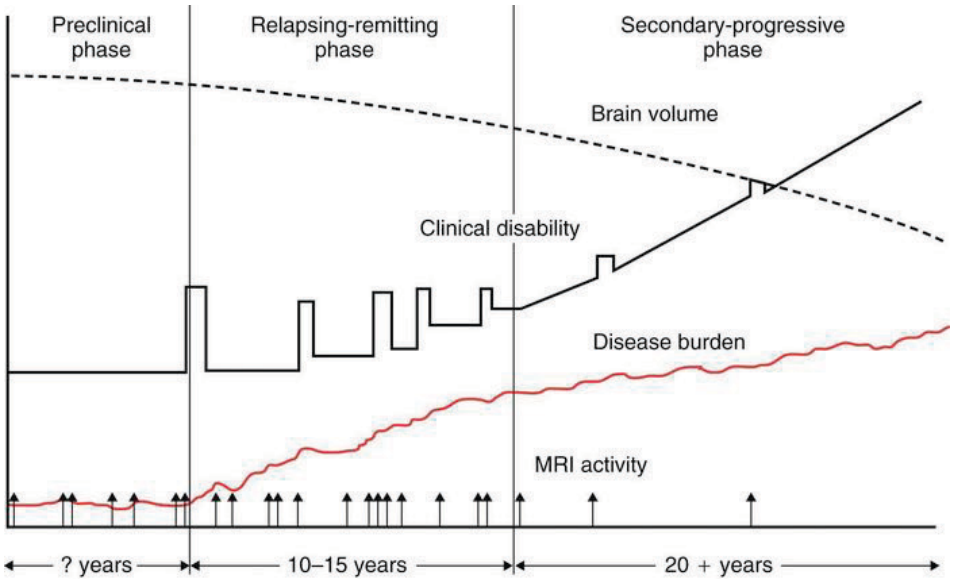

Figure 2.15 – Typical clinical and MR image course of MS. MR image activity (vertical arrows) indicates an inflammatory process as measured on brain MR imaging by a gadolinium-enhancing or a new T_2 hyperintense brain lesion. An increase in disease burden (red line) can be measured by the total volume of MR-imaged lesions and is an indication of permanent tissue damage. A loss in brain volume (dashed line) is also thought to occur early in the disease and gradually progresses over time. Importantly, MR image activity is typically more frequent than clinical relapses. Reproduced from the online source [105], which is an adaptation of the original figure in [106].

In addition to simply counting the number of clinical relapses as a marker of disease activity, the clinical consequences of MS are typically evaluated using the Kurtzke Expanded Disability Status Scale (EDSS) [107, 108].[xxviii] Studies investigating the relationship between EDSS-measured clinical disability and total T2w cerebral WM lesion load have yielded highly variable, and on average only moderate, correlations [109, 112, 113]. A recent paper by Caramanos et al. [114] optimized for a judicious investigation of this relationship found that, while T1w hypointense lesion load seemed to be a slightly better predictor of clinical disability than T2w hyperintense lesion load (concurrent with previous findings [115]), they were still at best able to account for only approximately 38% of the variance in the clinical disability.

The question of what accounts for the rest of the variance remains to be answered unequivocally, but explanations include the presence of extracerebral lesions (e.g., spinal cord, optic nerve, etc.), GM lesions, nonlesional WM or GM pathology that was not quantified by conventional means, cerebral redundancy and plasticity, and the fact that most of the brain performs functions that are not well reflected by routine clinical assessment. Of course, there also exists the possibility that the pathology may be present and could be accounting for the total variability we see in clinical disability, but we are simply unable at this point to image that pathology in vivo. As a prime example, consider the state of affairs some two decades ago when there was little to no capability of imaging cortical GM lesions in vivo.

Cortical Grey Matter Lesions

Recall the first modern pathological study of cGM lesions by Kidd et al. in 1999 [37]. In addition to their pathological characterization of the seven types of lesions they described, they also performed a conventional MR imaging analysis of the postmortem samples. Not surprisingly, very few of the lesions that were identified histologically

[xxviii] Invented in 1983 by physician and epidemiologist John F. Kurtzke, the EDSS is an ordinal scale with half-point increments ranging from zero (normal neurological examination results) to 10 (death due to MS). Steps 1.0 to 4.5 refer to patients that are fully ambulatory, while steps 5.0 to 9.5 are defined by the impairment to ambulation [107]. Despite the heavy reliance on ambulation and numerous criticisms [109-111], it is still the de facto measurement of clinical disability for patients with MS.

were visible on T2w images, in stark contrast to the high sensitivity of WM lesion detection. Among other things, Kidd et al. theorized that the insensitivity of the modality was due to the high cellular density of the cortex, which prevents sufficient expansion of the extracellular space to result in an increase in the relaxation time of the lesion—and thus sufficient contrast with the nonlesional GM.

Of course, it wasn't just Kidd et al., but the entire imaging community who had been unable to directly visualize cortical plaques on conventional modalities since the introduction of MR imaging into the clinic. Given the previously described pathological descriptions of the lesions (i.e., limited inflammation and BBB damage, minimal edema, sparse tissue loss due to the low amount of myelin and restricted axonal and neuronal injury), it was not surprising that these cortical plaques were being missed [2]. A marginal improvement in detection over T1w, PDw, and T2w imaging was noted in early studies that incorporated FLAIR, and those researchers attributed the increased sensitivity to the modality's ability to suppress CSF signal and its higher sensitivity to prolonged T_2 in tissue [116, 117]. But, without direct comparison with postmortem histology, the exact proportion of purely cortical lesions visible on MR images was unknown; luckily, in 2005, Geurts and colleagues performed just such a study that combined postmortem MR imaging with histopathology [54].[xxix] They studied 49 samples from nine cases of chronic MS, and of the 63 immunohistochemically identified purely intracortical lesions (Types II and III), only 3% and 5% were visible on T2w and FLAIR images, respectively. The authors even noted that, after side-by-side review of the histopathology and MR images, many of the intracortical lesions could not be identified retrospectively. Thus, while a slight improvement was seen with FLAIR imaging, it was clear that the majority of the cortical pathology in MS was being missed.

In vivo intracortical lesion detection wasn't really considered viable until a second study by Geurts and colleagues (also published in 2005) that used 3D DIR [119]. Although they

[xxix] Perhaps the first combined histology and MR imaging study that commented on cortical lesion detection in MS was done in 1991 by Newcombe and colleagues [118]. Unfixed postmortem brains were imaged using T2w imaging on a 0.5T magnet, and none of the 39 cortical plaques from eight brains identified on histology were detected.

did not have the luxury of histological confirmation, the authors evaluated the sequence's sensitivity to cortical lesion detection compared with 3D FLAIR and T2w images. With respect to Type I lesions (i.e., mixed WM and GM), they found a mean gain in sensitivity of 165% for DIR compared with T2w imaging (95% CI: 43%, 390%) and a less significant 39% gain over FLAIR (95% CI:-25%, 157%). Where the strength of DIR became obvious was in the improvement in detecting purely intracortical lesions (Types II and III), with a 152% gain over FLAIR (95% CI: 15%, 453%) and a 538% gain over T2w images (95%: 191%, 1296%). While considered a breakthrough for imaging cGM pathology in MS, the study by Geurts et al. and DIR in general is still considered far from ideal; the DIR sequence benefits from increased GM contrast, but it is hindered by the low signal-to-noise ratio (SNR) in addition to the many artifactual hyperintensities that make the specificity of lesion detection difficult.

Offering an improvement on some of the technical issues of the previously used DIR sequences was the work by Pouwels et al. in 2006 [120]. In addition to suppressing flow artifacts and signal intensity differences between slabs and reducing the scanning time, the authors were able to marginally increase the SNR and contrast-to-noise ratio in GM. Another study by Bagnoto et al. investigated the effect of SNR by averaging the signals from 12 sequential scans of the same subject [121]. Though they investigated only FLAIR and T1w sequences and not DIR, the authors found that the higher SNR (which translated to a threefold increase in the contrast-to-noise ratio) from the averaged scans led to increased lesion conspicuity. Unfortunately, obtaining 12 sequential scans of a subject is not feasible in a clinical setting. Still, this study paved the way for a flurry of studies that would try to increase cortical lesion detection by increasing the SNR via higher field strengths (i.e., 4.7 T to 8 T) [122-129].

The (staggering) 538% gain in cGM lesion detection of DIR over T2w imaging at 1.5 T led to the modality's dominance in detection and elucidated the insensitivity of T2w imaging [119]; however, it wasn't until a thorough postmortem verification study by Seewann et al. that the true sensitivity (or rather, insensitivity) of DIR became

apparent [130]. Overall, the authors found the sensitivity of DIR for all cortical lesions to be a disappointingly low 18% when evaluated prospectively, which increased to 37% retrospectively (i.e., after looking at the postmortem histopathology). They also reported that the pathological specificity for DIR was 90%. Both of these findings led to the fitting maxim regarding MS cortical lesions on DIR: "few but true" [131]. Indeed, most cortical lesions were, and still are, being missed, and when investigated further, it became apparent that subpial lesions (Type III) were the least detected (i.e., a sensitivity of only 7%) even though they are the most prevalent in postmortem studies. Thus, in vivo cGM lesion detection with DIR represents only a limited snapshot of the true cGM pathology.

Ultra-High Field Imaging

Standard clinical imaging is executed with magnetic field strengths of 1.0 T to 3.0 T. Ultra-high field magnets (i.e., main magnetic field strength $B_0 \geq 7.0$ T) have substantial advantages in terms of SNR, and thus, obtainable resolutions in a given period of time. Though dependent on many factors, the SNR is considered to be approximately proportional to the field strength (i.e., $SNR \propto B_0$).[xxx] As already described, Bagnato et al. reported a qualitative increase in cGM conspicuity on conventional images by increasing SNR via averaging at low field strength (1.5 T) [121], providing the impetus for increasing SNR to achieve higher sensitivity. Indeed, at 8.0 T, where the SNR is somewhere between 5 and 20 times greater than that of a 3.0 T system, this qualitative increase in sensitivity was also seen [126]. Importantly, though, a key study by Geurts et al. found that, when compared with histology, even the conventional images (i.e., T2w, PDw, FLAIR) at higher field strengths had a sensitivity for Type II and Type III cortical lesions below 10% [125]. Together, these studies suggest that, while SNR may contribute to increased sensitivity, it does not appear to be the dominating effect for detection. In

[xxx] The SNR in MR imaging, approximated as the ratio of the signal amplitude to the standard deviation of the noise, has two basic dependencies: (i) physical/instrumental parameters and (ii) image sequence parameters. The main magnetic field strength B_0 is one of the dominant instrumental parameters. The signal amplitude is proportional to the square of B_0, while the standard deviation of the noise is proportional to the square root of a combination of two primary noise sources: (i) the receiver coil and (ii) the body. If the noise from the body dominates, then the SNR is simply linearly proportional to B_0; however, if the coil noise dominates, then the SNR is proportional to $B_0^{7/4}$. As the latter is often the case, the SNR is most commonly approximated as increasing with the square of B_0. A more complete characterization of SNR in MR imaging is available in Nishimura's textbook [62].

fact, even when compared with histology, the majority of cortical lesions—especially the Type III subpial lesions—are still being missed.

Luckily, ultra-high field magnets have effects on parameters other than just SNR. Kangarlu and colleagues, in their 8.0 T study [126], suggested that effects other than SNR enabled improved lesion recognition—effects such as the high contrast-to-noise ratio created as a consequence of the changes in relaxation parameters at high field strengths (e.g., longer T_1 and shorter T_2 and T_2* relaxation times [132-134]). In particular, data at ultra-high field strengths also demonstrated greatly enhanced T_2* contrast resulting from the increased sensitivity to magnetic susceptibility at higher field strengths [134, 135]. Moreover, in a 2007 study by Mainero and colleagues that looked at T_2*-weighted (T2*w), T2w, T1w, and phase images at 7.0T, the authors found that the T2*w magnitude images had the highest sensitivity to cortical pathology in MS [127]. What sets this study apart from many others that claim high sensitivity to cortical pathology is that the distribution of the cortical lesions Mainero et al. identified in vivo matched that of previous postmortem studies. That is, of the 199 cortical lesions they identified, Type III subpial lesions were the most frequent (50%), followed by Type I (36%) and Type II (14%) lesions. The utility of T2*w imaging has yet to be fully explored, especially at clinically practical field strengths, but a recent postmortem validation study found that, with a 7.0 Tesla T2*w sequence, 93% of all cortical lesions could be identified retrospectively and that the relatively low prospective identification rate of 40% could be attributed simply to lesion size [136]. Overall, T2*w imaging shows great potential for cortical lesion identification in MS and is the focus of some of the most recently published studies on the subject, even if only at ultra-high field strengths [128, 137, 138].

Finally, a 9.4 T postmortem study by Schmierer and colleagues investigated the association between histological features (i.e., myelin content, neuronal density, and axonal preservation) and quantitative MR measures (i.e., T_1 and T_2 relaxation times and MTR) in cGM lesions [52]. The data showed that, compared with nonlesional GM, cGM lesions had significantly lower MTR, and higher T_2 relaxation times. Both the T_2 times

and MTR values were predictors of the histologically quantified myelin content; however, after regressing the myelin content on both MTR and T_2 relaxation time, the partial *r*-values indicated that the association between myelin and T_2 was stronger.[xxxi]

Extralesional Pathology

Though not as easily identifiable as lesional pathology, extralesional abnormalities can be observed on MR images. Normal-appearing white matter (NAWM) is defined pathologically as WM that appears macroscopically normally myelinated and is at least 1 cm away from the edge of a plaque [55]. Normal-appearing WM is perhaps the easiest nonlesional tissue to investigate with MR images, if only because WM lesions are so conspicuous on conventional MR imaging sequences. Because the microscopic changes in NAWM are not as severe as those in their lesional counterparts, it follows then that the change in signal from MR images of NAWM is also not as readily visible. Still, NAWM abnormalities detectable on MR images are present in all MS subtypes, and though they appear widespread, they may be more apparent adjacent to lesions [139-142]. Normal-appearing WM abnormalities are usually detected using nonconventional sequences. Of these, MTR has been suggested as the most sensitive and reproducible measure for investigating changes in NAWM [142].

The term normal-appearing grey matter (NAGM) has, until recently, been used almost synonymously with GM because lesional GM pathology has been largely undetectable by MR imaging, especially at the more common field strengths of 1.5 T and 3.0 T. Still, investigations of intensity deviations in NAGM on MR images have been reported, with MTR being the predominantly used modality for detecting such abnormalities [139, 142-147]. Volumetric GM studies, that is, studies wherein the volume of the GM is quantified using various image processing methods and is the subject of investigation (rather than the MRI intensity value), have become highly valuable tools for studying so-called NAGM pathology in MS and will be the subject of further discussion in the next section.

[xxxi] The described 9.4 T study used formalin-fixed brain tissue. In a previous postmortem study by the same authors but at 1.5 T [97], they also found that T_2 was a better predictor for myelin content in fixed-brain WM; however, it should be noted that, in the same study, they also showed MR was the best predictor in unfixed-brain WM.

2.3 IMAGE PROCESSING IN MS

Magnetic resonance imaging has become an integral part in the diagnosis of MS as well as an invaluable tool to the neurologist for assessing treatment and prognosis. Neurologists have come to appreciate its sensitivity to pathology and, as technology advances, its increasing specificity. For the qualitative assessment of the brains of patients with MS, there is no better choice than the insight, reasoning, and intelligence of the educated mind and trained eye of a neurologist/neuroradiologist, armed with MR imaging. Yet, even the most specialized and skilled physician can only do so much. Bound by the physical limitations of their senses, even they have turned for aid to computers. Digital image processing, a subfield of signal processing, employs computer algorithms on digital images, or in the case of 3D MR imaging, digital volumes, to yield metrics that are sensitive and relatively specific to the presence of MS pathology and its temporal evolution. For the in vivo quantitative assessment of the brains of patients with MS, it knows no rival.

2.3.1 PREPROCESSING

Each of the different types of MS pathology described above (section 2.1.2) can be identified via various methods. Perhaps the most rudimentary is a human's evaluation of an MR image by simple visual assessment. Additionally, specialists can attempt to quantify the size, location, number, and/or degree of severity of MS lesions to the best of their ability. This manual identification technique, termed manual segmentation, has a number of drawbacks, but perhaps the most obvious is the fact that it is manual and not automated. Automated techniques have the advantage of being faster—and in many instances more precise then their manual counterparts. What they lack, however, is the superior reasoning ability of the human mind. Thus, a number of steps are taken in order to ensure high-quality standardized input into these automated methods so that the complexity of human intelligence can be reduced to a simpler set of programmed rules.

The first step for nearly all automated methods is some form of preprocessing. In the context of MS lesion segmentation, preprocessing usually involves optimizing the original MR image(s) that are output from the scanner and providing standardized input image(s) that will be fed into a classifier or pipeline (i.e., the actual program(s) that will identify the tissue, or structure, of interest). Preprocessing can include anything from various filtering techniques such as those that attempt to reduce the random noise arising from the acquisition process (termed denoising [148-150]), to removing artifacts such as intensity nonuniformities [151, 152], to removing the skull and other nonparenchymal tissues of the brain from the image [153-157], to the alignment of two or more images into spatial correspondence, termed registration.[xxxii]

Intensity nonuniformities (Figure 2.16)—also referred to as bias fields, inhomogeneities, gain fields, or illumination nonuniformities—are usually attributed to gradient-driven eddy currents, poor RF coil uniformity, and patient anatomy. While these 10%-20% intensity variations have little impact on visual diagnosis, they seriously degrade the performance of automatic segmentation techniques that rely on the intensity being homogeneous within a particular class [151]. The Nonparametric, Nonuniform intensity Normalization procedure (termed N3) is fully automatic, requires no a priori knowledge, and can be applied to almost any MR image. The preprocessing of data using N3 has been shown to substantially improve the accuracy of anatomical analysis techniques such as tissue classification, registration, and cortical surface extraction [158-160]. For these reasons, and of course for its high performance, N3 has been the industry standard for nearly 15 years, though recent improvements have led to a successor, N4ITK, which is slowly being adopted [152]. Nonetheless, virtually all image processing methods rely on some form of intensity nonuniformity correction as part of their preprocessing protocol. A nice review of the many different nonuniformity correction algorithms is available in [161].

[xxxii] The publications referenced above [148-157] are by no means a comprehensive list and represent but a tiny fraction of the methods available. Registration is an image processing subfield unto itself, with over 9,000 publications currently listed on PubMed, and will be a topic of later discussion.

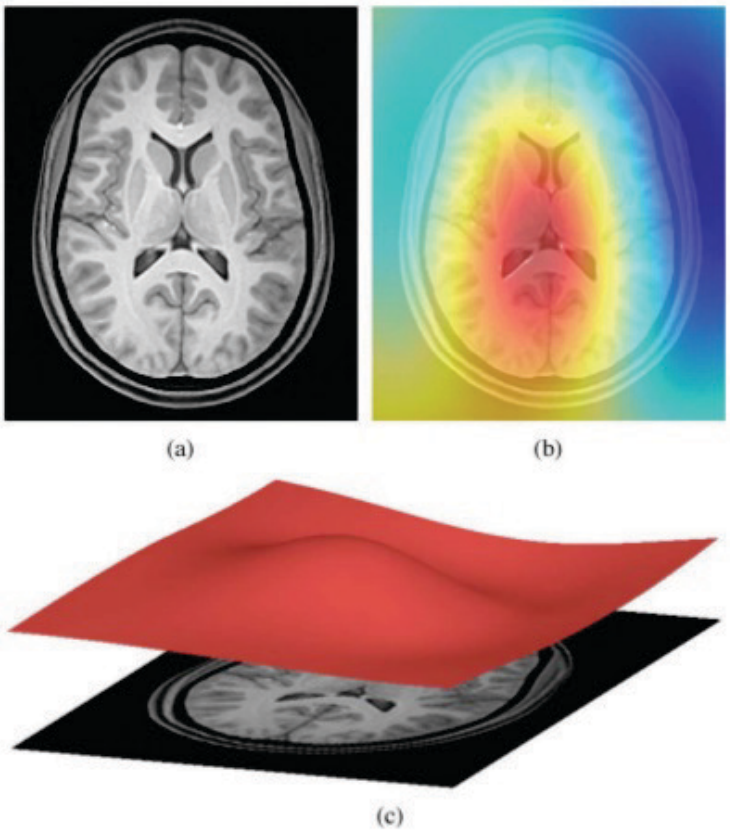

Figure 2.16 – Intensity nonuniformities. (a) T_1-weighted MR image exhibiting bias field. Notice the lowered intensity on the right side of the image. (b) Several algorithms have been proposed to estimate the bias field, which can then be used to "correct" the image. (c) Viewed as a surface, the low frequency modulation of the bias field is readily apparent. Figure and caption reproduced from [152].

2.3.2 REGISTRATION

Though there is a seemingly innumerable[xxxiii] amount of registration methods available, generally, all the procedures developed either implicitly or explicitly involve the following steps: (i) identification of features; (ii) a measure of similarity based on these features; (iii) a method to identify the parameters of the spatial transformation function; and (iv) a definition of the nature and domain of the spatial transformation function.

[xxxiii] An attempt to formally classify the many variants of registration can be seen in Maintz and Viergever's 1998 survey of medical image registration [162] or in Van den Elsen et al. [163]. For a review of the more current algorithms being used today, the reader is referred to recent reviews by Oliveira and Tavarez [164], and Glocker et al. [165].

Identification of features

The term feature is used in this context as a deliberately vague term. In terms of MR images, a feature could be anything from the intensity at a given location to specific anatomical landmarks. Features can be extrinsic, that is, based on foreign objects that are either invasively or noninvasively introduced into the image space. Invasive markers include fiducial screws, or a stereotactic frame that is often used in surgical planning. Fiducial skin markers, a head mould, or a dental frame, are examples of noninvasive extrinsic markers. Much more popular are the intrinsic markers, that is, information already in the image that is generated by the patient. These can range from geometrical (e.g., points, lines, and surfaces) or anatomical landmarks to segmented structures to the image properties of the voxels[xxxiv] themselves. The latter is perhaps the most popular in automatic registration-based methods and can in turn range from the simple grey scale value of the voxel to the intensity of that voxel after a morphological operation such as blurring, or gradient calculation. In the case of automated registration algorithms, any feature used needs to be identified automatically.

A measure of similarity based on these features

The measure of similarity is used in the context of an objective function, where the similarity of one or more features is used to evaluate the match between the two volumes being brought into spatial correspondence. When trying to maximize a function, as is the case with similarity, the objective function may be more specifically referred to as a utility function. By contrast, if one seeks to minimize the objective function, it could be called a cost function. An objective function can also comprise a combination of utility and/or cost functions and are as custom as the minds of the users who create them.

The goal of registration algorithms is to find an appropriate mapping from a source image to target image. By incorporating this mapping as a term in the objective function, algorithms are able to iteratively modify the mapping and then evaluate if the current

[xxxiv] A pixel is the unit element of a 2D image, whereas a voxel is simply its 3D counterpart.

iteration of said mapping is better or worse than its predecessor, eventually leading to the convergence of an appropriate mapping between source and target, and thus, a good registration.

Choosing the objective function that will yield the best registration is a difficult task as there are many objective functions available. For a comprehensive review, the reader is referred to book chapters by Hill [166, 167], and Woods [168], as well as these review publications [163-165]. Briefly, different objective functions have traditionally been employed depending on the context of the registration. In the case of intramodal registration (i.e., between two images with the same intensity contrasts and range), the sum of squared distances, sum of absolute differences, least trimmed squares, variance of intensity ratios [169], and cross-correlation [168] objective functions have been used, with cross-correlation being amongst the most widely used.

When the images do not have the same intensity contrasts and range (i.e., inter-modal registration), the predominantly used objective function is that of mutual information. Based on the Shannon-Weiner entropy measure from communication theory developed c.1940 [170], mutual information evaluates histogram probabilities in order to reduce the amount of information in a combined, or overlapping, image of the two images being registered.

A method to identify the parameters of the spatial transformation function

Once the appropriate objective function has been selected, a method for solving that objective function (i.e., finding the parameters that yield the maximum of a similarity function) must then also be determined. In the simplest case, direct solutions can be computed (e.g., by singular value decomposition), but in most cases, a more sophisticated optimization procedure is necessary (e.g., gradient descent, genetic algorithm, Levenburg-Marquardt, etc.). Regardless of the method, the result of the optimization is the identification of the parameters associated with the (best) spatial

transformation function. The number of parameters depends on the nature and domain of that function.

The nature and domain of the spatial transformation function

The spatial transformation function (i.e., the mapping of source to target) can be defined according to both its nature and domain. A comprehensive review can be found in [162] or [163], among others.

Briefly, there are two possible domains for spatial transformations. The global domain is where the transformation is applied globally; that is, the same mapping is applied to each pixel in the image or to each voxel in the volume. This is also referred to as linear registration. The second possibility is a transformation operating in the local domain, which means that there exists a separate mapping for each subsection of the whole (e.g., each pixel in an image or each voxel in a volume). This is also referred to as nonlinear registration. The mapping need not be unique, and in the case in which they are all identical, we have the encompassed special case of a global transformation.

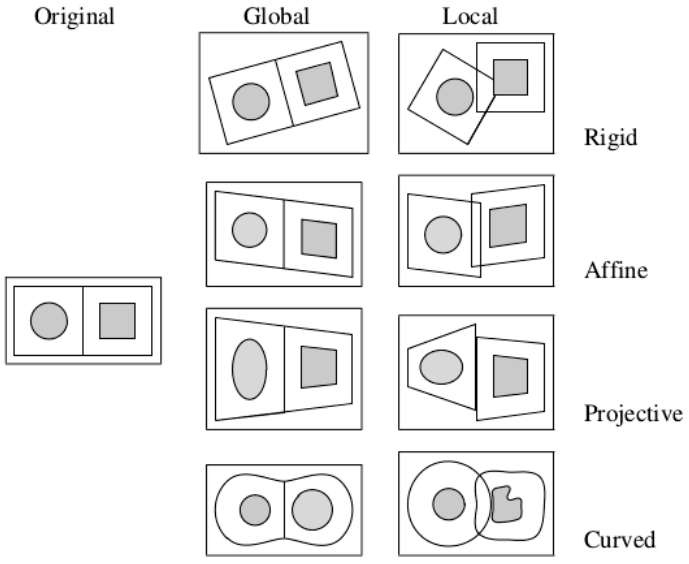

Figure 2.17 – Two-dimensional examples of the nature and domain of spatial transformation functions. Each type of transformation contains as special cases the ones described above it (e.g., the rigid-body is a special kind of affine transformation). Also, a composition of more than one transformation can be categorized as a single transformation of the most complex type (e.g., the composition of an affine and a projective transformation is a projective transformation). Reproduced from [162].

The nature of the mapping concerns the degrees of freedom used and can be any of the rigid-body, affine, projective, or curved types, though in practice, a rigid-body (i.e., 6 parameters) or affine (i.e., up to 12 parameters) transformation is all that is typically used for the linear registration of MR images. In the case of nonlinear registration, the curved type, most often represented in terms of a unique displacement vector at each voxel, is the norm. Whereas a linear registration is appropriate for intrasubject registration, nonlinear registration is better suited for intersubject or subject-atlas registration.

To summarize registration then, an optimal mapping can be found by optimizing an objective function based on certain features, and that mapping can be applied according to a defined nature and domain in order to align one image with another.

2.3.3 White Matter Lesion Detection

The manual segmentation of WM lesions on MR images is cumbersome, time-consuming, objective, and prone to human error. Many groups have developed automated tissue classification[xxxv] and lesion segmentation methods in order to improve upon these drawbacks. A recent review of the various methods for WM lesion segmentation can be found in [171]; in general, they fall into one of the following four categories:[xxxvi]

1. Data-driven methods: Thresholding, region growing, and/or other spatial approaches. These methods were some of the first developed and are perhaps the most intuitive. For example, thresholding in its most basic form involves defining the range of intensities that are to be considered lesion. For the case of a T2w hyperintense lesion on an MR image, this entails finding the correct intensity value such that all voxels with intensities below this value represent

[xxxv] Classification broadly refers to the practice of labelling each voxel in a volume to one of several predefined classes, whereas segmentation is the delineation of a specific structure, though segmentation can be achieved as a result of tissue classification, and vice versa, if all structures in the volume were segmented such that every voxel belonged to some structure.

[xxxvi] Segmentation algorithms need not be restricted to using just one of the above four methods, and in practice, this is rarely the case. Those that use a combination of methods can be classified according to the category that has the most relevance to the core method of the algorithm.

nonlesional tissue, while those with intensities above it represent the lesions. Though not overly complex, these have proven, unfortunately, not to be the most accurate methods available and are now more commonly used in conjunction with other techniques. Notable examples of data-driven lesion segmentation software include an intensity-based global thresholding scheme [172], histogram matching algorithms [173, 174], adaptive thresholding [175], region growing [176], and hierarchical techniques [177].

2. Statistical methods: Can be further divided into nonparametric and parametric techniques, depending on the normality (i.e., if the distribution resembles a Gaussian distribution) of the underlying data. These represent the most popular methods described in the literature as well as the more recent and include techniques such as k-nearest neighbours (kNN) [178-180], support vector machines [181], Bayesian classifiers [182-184], expectation-maximization algorithms, and combinations thereof with other statistical modelling techniques (e.g., Markov random fields (MRF), Gaussian mixture models, etc.) [185-187].

3. Intelligent methods: Cover a wide range of neural networks and/or fuzzy logic. These methods are also popular in the literature and are considered to have a high degree of complexity. Notable examples include the artificial neural network employed by Zidjdenbos et al. [188], the fuzzy connectedness method proposed by Udupa et al. [189], and the fuzzy segmentation method of Shiee et al. [190].

4. Deformable models: Methods concerned with volume estimation, usually via some sort of deformation. These represent the least popular in the literature, and are still lagging behind the performance of statistical methods. Examples include [191-193].

Table 2.2 – Comparison of WM lesion segmentation methods. ICM = Iterated Conditional Modes, kNN = k-Nearest Neighbours, EM = Expectation Maximization, MRF = Markov Random Fields, AMM = Adaptive Mixture Model, SVM = Support Vector Machine, ANN = Artificial Neural Network, FCM = Fuzzy c-Means, FIS = Fuzzy inference Systems. For a full description of each method see [171].

	Median Pub. Date	Overall popularity (%)	Method	Popularity (%)	Accuracy	Complexity
Data-driven	1998	10	Thresholding	6	Low	Low
			Region growing	2	Low	Low
			Hierarchical	2	Medium	Medium
Statistical	2006	48	ICM	2	Medium	Medium
			kNN	18	Medium	Medium
			EM	20	Medium	Medium
			kNN+EM+MRF	4	High	High
			AMM	2	Medium	Medium
			SVM	2	High	Medium
Intelligent	2002	38	ANN	12	Medium	High
			FCM	18	Medium	Medium
			Fuzzy connectedness	6	Medium	Medium
			FIS	2	High	Very high
Deformable	2004	4	Deformable contours	4	Medium	Medium

Regardless of the methodology used, WM lesion detection in MS has come to play a central role in the development and evaluation of new therapies. Despite the clinico-radiological paradox seen at the patient level, WM lesion activity on MR images has become the accepted surrogate[xxxvii] primary outcome measure in proof-of-concept placebo-controlled clinical trials of new immunomodulating therapies in (RR) MS. The reasoning behind this is twofold: (i) that MR imaging markers and clinical outcomes are generally poorly correlated reflects variability of both measures rather than a lack of a biologically meaningful relationship [194, 195] and (ii) that MR imaging markers being weak at the patient level does not preclude the use of MR imaging measures in clinical trials. The importance of this distinction between individual-level and trial-level surrogacy is discussed by Korn et al. [196], and Sormani et al. [197, 198].

[xxxvii] The Prentice criteria for formal surrogacy validation [Prentice 1989] stipulate that a given treatment must be effective on both the surrogate and the clinical endpoint, the surrogate and clinical endpoints must be significantly correlated, and the effect of treatment on the clinical endpoint must be mediated through the effect on the surrogate.

2.3.4 Cortical Grey Matter Lesion Segmentation

All of the WM lesion segmentation methods described above, regardless of their category, relied on the fundamental principle that there exists a detectable difference between lesions and the surrounding tissue. In the case of cGM lesions, in particular subpial demyelinating lesions, these differences in contrast simply do not exist. An exception is seen with DIR, where Type I and Type II cGM lesions appear hyperintense compared with the surrounding GM. For such lesions, it is conceivable that many of the automated methods for T_2 hyperintense lesion segmentation would perform well with some minor modifications. Indeed, such automated techniques would be useful to help differentiate lesions from artifact. At this time, however, such endeavours might be considered premature since DIR misses the most abundant cGM lesion type (i.e., the Type III subpial demyelinating lesion) and the first consensus recommendations for what constitutes a lesion on DIR have only just been established [199].

Given that the imaging community is still in the process of defining cGM lesions on MR images, not to mention still struggling to identify them via manual segmentation by a trained expert, it comes as no surprise that there are presently no automated techniques that accurately and reliably segment cGM lesions. The need for such techniques is immense, as researchers seek to understand the role of cGM lesions in MS, and serves as the motivation for the work of the automated method presented in Chapter 4.

2.3.5 Cortical Grey Matter Segmentation

Cortical lesions aside, the automatic segmentation of the cerebral cortex is a challenge unto itself. The cortex is a thin layer of GM surrounded by WM on one side and CSF on the other, both of which produce partial volume effects that confound its delineation. Still, many of the above-mentioned WM lesion segmentation methods include a tissue classification stage whereby the cerebral cortex is segmented.[xxxviii] Achieving an accurate segmentation is important because the cerebral cortex serves as a marker for

[xxxviii] What most of the algorithms do not offer, however, is a separate segmentation for both the cGM and the subcortical GM structures (e.g., caudate, putamen, globus pallidus, thalamus, etc.); that is, they provide one all-encompassing GM segmentation, sometimes even including an estimate of the cerebellar GM.

neurodegeneration in MS; an analysis of the accuracy of several of the most commonly used volume-based segmentation methods is presented in Chapter 3. Volume-based methods refer to those that operate in the 3D domain (i.e., segmenting on a voxel basis; using tissue classifiers) and are in stark contrast (i.e., use different assumptions and methods) to cortical thickness approaches that attempt to quantify the GM volume at a subvoxel resolution via a 2D cortical surface extraction.

Volume-based methods

In a recent survey of studies that quantified GM volumes in MS, the two most prominent voxel-based methods used were that of SPM and SIENAx. Thus, they, along with FreeSurfer (which has been gaining popularity in the literature)[xxxix], will be the subject of further review.

SIENAx

Created at the Oxford Centre for Functional MR Imaging of the Brain (FMRIB) and distributed as part of the publicly available and widely used FMRIB Software Library (FSL; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/) [200, 201], the Structural Image Evaluation, using Normalization[xl], of Atrophy cross-sectional method (SIENAx) is used to provide brain tissue volumes [202]. The heart of the technique is FMRIB's Automated Segmentation Tool (FAST), which segments the brain into GM, WM, and CSF while also correcting for spatial intensity variations and partial volume effects. The method employs a straightforward k-means segmentation in order to first estimate the initial parameters of the three tissue classes and then uses a hidden MRF model[xli] and an associated expectation-maximization algorithm to calculate the parameters of that model for the final classification [203]. Although it does not provide WM lesion segmentation,

[xxxix] The unpublished 2010 survey of 25 studies was conducted by the author while doing the literature review and researching the motivation behind Chapter 3.
[xl] Brain tissue volumes are normalized for subject head size.
[xli] The finite mixture model (e.g., assuming separate Gaussian distributions for the intensities of each tissue class in an image) is perhaps the most commonly used model for statistical segmentation because of its simple mathematical form; however, as a histogram-based model, it does not take into account any spatial information. Markov random field models, on the other hand, incorporate spatial information by placing contextual constraints on the segmentation of a voxel based on the characteristics of its neighbours [203].

the taxonomy of section 2.3.3 can be used to categorize SIENAx as a statistical segmentation method.

Until recently, SIENAx offered a segmentation only of whole-brain GM. To a certain extent, it was the work presented in Chapter 3 that prompted the separate reporting of cGM volume.

SPM

The Statistical Parametric Mapping software suite (SPM) is a collection of MATLAB functions and subroutines designed for the analysis of brain imaging data. Originally developed by Karl Friston for the routine statistical analysis of functional neuroimaging data [204, 205], the software has grown considerably to include several analysis pipelines, is publicly available (http://www.fil.ion.ucl.ac.uk/spm), and has been heavily adopted by the imaging community. One of the reasons for the popularity of SPM is the development of the voxel-based morphometry method by Ashburner and Friston [206]. Their method has been widely used for the quantification of GM atrophy, and will be covered in section 2.3.6, but fundamental to that method is a voxel-based tissue classification where the GM is segmented. As in FAST, there is no separate class for cGM.

The current version of SPM (SPM8) uses the same tissue classifier that was developed and released in SPM5 and is the unified segmentation algorithm also developed by Ashburner and Friston [207]. The algorithm uses a Gaussian finite mixture model (i.e., the probability of the intensity distribution of a tissue class is modelled by a Gaussian), where the parameters of the model are estimated by minimizing a custom cost function. This model is then extended to incorporate a correction for intensity nonuniformity as well as spatial information based on other subjects' brain images (i.e., a nonlinear registration to a tissue probabilistic atlas) to provide a final, unified tissue classification.

FreeSurfer

FreeSurfer currently comprises a suite of tools that are publicly available (http://surfer.nmr.mgh.harvard.edu) and are maintained under the supervision of Bruce Fischl at the Massachusetts General Hospital and Anders Dale at the University of California, San Diego. FreeSurfer contains two main workflow streams, the first being a whole-brain segmentation, or volume-based stream, that classifies the brain into over 50 anatomical structures (for reference, most methods similar to those described above commonly use only three classes). The classification is based on the intensity information as well as several heuristic neighbour constraints (e.g., the hippocampus should not be adjacent to the caudate; the amygdala should reside anterior and superior to the hippocampus, etc.). A detailed description of the method, along with an analysis of the performance, is given in [208], but briefly, prior intensity and spatial arrangement probabilities for each anatomical structure are obtained from the training set that comprises the atlas used in the method. A high-dimensional nonlinear (i.e., curved in nature and local in domain) registration to the atlas is performed to ensure proper alignment. The probabilities are used with an MRF approach, which simply means that the probability of a class at a given voxel is computed not just in terms of the intensities and prior probabilities at that voxel, but also as a function of the class in a neighbourhood around the voxel in question. Importantly, a separate histogram is modeled for each GM structure within the brain; therefore, in contrast to SIENAx and SPM, FreeSurfer does provide a unique segmentation for the cortical and subcortical structures of the brain.

The volume-based stream is mostly recommended for subcortical segmentation, but can be used to provide a satisfactory segmentation of the cerebral cortex; however, for a more accurate and precise segmentation of the cortex, the surface-based stream within FreeSurfer is recommended.

Surface-Based Methods

Several automated surface-based methods exist for cortical extraction, but common to all of them is the extraction of both an inner surface—often called the WM surface—which represents the boundary between WM and GM, and an outer surface—often termed the pial surface—which represents the outer limit of the GM, or the boundary between GM and CSF.

It should be noted that each of the methods described below uses its own preprocessing steps (i.e., registration, skull-stripping, cutting the brain into two separate hemispheres, tissue classifiers, etc.), but that the focus presented here will be on the actual surface extraction algorithm used.

CLASP

The Constrained Laplacian-based Anatomic Segmentation using Proximity algorithm (CLASP) [209] is a modification of the original cortical surface extraction method developed at the Montreal Neurological Institute [210].[xlii] CLASP first extracts the inner surface by deforming a sphere polygon model to the boundary between GM and WM. This deformation is done in a hierarchical manner, increasing the number of polygons at each iteration, while explicitly preventing self-intersections, thus preserving the topology of the sphere for the modelled surface.[xliii] The WM/GM boundary is determined using the information from a discrete tissue classifier (INSECT) [188]. This WM surface is then expanded to the boundary between GM and CSF along a Laplacian map to create the pial surface. At most places, the boundary between GM and CSF is also defined using the discrete tissue classification [188], but a probabilistic partial volume classification is used to detect buried CSF voxels in the depths of sulci. The

[xlii] The original method developed by MacDonald et al., Anatomical Segmentation using Proximities (ASP), deformed the GM and WM surfaces simultaneously using the restriction that the cortical thickness should fall within a predefined set of values, or that the surfaces have a somewhat fixed proximity [210].

[xliii] Geometrically, the human cerebral cortex is a thin folded sheet of GM; so strictly speaking, the cortex of each hemisphere is topologically a sheet. But, by artificially capping, or closing, the opening at the brainstem, the surface of the cortex becomes topologically equivalent to a sphere [211, 212]. Thus, to be accurate representations of the cortex, models should be isomorphic to either a sheet or a sphere.

reasoning for the partial volume classification is that pure CSF voxels are considered rare in folded sulci because sulcal walls are often so close together [209].

FACE

The Fast Accurate Cortex Extraction method (FACE) was created by Eskildsen and Ostergaard [213-215]. FACE differs from CLASP in terms of creation of the WM surface; that is, instead of deforming a polygonal sphere, it first tessellates the cerebral WM of each hemisphere. At this point, holes, handles, tunnels, and/or other topological defects may exist; thus, to ensure that the tessellated surface is isomorphic to a sphere, a topology correction algorithm is applied [211]. This initial estimate of the inner surface is then deformed by moving vertices so as to minimize an energy function. In this case, the function is expressed in terms of both internal energies, which control the behaviour of the deformable surface (e.g., the tension of the surface), and external energies, which are used to guide the surface toward the WM/GM boundary (e.g., gradient information, distance from the initial surface). In addition, two hard constraints are applied: one that prevents the surface from self-intersection and another that ensures a certain minimum distance between neighbouring vertices, achieved by remeshing (i.e., adding or deleting vertices where necessary). This final WM surface is then deformed to create the outer pial surface, a process which, much like the inner surface deformation, tries to minimize its own energy function. The deformation is weighted to be along the surface normal, which is generally in the direction of the pial surface; however, a specific energy term called the generalized gradient vector flow force (GGVF) [216] is also used to help point the surface towards the nearest image boundary.

FreeSurfer

The other workflow stream available within FreeSurfer is the cortical surface extraction, or surface-based stream. The segmentation and surface reconstruction work was originally proposed in 1999 by Dale et al. [217]. The surface extraction is similar to the FACE method described above in that the WM surface is created from an initial

tessellation of the cerebral WM output from a tissue classifier and then refined to follow the intensity gradients between WM and GM. The surface deformation for both the inner and outer surfaces is accomplished by minimizing a custom energy function, which is a weighted sum of three terms:

i. A tangential spring term (J_T), which acts to redistribute vertices to regions where they are needed, encouraging a uniform spacing of vertices;

ii. A second spring term in the orthogonal direction, termed the normal spring term (J_N), which imposes a smoothness constraint on the surface, with each vertex being penalized for being too far in the direction normal to the surface from its neighbours; and

iii. An intensity-based term (J_I), which is computed differently for the WM/GM interface than for the pial surface. For the WM surface, the target intensity value is determined using the local mean value of the voxels in a 5 mm neighbourhood. For the pial surface, a simple global target based on the contrast between CSF and GM is used.

Although there are many cortical surface extraction algorithms, some of which are publicly available, FreeSurfer has grown to become the most widely used. This is in part due to its high level of documentation and support, but also because it has been vetted by the community and found to be quite robust (i.e., shows good test-retest reliability across scanner manufacturers and field strengths [218, 219] and performs well for pathological, pediatric, and nonhuman brain cases; see the FreeSurfer website for a considerable list of references). FreeSurfer's cortical extraction stream has also been validated by comparing the extracted cortical thickness values with those from manual measurements [220, 221] as well as postmortem histology [222].

Since cortical surface extraction has the potential to provide more than just a measure of cortical thickness, and because it is an integral part of the methodology presented in Chapter 4, further aspects beyond just the extraction itself warrant discussion.

Surface-Based Registration

Perhaps one of the reasons FreeSurfer is so popular, and what also sets it apart from the other methods described herein, is that, in addition to the cortical surface extraction, it provides a high-dimensional (i.e., local in domain and curved in nature) intersubject surface-based registration [223, 224].[xliv] Much like the surface deformations, the surface-based registration is achieved by minimizing a custom energy function comprising a weighted sum of the following terms [224]:

i. An alignment term (J_P), which helps to match the (primary) folding patterns of the two surfaces by using the average convexity information at each vertex;

ii. An area term (J_A), which minimizes distortion by penalizing vertex movement that results in a smaller area (i.e., three vertices will make up a triangular face) than the original area; and

iii. A distance term (J_D), also used to minimize distortion, which ensures that the difference between a vertex and its neighbours at the current iteration is not drastically different from that same distance computed on the original surface (i.e., iteration zero).

The advantages of using a surface-based registration that explicitly aligns the major cortical folding patterns over even a high-dimensional volume-based registration (Figure 2.18 and Figure 2.19) stem from the fact that the curvature patterns (i.e., the gyral and sulcal foldings) to be aligned are a property of the 2D cortical mantle and therefore

[xliv] This surface-based registration allows for good intersubject alignment—a necessity for any group-based study. In addition, a single subject's surface can be aligned to an atlas, as is the case in [225] where a detailed parcellation of the cortex, based on gyral and sulcal structure and defined in [226], is obtained.

can only be determined from an explicit representation of the surface itself. A full discussion is offered in [224].

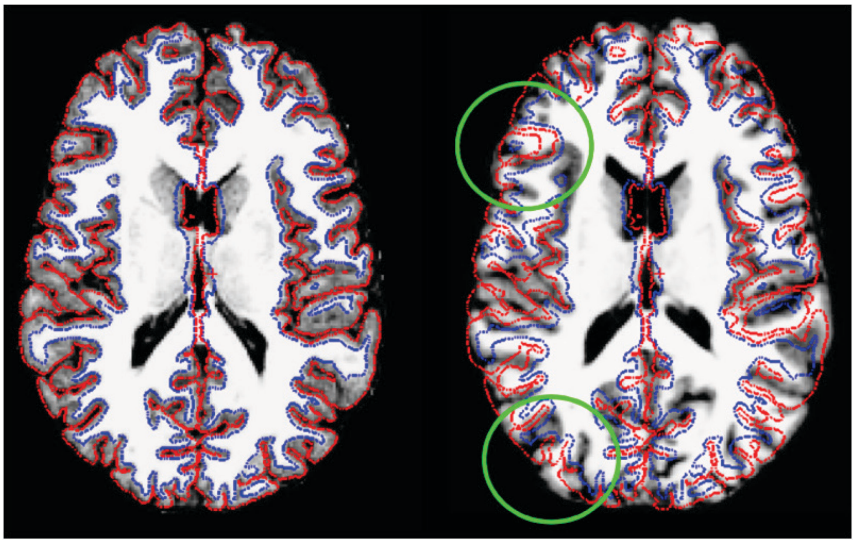

Figure 2.18 – Volume-based (3D) registration with 12 degrees of freedom. Notice how the cortical folding patterns are not aligned (green circles) [227].

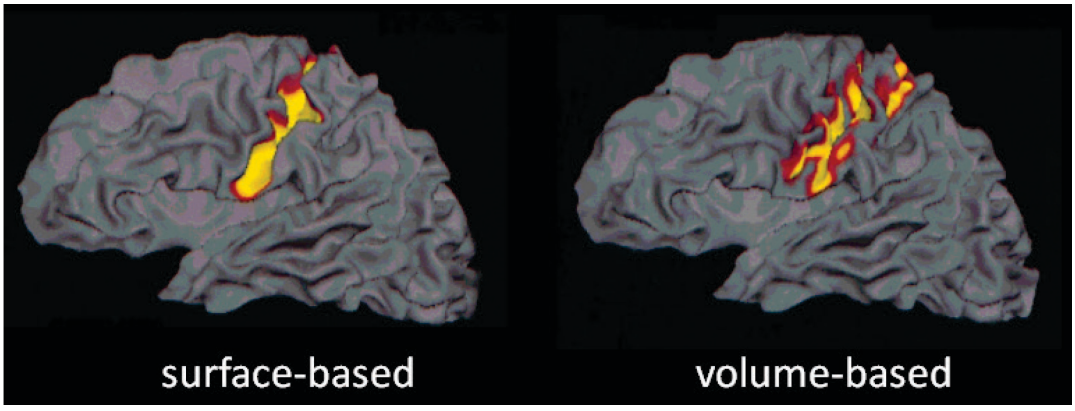

Figure 2.19 – Surface- versus volume-based registration. The mapping of the central sulcus of 11 subjects onto an individual WM surface using both registration methods. Note the superior localization of the sulcus using surface-based registration. Modified from [224].

Cortical Thickness Metrics

Once both the inner and outer surfaces have been extracted with the method of choice, the cortical thickness can be measured using various distance metrics between each of the surfaces. A full evaluation of several methods is available in [228], but they are briefly described below and illustrated in Figure 2.20:

a) Linked distance: The cortical thickness is measured as the distance between the linked vertices. That is, vertex *n* on the WM surface is inherently linked to vertex *n* on the pial surface in the cases where the pial surface is an expansion of the WM surface. The magnitude of the vector between these two vertices is taken as the cortical thickness. This method requires that the two surfaces be topologically equivalent with the same number of vertices.

b) Nearest distance: Distance to the nearest vertex on the adjacent surface. This method performs a simple search across the opposite surface and picks the vertex that is the shortest Euclidian distance away, which, it should be noted, is not necessarily the same as the linked vertex. This method has the potential for some gross errors, such as jumping across gyri. A modification of this method is to compute the nearest distance twice, once from the WM to the pial surface and again from the pial surface to the WM surface, and to then report the average of the two distances.

c) Normal distance: The distance along the surface normal (i.e., perpendicular to the surface) from a vertex to the opposite surface. Variations, such as the averaged method described above or ones that first create intermediate surface(s), are also possible and are in fact preferable since simply taking the normal distance is prone to large errors when the inner and outer surfaces are not parallel (see Figure 2.20 for an example).

d) Laplacian distance: The sum of the distance between all nonintersecting intermediate equipotential[xlv] surfaces that lie between the two boundary surfaces. Laplace's equation is a second-order partial differential equation that describes just such a set of layered equipotential surfaces that make a smooth transition from one boundary surface to the other. For a more thorough description, see [212].

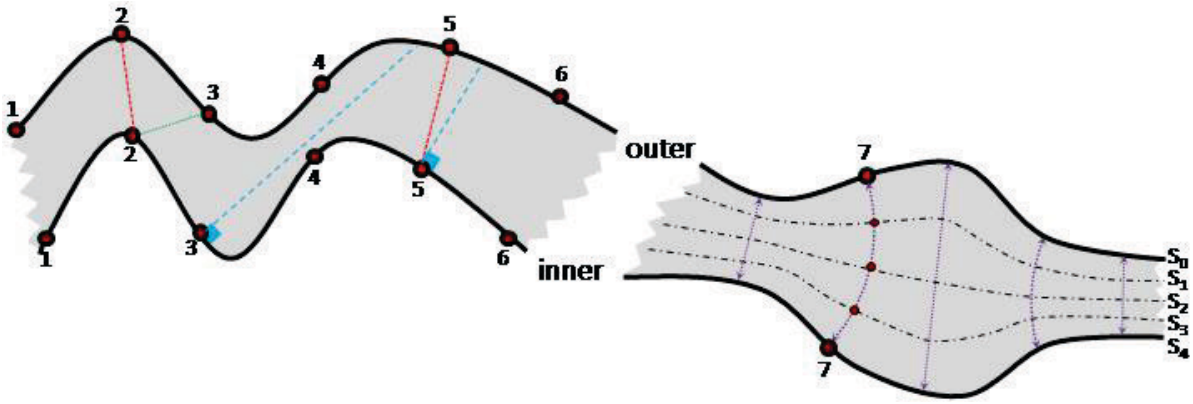

Figure 2.20 – Cortical thickness metrics. Numbers 1 to 7 represent vertices on the inner and outer surfaces. Red dashed lines connecting same-numbered vertices represent linked distances. The green dotted line from vertex 2 to 3 represents the nearest distance. The blue dashed lines represent the normal distances from the inner surface to the outer surface. Note how with this method, at vertex 3, the thickness would be erroneously reported. The Laplacian distance is represented at vertex 7 and is the sum of the distances from point to point along the purple dotted line. The equipotential surfaces are represented by the inner and outer surfaces as well as the dashed black lines (S_0 to S_4). The purple dotted lines are termed field- or streamlines and connect the inner and outer surfaces while being everywhere orthogonal to all equipotential surfaces.

Surface-Based Blurring

In addition to the variety of cortical thickness metrics, different methods are also available for blurring surface data, which is most often the cortical thickness values. Blurring is generally done in order (i) to reduce the noise in the measurement, thus increasing the SNR; (ii) to increase the validity of statistical tests by rendering the data more normally distributed and reducing the number of comparisons that need be

[xlv] The term equipotential is borrowed from the mathematical description of the applications of Laplace's equation to electrostatic fields. In such a system, the two bounding surfaces are given two arbitrary voltages (potentials), and a significant property of Laplace's equation is that nonintersecting intermediate surfaces are guaranteed to occur and these surfaces have a constant voltage value; that is, they are equipotential. Though the description is of a vastly different physical system from that of surface extraction, the systems share the same properties of Laplace's equation [212].

statistically controlled for (i.e., blurring increases the interdependence of the neighbouring vertices); and (iii) to reduce the impact of imperfect anatomical alignment when comparing the same vertices between cortices.

Of course, the SNR and statistical benefits come at the cost of reduced image resolution; thus, it is necessary to choose an optimally sized blurring kernel, which is in itself an active area of study. An in-depth analysis of blurring kernel choice for cortical thickness analysis is presented in [228], but matched filter theory suggests that, if there is prior information about the extent of the signal to be detected (e.g., the area of cortical thinning), then the size of the blurring kernel should match the size of the putative area of change [228, 229].

Perhaps the most commonly used blurring scheme in image processing is Gaussian blurring, also called smoothing, which is the result of blurring an image by a Gaussian function of the same dimension as the image. The equation of a 1D Gaussian function is:

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

where σ is the standard deviation. The 2D (or 3D) functions are simply the product of two (or three) such functions, one in each dimension. Though σ can be used to define the Gaussian, blurring functions in general are usually expressed in terms of their defined full-width-at-half-maximum height (FWHM; Figure 2.21).

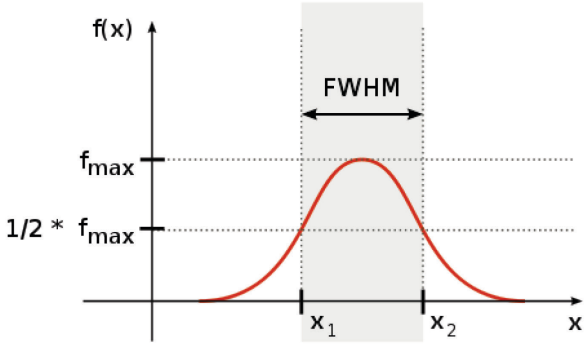

Figure 2.21 – Full-width-at-half-maximum height.

Gaussian kernel smoothing is a special case of diffusion smoothing restricted to Euclidean space. For curved non-Euclidean spaces (e.g., the curved cortical surface), the generalized diffusion smoothing is preferred.

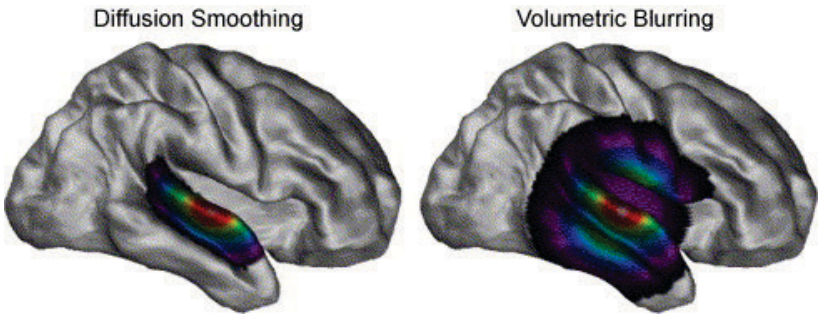

Figure 2.22 – Diffusion smoothing. An illustration of the difference between the geometry-preserving diffusion smoothing blurring over a 2D surface manifold and the more commonly employed 3D volumetric blurring kernels. The FWHM was set at 30 mm in both cases. One can see how anatomically disparate areas such as the inferior motor and sensorimotor areas are influenced by the volumetric kernel but not by diffusion smoothing. Figure and caption reproduced from [228].

2.3.6 Quantifying Grey Matter atrophy in MS

The automated segmentation of GM is a feat in and of itself, but in the context of MS research, it is often used as a tool to assess the degree of neurodegeneration. This can be done by measuring the change seen within a single subject over time or by examining the GM volume of two or more populations at any given instance. With either method, the quantification of the change in GM can be done by (i) computing the change from two independent measurements or (ii) measuring that change directly via a single measurement that incorporates the features of both time points. All of the methods described above (i.e., surface- and volume-based) fall into the category of methods that quantify the degree of atrophy by measuring the GM volume at two time points independently. It has been shown that a more precise measurement can be obtained when serial scans from an individual are accurately registered and the volume changes are derived directly using a single measurement [230]. Some examples of techniques that make use of this principle are the brain boundary shift integral [231], SIENA [232] (part of FSL), and voxel-based morphometry (VBM; part of SPM). As the latter two are widely used, they will be described in further detail.

SIENA

Structural Image Evaluation, using Normalization, of Atrophy (SIENA) starts by extracting the brains from the two images to be analyzed [156]. The two brain images are then aligned to each other [233], using the skull images to constrain the scaling parameters in the registration. Next, a volume-based classification (FAST; described above) is done in order to find brain/nonbrain edge points so that a perpendicular edge displacement (between the two time points) can be estimated at these edge points. Finally, the mean edge displacement is converted into a global estimate of percentage brain volume change between the two time points.

Because of its accuracy and high level of reproducibility (~0.1% error) [202], SIENA has been used in many clinical studies. Unfortunately, SIENA provides a measurement only of whole-brain atrophy (i.e., GM plus WM).

VBM

Voxel-based morphometry is fully described in [206], but can be thought of as a method that looks for a change in GM "density" between two images on a voxel-by-voxel basis. Briefly, the GM of the two images to be analyzed are segmented as previously described (SPM; sectionChapter 0) and are then blurred using a 3D Gaussian kernel. Voxel-wise parametric statistical tests[xlvi] that compare the smoothed GM images are then performed, with corrections for multiple comparisons using the theory of Gaussian random fields [234]. The result is a voxel-wise statistical parametric map comprising the results of the tests. While this method does provide a quantification of GM atrophy, it does not distinguish between cGM and subcortical GM, and is highly dependent on the accuracy of the GM segmentations for each image.

[xlvi] Standard parametric statistical procedures, that is, *t* tests and *F* tests, are valid for testing the hypotheses, provided that the residuals after fitting the model are both independent and normally distributed. The appropriateness of the model ensures the residuals are independent, but there are reasons the residuals may not be normally distrusted. Only by smoothing the segmented images does the behaviour of the residuals become more normally distributed [206].

Magnetic Resonance Studies of Atrophy in MS

Given the abundance of methods available for the quantification of whole-brain and GM atrophy, it comes as no surprise that there is a large body of literature on atrophy in MS. Two recent review articles on MR imaging in MS [235, 236], as well as a more focused review of GM imaging in MS [237], are offered to the interested reader.

Whole-brain atrophy in patients with MS, across all phenotypes, has been found to be on the order of 0.7% to 1% per year on average. This was quantified using a variety of methods (some of them described above), and the details are enumerated in [238]. Importantly, brain atrophy appears to be more pathologically specific than WM T_2 hyperintensities; still, it is at best only moderately correlated with clinical disability in patients with RRMS and SPMS [238, 239], with the strength of the correlation increasing when neuropsychological impairment is considered [240].

Through the use of SPM and SIENAx, GM atrophy has been shown to occur across all phenotypes and even in the early stages of the disease [241, 242], is associated with clinical disability [243-245] as well as cognitive deterioration [246, 247], has been shown to worsen over time [248], and accumulates to a greater degree over time than whole-brain and WM atrophy [235, 242, 249, 250]. Widespread cortical thinning (i.e., measured using FreeSurfer's surface-based stream) has been shown in patients with MS [251, 252], and along with volume-based GM atrophy, are significantly associated with physical disability and cognitive decline [242, 245-247, 251-254]. Importantly, the measures of GM atrophy showed stronger correlation with clinical parameters than WM damage [237, 250, 253, 255, 256].

Using regional volumetry analysis (i.e., VBM), distinct topographical distributions of GM atrophy have been shown to differentiate patients with MS according to their phenotypes and clinical status (i.e., cognitive dysfunction and fatigue) [257, 258]. It should be noted though, that GM atrophy in MS is not limited to the neocortex. Voxel-based morphometry has also been used to draw conclusions about the degree of atrophy

in the subcortical structures of patients with MS [259-262]. Of all the GM structures, the thalamus is perhaps the most studied in MS [237]. It has been shown to be pathologically subject to degeneration, has reasonably well-defined boundaries thus minimizing its potential for partial volume effects on MR imaging, and has extensive reciprocal connections with the cortex and other subcortical structures making it particularly sensitive to pathological changes in other areas of the brain [61]. Further summaries of non-neocortical GM damage can be found in [237].

2.4 COGNITIVE DYSFUNCTION IN PATIENTS WITH MS

The body of knowledge on the role of cognitive dysfunction in patients with MS, unfortunately, suffered the same epistemic history as that on cortical lesions. That is, it was noted very early on, was then forgotten or even erroneously believed not to exist in patients with MS, and garnered little interest in the clinical and neuroscientific literature until a catalyst begat its scientific resurrection. In this case, interest in the field was renewed as a result of the advent of MR imaging in the 1980s—which provided new metrics of neuropathology—and, to a lesser degree, by improved psychometric procedures—which enabled reliable quantification of cognitive abilities.

2.4.1 FREQUENCY AND PREVALENCE

Jean-Martin Charcot, the seemingly omniscient neurologist who endowed MS with nosological status, wrote the following on the cognitive manifestations of the disease: "There is marked enfeeblement of the memory; conceptions are formed slowly; the intellectual and emotional faculties are blunted in their totality." [263]

Although his clinical observations of a few case studies were insufficient to draw any epidemiological conclusions, Charcot did recognize cognitive dysfunction as part of the disease over a hundred years ago. Despite this insight, for the better part of the century that followed, patients were told that MS does not cause memory problems [264]. This assertion was driven mostly by a couple of highly influential studies that grossly underestimated the prevalence of the problem. The first was the 1926 study of 100 patients with MS by Cottrell and Wilson wherein they concluded that cognitive dysfunction was "minimal and negligible" [265]; the second was Kurtzke's 1970 study in which, on the basis of clinical examination, he estimated that cognitive difficulties affected less than 5% of the patients, and if present, were generally confined to patients with severe physical disability [266].[xlvii] By the 1990s, however, the view on cognitive

[xlvii] Not all studies were as dismissive of cognitive dysfunction. In 1951, Aaron Herman Canter found a significant drop of 13.5 IQ points in a group of 23 men that had developed MS after joining the military [267]. To this day, it remains the only study to have the unique opportunity of directly comparing premorbid IQ with that of post-MS onset IQ.

impairment in patients with MS had changed drastically, cemented by the National MS Society's Cognitive Function Study Group's publication that estimated the prevalence of MS cognitive impairment to be 54% to 64% [268]. Two subsequent studies that used more representative community-based samples, as opposed to the almost certainly biased clinical sample of Peyser et al. [268], found slightly lower figures for prevalence: 43% [269] and 46% [270] respectively. Today, after countless studies have examined the frequency and prevalence of cognitive dysfunction in MS, it is generally agreed that approximately half of patients with MS suffer from cognitive dysfunction, though the numbers vary greatly depending on the specific cohort studied.[xlviii]

2.4.2 NATURE

Over the years, it has become oxymoronically clear that no single cognitive domain is always affected, and that no domain is consistently spared in patients with MS who suffer from cognitive dysfunction. The pattern and severity of impairment varies from individual to individual, and yet, there appears to be some functions that are more affected than others. What follows then is a review of the salient aspects of the cognitive nature and domains affected in MS. For a more comprehensive review, the reader is referred to Feinstein's textbook on the clinical neuropsychiatry of MS [272].

In addition to general intelligence, the nature of cognitive deficits in patients with MS is often pigeonholed into the following cognitive domains: attention, memory, executive function, and language.

Attentional deficits refer to both visual and auditory deficits as well as delayed information processing. This limitation in lay terms is referred to as a multitasking problem, though strictly speaking, the term is somewhat of a misnomer as most complex attention tasks involve *switching* between two or more tasks. Indeed, patients with MS are able to fix and maintain their attention, even on a seemingly complex task, but are

[xlviii] For example, in a study that targeted a less disabled subgroup of patients (i.e., EDSS less than 4, relapse-free for at least 2 months, etc.) [271], they found the prevalence of cognitive impairment to be less than half that reported in the samples analyzed by Rao et al. [269] and McIntosh-Michaelis et al. [270].

severely limited when required to alternate between tasks [264, 273, 274]. The problem seems to stem from a slowed information processing speed, and it is this delayed processing speed that is the most common cognitive impairment in patients with MS and may be regarded as a (if not the) core, defining feature of cognitive dysfunction in patients with MS [264, 272, 275-278].

Short-term versus working versus long-term, procedural versus declarative, implicit versus explicit, and episodic versus semantic memory are just some of the labels given to the many constituents that fall under the all-encompassing banner of memory[xlix]; thus, generalizations about memory can often obscure a more complex picture. With that in mind, a substantial number of patients with MS have been found to have some degree of memory impairment, with one particular well-sampled study reporting that 30% have severe memory impairment, 30% are moderately impaired, and only 40% present with little or no impairment [279]. Deficits in working and long-term memory (i.e., recall memory) and metamemory (i.e., the ability to accurately appraise one's own memory) are well documented [272, 280, 281], whereas implicit memory (i.e., memory not reliant on conscious recall; recognition memory) is generally spared [272, 282].

Executive, or higher-order, function involves planning, abstract thinking, goal setting, problem solving and judgment, and while a deficit in such function is often noted in patients with MS, that so many factors contribute to the complexity of these abilities has made it difficult to estimate the frequency and severity of these problems in patients with MS [264, 283-285]. It should be mentioned though that, whereas attentional and even some memory deficits can be overcome quite easily in daily life, executive function impairment can have a drastic effect on the quality of life of patients with MS. For example, poor decision-making abilities can lead to serious financial troubles, and most

[xlix] Neuropsychologists differ when it comes to the classification of memory, leading to the obtuse taxonomy. Perhaps the broadest division cleaves memory into explicit (also known as declarative/conscious/effortful) and implicit (procedural/automatic) memory. Explicit memory is further split into short-term, working, and long-term memory, with long-term being subdivided according to episodic and semantic subtypes.

workplace tasks require executive functioning skills—thus making it difficult for patients with MS to sustain employment [264, 286].

Mercifully, language is usually preserved in patients with MS [287]. Aphasia is extremely rare [2]; however, in keeping with results from the other cognitive paradigms, a minority of patients with MS have been found to perform poorly on tests of language, and when the cognitive domain of language is broadened to include verbal and/or lexical memory, the impairment can be quite profound [269, 272].

2.4.3 ASSESSMENT

Unfortunately, cognitive ability is not accurately observable during routine neurological examinations, and self-reported questionnaires are plagued by subjectivity and often confounded by the subject's mood [288]. Acquiring a full neuropsychological assessment from a trained specialist can be time-consuming and costly and, in reality, is just not readily available in many clinical centers. Because the quest for a brief, reliable, and easy-to-score method for eliciting cognitive deficits has no formal beginning or end, the remainder of this section will focus on some of the more recent testing batteries that have been used, as well as their pitfalls.

The Mini-Mental State Examination

Developed in 1975, the Mini-Mental State Examination (MMSE) examines orientation to time and place, attention, short-term memory, constructional ability, and language, with the individual scores summed to give a total cognitive index (out of 30). Cut-off scores of less than 20 [289] and 24 [290] have been used to signify dementia. The MMSE's strength is in its brevity, taking approximately 5 to 10 minutes to complete, though its sensitivity as a screening tool has been vetted and found lacking[l], making it a less than desirable choice for clinicians.

[l] Huber et al. and Rao et al. found important differences between patients with MS and controls, yet the mean MMSE scores for the MS samples in both studies were well above the cut-off point for dementia [269, 291]. Attempts to change the cut-off point were also investigated, but still yielded only 20% sensitivity [292].

One of several attempts to rescue the MMSE was made by Beatty and Goodkin [293], who recognized the need for brevity but set about increasing the diagnostic sensitivity of the testing battery. They evaluated the usefulness of the Boston Naming Test (BNT), the Wisconsin Card Sort Test (WCST), the Symbol-Digit Modality Test (SDMT), a verbal fluency test (FAS), and various memory paradigms. What they found was that the SDMT, followed by the WCST and a test of immediate recall, was the most sensitive; thus, they advocated that the SDMT, which only takes a couple of minutes to administer, be added to the MMSE.

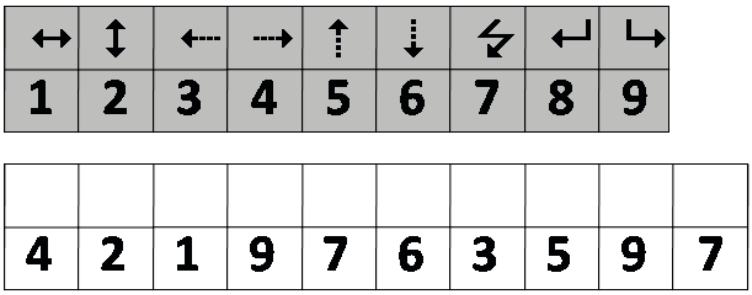

Figure 2.23 – Symbol-Digit Modalities Test (SDMT). Participants are first presented with a legend of symbols that correspond to digits (top, grey box) and are then asked to transcode as many symbols as possible within a given period of time [294]. The simple substitution task was designed and is purported to be a measure of processing speed. The test can be administered orally in order to remove dependence on motor deficits, though strictly speaking, in today's parlance, the written version of the test is referred to as the digit-symbol coding test. Note: symbols displayed in this figure were custom selected by the author and are not the actual symbol set used in the copyright-protected test.

The Brief Repeatable Neuropsychological Battery

In 1991, Rao and colleagues published two seminal papers on cognitive dysfunction in MS [269, 286]. In them they described the frequency (43% of a community-based sample) and patterns (nonuniform and impaired more frequently on measures of recent memory, sustained attention, verbal fluency, conceptual reasoning, and visuospatial perception and less frequently on measures of language and immediate and remote memory) of cognitive dysfunction in patients with MS. They also described the impact of cognitive impairment on employment (patients with cognitive impairment were less likely to be working) and social functioning (engaged in fewer activities, greater sexual dysfunction, and greater psychopathology), suggesting that cognitive dysfunction is a

major factor in determining the quality of life of patients with MS. Despite these informative and highly relevant findings, perhaps the greatest contribution of those papers was the development of a brief screening battery to aid in the prediction of cognitive impairment in patients with MS. Originally comprising four tests[li], the now five-test Brief Repeatable Neuropsychological Battery (BRNB) [295] comprises the Consistent Long-Term Retrieval test (CLTR), Total Recall from the 10/36 Spatial Recall Test, the Controlled Oral Word Association Test (COWAT), the Paced Auditory Serial Addition Test (PASAT), and the SDMT and takes about 30 to 35 minutes to complete. Despite its much-improved sensitivity over the MMSE, the BRNB has been criticized for its high false negative rate and the narrow range of cognitive abilities assessed. Still, with 15 alternative, equivalent versions facilitating its use in longitudinal studies and normative data corrected for age, gender, and education, the BRNB for better or worse has come to shape the current state of cognitive screening tools used today and, in many cases, is still being administered.

The Screening Examination for Cognitive Impairment

The 20- to 25-minute Screening Examination for Cognitive Impairment (SEFCI) boasts an impressive 86% sensitivity and 90% specificity to detect impairment [296]. It comprises the SDMT, measures of learning and delayed verbal recall, as well as tests of vocabulary and verbal abstraction. In head-to-head comparisons, the SEFCI has proven superior to the BRNB [297] and by all statistical inferences should be more popular than its predecessor, yet the BRNB is cited eight times more often than its equally worthy alternative [272]. Though the dominance is perhaps unwarranted, the popularity of the BRNB can be attributed to (i) its multiple, equivocal forms, which make it preferable for longitudinal studies, and (ii) its publication date (five years before the SECFI), which make it the de facto screening test and yardstick for all subsequent batteries.

[li] The four-test battery, termed the Neuropsychological Screening Battery for Multiple Sclerosis (NPSBMS), differs from the BRNB in that it does not include the SDMT, and the 10/36 recall test used less items (7/24).

The Minimal Assessment of Cognitive Function in MS

In contrast with the previous "brief" screening batteries, the Minimal Assessment of Cognitive Function in MS (MACFIMS) is a much more comprehensive battery, requiring approximately 90 minutes to administer. The MACFIMS resulted from a consensus by a group of neuropsychologists from the USA, Canada, the UK, and Australia who aimed to provide guidance on the minimal assessment required by MS patients [298]. The test assesses cognition in the following domains: processing speed and working memory (PASAT, SDMT), learning and memory (California Verbal Learning Test-II, Brief Visuospatial Memory Test-Revised), executive function (D-KEFS Sorting Test), visual perception/spatial processing (Judgment of Line Orientation Test), and language (COWAT). Although largely based on the BRNB, importantly, the MACFIMS also includes an assessment of spatial processing and higher executive function, two domains that the BRNB was recognized to neglect. The MACFIMS is increasingly utilized, and a 2009 cross-sectional study deemed the two batteries as having comparable sensitivity and diagnostic specificity [299].

Full neuropsychological batteries

Given the heterogeneous nature of cognitive dysfunction in patients with MS, in many cases, the only way to adequately evaluate cognitive impairment is by means of a full neuropsychological battery that takes more than two hours to administer. Considerable resources are required to administer such lengthy tests, and because slightly different tests are preferred depending on the clinic or the clinician administering the test (i.e., neuropsychologist, clinical psychologist, speech-language pathologist, occupational therapist, etc.), these full batteries, while extremely useful for individual diagnosis and monitoring, make it difficult for cross-centre validation/comparative studies and clinical trials to be performed.

2.4.4 IMAGING CORRELATES OF COGNITIVE DYSFUNCTION

As with the quest for a single all-encompassing (and ever-elusive) cognitive screening test, the MR imaging community also seeks that holy grail of surrogates. As T2w WM hyperintensities were (and still are) the most conspicuous on MR imaging, it is no surprise that these were the first metrics to be correlated with cognitive performance. Rao and colleagues' early paper [300] was one of many that suggested that total hyperintense lesion area was a robust imaging predictor of cognitive dysfunction. Today, after many improvements in image acquisition and lesion identification, the correlation hovers around an *r* coefficient of 0.40 to 0.50 at best [301].

Another striking aspect of the MR images of brains affected by MS is the amount of grossly observable brain atrophy. Various brain volume correlates have been tested as predictors of cognitive impairment; in general, they are better predictors of cognitive performance than is lesion volume, both in cross-sectional [302] and longitudinal studies [303]. Correlations have been achieved by measuring whole brain atrophy [302-306] or the atrophy of specific brain regions such as the deep GM [306-308], corpus callosum [309, 310], cerebral cortex [246, 247, 311], mesial temporal lobes [312, 313], and other subregions of the cerebral cortex [254, 314].[lii] The body of work on cortical atrophy alone underscores the clinical relevance of GM atrophy as a risk factor for cognitive impairment.

While lesion burden and brain volume measurements are revealing, they are not very informative about the pathological substrates (i.e., demyelination, edema, inflammation, axonal loss, etc.). Magnetization transfer ratio, though not wholly specific to demyelination, still provides more insight on the nature of the damage present. It has also been found to be a predictor of cognitive dysfunction [315-317], to the point that, in a multivariate regression study, average cortical/subcortical MTR emerged as the most robust predictor, exceeding both T1w and T2w lesion load [318], a finding supported by

[lii] In fact, so many different MR measures of focal, regional, and/or global damage have been correlated with cognitive performance in patients with MS that, unfortunately, the literature has become diluted and difficult to navigate.

another study that found that MTR accounts for slightly more variance than WM lesion burden [319]. Decreases in the MTR of specific cortical regions have also been found to correlate with cognitive dysfunction (specifically, poor PASAT performance) in both patients with RRMS [320] and PPMS [321]. In a 2005 functional MR imaging (fMRI) study of patients in the earliest stage of the disease (i.e., clinically isolated syndromes), significant correlations were also found between NAWM MTR values and PASAT performance. Interestingly, the fMRI data revealed an inverse relationship between regional activation and MTR, suggesting that there are compensatory brain mechanisms working to offset structural damage [322].

Magnetic resonance spectroscopy (MRS), which provides a quantitative measure of important metabolites within the brain, can also be used as a more specific pathological indicator. The measurement of the ratio of N-acetylasparate to creatine via MRS is a putative indicator of axonal integrity and has been found to correlate with cognitive dysfunction. Specifically, there is an association with processing speed (assessed via the SDMT; $0.39 \geq r \geq 0.43$) and, to a lesser degree, a relationship with verbal and visual memory and attention [323-325] and with verbal learning [326].

Finally, even though the reliable identification of cGM lesions on MR images is still in its infancy and that a large proportion of these lesions are missed (i.e., Type III subpial demyelinating lesions), several studies have already attempted to correlate this marker of cortical pathology with measures of cognitive impairment. A longitudinal study using DIR for the detection of cGM lesions found that the number of total cortical lesions at three-year follow-up was related to visuospatial memory and processing speed [327]. Hippocampal lesion count was also related to visuospatial memory deficits, but, unfortunately, so was WM lesion number. Since the number of WM lesions was not controlled for in the study's statistical models, it is difficult to distinguish the independent contributions of the different types of pathology. The study was also relatively small (N = 13). A subsequent, and larger, study (N = 70) also used the DIR sequence to identify cGM lesions and found a relationship with both the number and

volume of cortical lesions and a measure of global cognitive impairment (i.e., patients who score two standard deviations below the mean normative values on at least one test of the BRNB were considered cognitively impaired) [328]. The study also found a modest correlation between the number and volume of cortical lesions and normalized cortical volume, but importantly, when each of these was controlled for, it was found that both the cGM lesion volume and the normalized cortical volume were independent predictors of cognitive dysfunction in patients with MS (partial $r = 0.597$ and 0.444, respectively, $p < .001$).

2.5 MOVING FORWARD

To summarize, MS is an immune-modulated disease of the human CNS that is particularly relevant in Canada, where it is considered the most common neurological disease affecting young adults. The traditional pathological hallmark of MS has been the WM lesion; however, recent studies have highlighted the nontrivial presence of previously overlooked GM pathology. Studying GM pathology in vivo, though, has thus far presented a challenge to the imaging community, as various techniques capture but a limited view of the total picture. As the appreciation grows for cortical GM pathology in patients with MS, so does the need for imaging methods that capture said pathology and for understanding the relationship of those methods to the clinical outcome of the patient.

With this as the rationale behind this work, and to reiterate, the main objective of this thesis was to evaluate and develop image processing methods for the in vivo quantification of GM pathology (i.e., atrophy and lesional pathology) in patients with MS.

The background chapter and literature review now concluded, the reader should be sufficiently armed to fully absorb the material presented in the remaining chapters of this thesis, beginning in Chapter 3 with an evaluation of commonly used image processing methods for the quantification of GM atrophy in patients with MS.

Chapter 3

AUTOMATED GREY MATTER SEGMENTATION METHODS

As discussed in the previous chapter, the last decade bore witness to a marked shift in the focus of MS research from WM pathology to that of GM. With a prolific research community seeking to sift through the numerous and sometimes contradictory studies of MR imaging metrics of neurodegeneration and their relationship to the clinical characteristics of MS, an obvious first step to understanding the role of GM in the disease is to evaluate some of the commonly used methods that have been responsible for shaping our modern day view of GM in MS.

The following manuscript, which was published in *NeuroImage*, sought to evaluate the accuracy of the GM segmentations performed by several automated techniques. Of several important findings, first and foremost was the large discrepancy between automated techniques in terms of their ability to accurately segment deep GM versus cortical GM—a result the author has seen firsthand influence the design of large-scale clinical trials of disease-modifying therapies for patients with MS.

An earlier form of this work was published as a conference proceeding and presented at the Medical Image Analysis on Multiple Sclerosis (MIAMS) workshop of the 2009 Medical Image Computing and Computer Assisted Intervention (MICCAI) conference [329].

EVALUATION OF AUTOMATED TECHNIQUES FOR THE QUANTIFICATION OF GREY MATTER ATROPHY IN PATIENTS WITH MULTIPLE SCLEROSIS

MISHKIN DERAKHSHAN[1,2], ZOGRAFOS CARAMANOS[1,2], PAUL S. GIACOMINI[1], SRIDAR NARAYANAN[1,2], JOSEFINA MARANZANO[2], SIMON J. FRANCIS[1,2], DOUGLAS L. ARNOLD[1,2], AND D. LOUIS COLLINS[1]

[1]McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, Montreal, Canada H3A 2B4
[2]NeuroRx Research, 3605 University St., Suite #4, Montreal, Canada H3A 2B3

NeuroImage 52 (2010) 1261-1267

3.1 ABSTRACT

Several methods exist and are frequently used to quantify grey matter (GM) atrophy in multiple sclerosis (MS). Fundamental to all available techniques is the accurate segmentation of GM in the brain, a difficult task confounded even further by the pathology present in the brains of MS patients. In this paper, we examine the segmentations of six different automated techniques and compare them to a manually defined reference standard. Results demonstrate that, although the algorithms perform similarly to manual segmentations of cortical GM, severe shortcomings are present in the segmentation of deep GM structures. This deficiency is particularly relevant given the current interest in the role of GM in MS and the numerous reports of atrophy in deep GM structures.

3.2 INTRODUCTION

Multiple sclerosis (MS) is a chronic inflammatory disorder of the central nervous system. Focal white matter (WM) lesions represent the hallmark pathological finding of MS; however, increasing evidence from pathological studies has underscored the importance of grey matter (GM) involvement as well [32, 39, 330]. GM pathology does not seem to correlate with focal WM lesions [331, 332], and neocortical GM volume loss has been shown to be related to worsening cognition [246]. As our appreciation for the importance of GM pathology grows, reliable imaging methods are essential to accurately measure and analyze GM pathology in MS.

One of the challenges in classifying GM and WM in the brains of patients with MS results from the presence of WM lesions. Previous studies [241, 244] have shown that lesions lead to misclassifications of other tissues, the majority of which are WM erroneously labeled as GM. Even the most basic correction method of adding the lesion volume to the segmented WM volume may be insufficient to obtain accurate volumes for WM and GM compartments, as all segmentation failures are presumed to involve only WM lesions.

In the present study, we examine the GM classification results of six automated methods used to detect GM atrophy in the brains of MS patients. These include the two most commonly reported techniques: (a) a voxel-based morphometry (VBM) approach, executed most commonly with the statistical parametric mapping (SPM) software suite [206] and (b) SIENAx [202], as well as (c) FIRST [333], (d) Freesurfer [208], (e) a classifier publicly available from the Montreal Neurological Institute (MNI) [188], and (f) a multispectral Bayesian classifier (MBC) designed specifically for segmenting the brains of MS patients [182]. In contrast to previous studies that focused on lesion misclassification in MS, our current work is specific to the accuracy of GM segmentations, both for cortical GM (cGM) and deep GM structures (dGM).

Given the complexity of the cerebral anatomy, combined with partial volume effects present in MRI data, it is well known that manual segmentation is difficult and time consuming. Furthermore, differences in interpretation of image intensity and contrast with respect to the anatomy can lead to significant variability in voxel labeling between readers. In order to minimize errors and reduce variability, we decided to solicit expert readers (i.e., radiologists, neuroradiologists, and neurologists) trained in manual segmentation on MRI to obtain the highest quality manual GM segmentations possible. Given that the time of these experts is limited, we were restricted to analyzing a small number of slices on a small number of subjects.

The focus of this study is to explore the validity and the variability of some of the freely available automated methods currently being used to segment GM and to estimate GM atrophy in MS. Although the assessment of the six techniques listed above is limited to three slices within the brains of three subjects, this was enough to demonstrate that (a) there is variability in GM segmentation between the different software packages; (b) this variability is quite high for deep GM structures; and (c) users must be careful when interpreting the results of automatic classification programs and when comparing results between studies.

3.3 METHODS

3.3.1 SUBJECT AND ACQUISITION DETAILS

Three subjects with secondary progressive MS were selected from a multicenter clinical trials dataset. The subjects were chosen at random for their low, medium, and high WM lesion loads of 2.4 cc, 8.6 cc, and 24 cc, respectively. Subjects' scans were acquired from three different centers, all at a field strength of 1.5 T, and included T1, T2, and proton density (PD)-weighted sequences with a voxel size of $0.98 \times 0.98 \times 3\ mm^3$. Consistent with previous reports of GM atrophy in MS, the T1w scan was used as the sole input to the automated techniques with the exception of MBC, for which all three modalities were mandatory. The images acquired were typical of those commonly obtained in clinical trials in terms of scan quality, contrast, and slice thickness, and were comparable to the data used in previous reports of GM atrophy in MS. While this small dataset is in no way representative of the full gamut of a typical MS population, given the constraints identified in the introduction, it still represents varying disease loads and levels of brain atrophy and is sufficient to demonstrate the variability in GM segmentation between the different software packages compared.

3.3.2 AUTOMATED TECHNIQUES

Statistical Parametric Mapping (SPM) is a software suite of MATLAB functions and subroutines designed for the analysis of brain imaging data sequences. We used the latest version, SPM8b (http://www.fil.ion.ucl.ac.uk/spm). Of the many pipelines in

SPM8b, we were interested in the VBM pipeline for MRI data, which performs region-wise volumetric comparisons between a population of subjects. The images must be spatially normalized, segmented into different tissue classes, and smoothed prior to the performance of statistical tests [206]. The default segmentation algorithm used in SPM8b is identical to that of the older SPM5 [207]. This segmentation produces subject-specific tissue probability maps from which a binary map of each class is obtained by assigning each voxel to the class in which the probability at that voxel was the highest.

SIENAx [202] is the cross-sectional version of the Structural Image Evaluation using the Normalization of Atrophy method (SIENA) [232] and is part of FSL (http://www.fmrib.ox.ac.uk/fsl/). With SIENAx, the brain is extracted from the volume using Brain Extraction Tool (BET) [156] and then segmented using FMRIB's Automated Segmentation Tool (FAST). This segments the image into GM, WM, and CSF, while also correcting for spatial intensity variations and partial volume. The default underlying method, which employs a K-means segmentation in order to estimate the initial parameters of the classes, is based on a hidden Markov random field model and an associated Expectation-Maximization algorithm [203]. SIENAx outputs spatially normalized and non-normalized volumes. In this study, we use the latter, that is, the tissue classification output from FAST in the subjects' native space.

FMRIB's Integrated Registration and Segmentation Tool (FIRST) [333] is a model-based segmentation tool also part of FSL (http://www.fmrib.ox.ac.uk/fsl/first/). Subcortical brain segmentation is performed using Bayesian shape and appearance models constructed from a set of manually segmented images. FIRST uses these models to search through linear combinations of shape modes of variation for the most probable shape instance, given the observed intensities in a T1w image. Since FIRST only segments the subcortical structures, we were only able to include this automated technique in our analysis of dGM.

Freesurfer (http://surfer.nrm.mgh.harvard.edu/) is a freely available image analysis suite that can be used for both cortical reconstruction and volumetric segmentation. In this study, only the volumetric segmentation is examined because of its similarity to the other methods being tested. Freesurfer's volume-based stream is designed to pre-process MRI volumes and label subcortical structures. The stream consists of multiple stages [208, 225]: in brief, the first stage is an affine registration with Talairach space specifically designed to be insensitive to pathology and to maximize the accuracy of the final segmentation. This is followed by an initial tissue classification and correction of the variation in intensity resulting from the B1 bias field. Finally, there is a high dimensional nonlinear volumetric alignment to the Talairach atlas where the final segmentation takes place.

As Freesurfer labels about 50 separate brain structures, in order to obtain our mask of cerebral GM, we combined the left and right labels for the following structures: cerebral cortex, thalamus proper, caudate, putamen, pallidum, hippocampus, amygdala, nucleus accumbens area, and ventral diencephalon.

Medical Imaging NetCDF (MINC) is a medical imaging data format and associated set of tools and libraries developed at the Montreal Neurological Institute (MNI) and freely available online (http://www.bic.mni.mcgill.ca/ServicesSoftware). The tool classify_clean, which is used to classify stereotaxic MINC volumes, involves a Bayesian labeling scheme and a set of standard sample points to compute an initial volume classification. This classification is then employed to purge incorrect tag points from the standard set, yielding a custom set of labels for the particular subject. Finally, this tag point set is used by an artificial neural net classifier to classify the volume [188].

Multispectral Bayesian Classifier (MBC) is software developed specifically for the identification of cerebral WM lesions in patients with MS [182]. MBC is distinct from the other methods because, as a multispectral technique, it uses T1w, T2w, and PDw images. While multiple tissue classifiers are available [183, 187, 188], we chose MBC for its

robustness, having been run on thousands of scans with minimal failures and having been shown to work well on subjects with lesions [88, 334-337]. MBC requires the following inputs: intensity range-normalized [338] multispectral images, spatial probability anatomic maps of CSF, GM, and WM, and a tissue/intensity-specific conditional probability. MBC automatically segments volumes as WM, GM, CSF, and T2w lesions.

3.3.3 MANUAL SEGMENTATIONS

Volumetric analysis was performed with the interactive software package, Display, developed at the McConnell Brain Imaging Center of the Montreal Neurological Institute. This program allows visualization of MR images in axial, coronal, and sagittal orientations to facilitate anatomical interpretation. Six expert readers manually segmented the GM on three separate axial slices: (a) an inferior slice through the temporal lobe, (b) a superior slice above the ventricles at the level of the centrum semiovale, and (c) an intermediate slice that intersected the insula and the basal ganglia. The extremely time-consuming nature of manually segmenting GM limited the number of slices that practically could be investigated. These three slices were chosen because they show anatomy where lesions are typically found and sample different regions of GM, which present different challenges to classifiers. For the purpose of assessing and quantifying segmentation accuracy in these challenging regions, we believe that three slices are sufficient. Accordingly, we acknowledge that our results do not provide a comprehensive picture of all the advantages and disadvantages of each technique for full-brain segmentation.

The readers were a combination of formally trained radiologists, neuroradiologists, and neurologists, each with specialized training and experience in the anatomical labeling of MR images and, on average, having reviewed on the order of thousands of MRI brain scans. The readers were blinded to each other's results, and the voxel-wise labeling process was completely manual; that is, it involved no thresholding. In addition, the window contrast levels were set to be identical for all readers.

From the six manual segmentations, a reference standard label set was established to create a ground truth for identifying GM in the brain. This was accomplished by using the Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm [339]. Briefly, STAPLE is an expectation-maximization algorithm that considers a collection of segmentations and computes a probabilistic estimate of the true segmentation. The end result is an optimal linear combination of the segmentations and is preferable to simple voting rules where the volume of the resulting ground truth image varies greatly depending on the number of votes required.

3.3.4 Performance Evaluation

Agreement between label pairs was measured using the Dice similarity coefficient (DSC) [340], a similarity index that is a special case of the Kappa coefficient. The DSC is the volumetric intersection of the two labels in question (L1 and L2), divided by the mean volume of the two labels:

$$DSC = \frac{2\,|L_1 \cap L_2|}{|L_1| + |L_2|}$$

Because each automated method produces a GM segmentation map of the entire brain, three separate labels were created that corresponded to the three slices segmented by each of the manual readers. To obtain the DSC, for each slice, the labels of each automated method were compared to the STAPLE reference standard. Furthermore, for the intermediate slice that included both cGM and dGM, the DSCs for each of these GM compartments were calculated separately. A manually created dGM mask was used to delineate the dGM from the cGM for each subject.

In addition to the DSC, the percent volume difference (PVD), the absolute value of the percent volume difference (APVD), and the positive predictive value (PPV) were calculated. The PVD exposes any under- or over-estimation in our volumes, while the APVD provides a measure of agreement that is complementary but does not necessarily correspond to the DSC. PPV was chosen because it is a good combination of sensitivity

and specificity; in this case, it reflects the probability that a voxel labeled as GM is in fact GM, and is defined as the ratio of true positives to the sum of true and false positives. Although additional geometric metrics for comparisons such as surface distances were considered, they were not included as they are best for surfaces defined with sub-voxel precision rather than the voxelated labels that were being compared.

3.4 RESULTS

3.4.1 MANUAL SEGMENTATIONS

Inter-reader variability was assessed by examining the mean DCSs for each pair of experts for each slice. The results are presented in Table 3.1. Two-way analysis of variance (ANOVA) was used to test for the effects of slice location (inferior, intermediate, superior) and WM lesion load (low, medium, high) on the experts' mean DCSs. No significant interaction was found between location and lesion load ($F_{df=4,126} = 1.85$, $p = 0.12$), but there was a main effect for both slice ($F_{2,126} = 15.32$, $p < 0.0001$) and lesion load ($F_{2,126} = 24.60$, $p < 0.0001$). As summarized in Table 3.1, post-hoc analyses with Tukey-Kramer HSD tests showed that, with respect to slice location, the expert readers had significantly higher mean DSCs ($p < 0.0001$) on the intermediate slice compared to either the superior or inferior slice. However, it is important to note that, because of the inclusion of the dGM on the intermediate slice, one would expect a higher DSC for those slices simply because of the nature of the DSC to be higher in cases with a larger volume/surface area ratio. Repeating the analysis using volume as a covariate seemed to confirm this relationship, as we no longer saw the increased DSC for the intermediate slice; however, the exact p values varied depending on which volume was used as the covariate (reader A, reader B, or the mean of readers A and B). With respect to WM lesion load, the expert readers had significantly lower DSCs ($p < 0.0001$) for the case of medium lesion load (8.6 cc). This relationship was preserved when using any of the volumes as a covariate. The mean DSC and standard deviation for manually segmented GM across all pairs of readers and all slices for all subjects was 0.80 (0.04), with a range from 0.60 to 0.96.

Table 3.1 – Inter-reader mean (standard deviation) Dice similarity coefficients.

Subject	Inferior	Intermediate	Superior	Per Subject Mean
Low	0.78 (0.03)^{a,1}	0.85 (0.03)^{a,2}	0.82 (0.04)^{a,3}	**0.82 (0.04)^{a}**
Medium	0.74 (0.04)^{b,1}	0.80 (0.03)^{b,2}	0.76 (0.08)^{b,1,2}	**0.77 (0.06)^{b}**
High	0.82 (0.03)^{c,1}	0.84 (0.04)^{a,1}	0.81 (0.04)^{a,1}	**0.82 (0.04)^{a}**
Per Slice Mean	**0.78 (0.05)^{1}**	**0.83 (0.04)^{2}**	**0.79 (0.06)^{1}**	

Subjects are identified by their WM lesion load: low = 2.6 cc, med = 8.6 cc, high = 24 cc. Subjects with different letter superscripts (i.e., a, b, c) differ significantly (Tukey-Kramer HSD, $p < 0.05$). Likewise, slices with different numeral superscripts (i.e., 1, 2, 3) differ significantly. For example, the experts' average performance across all subjects (bottom row of the table) for the inferior slice was the same as the superior slice (both have superscript 1) but different from the intermediate slice (which has a different numeral superscript, i.e., 2).

The manual segmentations by the six experts were used to create the STAPLE reference standard, which was compared to the performance of each individual expert; the results are shown in Table 3.2. The high DSCs, PPVs, and the low APVDs indicate good agreement between each reader and the reference standard, while the PVD between these values indicates a lack of either a significant over- or under-estimation in the reference standard (*t* test, $t_{53} = -0.95$, $p_{prob>t} = 0.83$, $p_{prob<t} = 0.17$).

Table 3.2 - Mean (standard deviation) of each evaluation metric calculated relative to the STAPLE reference standard for total GM across all three slices and subjects.

	Method	DSC	PVD	APVD	PPV
Manual segmentation	Expert 1	0.88 (0.02)^{b}	-12.59 (5.07)^{**,c}	12.59 (5.07)^{a}	0.95 (0.03)^{a,b}
	Expert 2	0.81 (0.03)^{**,c}	-1.76 (6.68)^{b}	5.58 (3.63)^{a,b}	0.82 (0.03)^{*,c}
	Expert 3	0.93 (0.02)^{a}	-6.44 (5.00)^{b,c}	6.85 (4.34)^{a,b}	0.97 (0.02)^{a}
	Expert 4	0.83 (0.07)^{*,c}	13.78 (11.47)^{***,a}	13.78 (11.47)^{a}	0.78 (0.08)^{***,c}
	Expert 5	0.92 (0.04)^{a,b}	2.58 (4.56)^{b}	4.27 (2.79)^{b}	0.91 (0.04)^{a,b}
	Expert 6	0.88 (0.02)^{b}	-3.83 (6.68)^{b,c}	6.44 (3.80)^{a,b}	0.90 (0.04)^{b}
		0.88 (0.06)	**-1.38 (10.61)**	**8.25 (6.71)**	**0.89 (0.08)**
Automated methods	Freesurfer	0.74 (0.05)^{*,b}	-0.35 (9.60)^{b}	6.70 (6.45)^{a}	0.75 (0.05)^{***,a}
	MBC	0.83 (0.02)^{a}	6.02 (11.97)^{a,b}	9.69 (8.85)^{a}	0.82 (0.05)^{a}
	MNI	0.80 (0.06)^{**,a,b}	18.93 (15.73)^{***,a}	18.93 (15.73)^{**,a}	0.74 (0.09)^{***,a}
	SIENAx	0.72 (0.09)^{***,b}	-7.79 (17.79)^{b}	16.66 (8.50)^{*,a}	0.76 (0.10)^{***,a}
	SPM	0.80 (0.06)^{**,a,b}	5.85 (13.88)^{a,b}	9.56 (11.32)^{a}	0.78 (0.08)^{***,a}
		0.78 (0.06)	**4.53 (13.79)**	**12.31 (10.17)**	**0.77 (0.07)**

Instances where the evaluation metric were significantly worse than the average expert manual segmentation were tested with Dunnett's control test and are marked accordingly ($p < 0.001$***, $p < 0.01$**, $p < 0.05$*). Only the MBC produced results that did not differ from the expert manual segmentation of total GM. As in Table 1, methods with different letter superscripts (i.e., a, b, c) showed significantly different performances for the particular evaluation metric (Tukey-Kramer HSD, $p < 0.05$). PVD = Percent Volume Difference, APVD = Absolute Percent Volume Difference, DSC = Dice Similarity Coefficient, PPV = Positive Predictive Value, MBC = Multispectral Bayesian Classifier, MNI = classifier developed at the Montreal Neurological Institute.

3.4.2 AUTOMATED RESULTS BY SLICE

The overall performance for total GM across all slices can be seen in Table 3.2. Figure 3.1 shows the DSCs of the STAPLE reference standard versus each of the segmentations generated by the automated methods, as well as by the expert readers; these are all grouped according to slice location. One-way analyses of variance suggested that these mean DSCs were comparable for all automated methods for the inferior ($F_{4,10} = 0.80$, $p = 0.56$) and superior slices ($F_{4,10} = 1.75$, $p = 0.21$), but not for the intermediate slice ($F_{4,10} = 4.05$, $p = 0.03$). In addition, in the intermediate slice, most of the automated methods did not perform as well as the experts ($p < 0.05$), as detected with Dunnett's multiple comparison test, which tests whether mean DSCs differed from the mean of the STAPLE reference standard (asterisks in Figure 3.1). This reduced performance prompted further examination of the automated GM segmentations at this level.

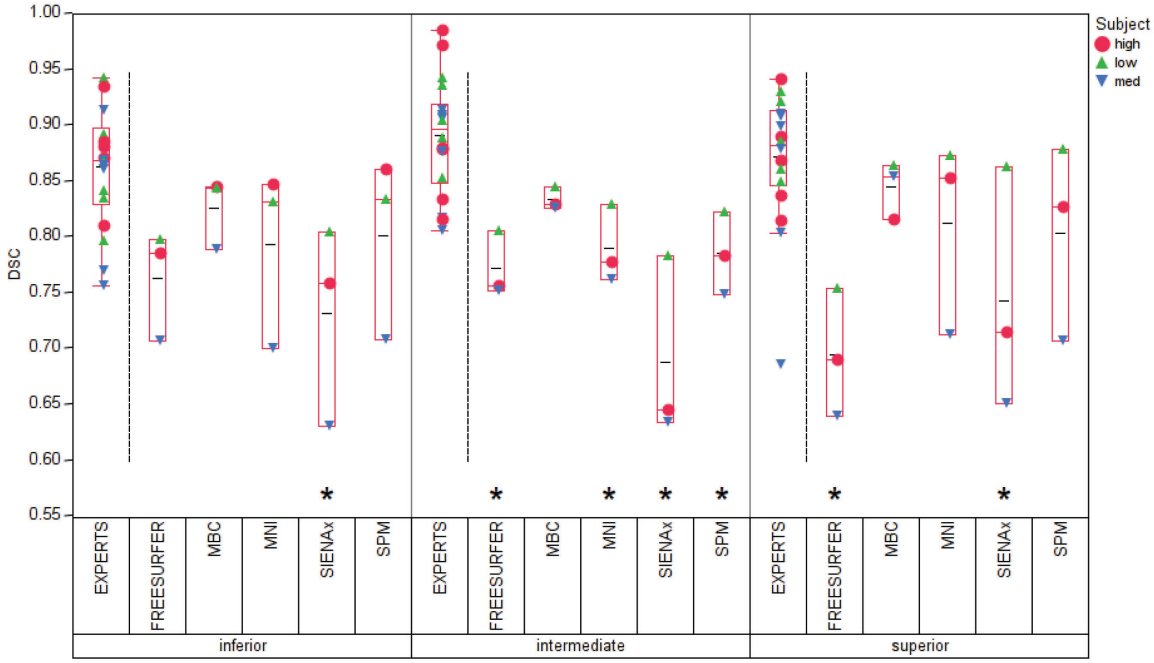

Figure 3.1 - Dice similarity coefficients (DSCs) of the experts' and each method's segmentation of total grey matter versus those of the STAPLE reference standard, grouped by slice. Methods that differed significantly from the experts are marked with an asterisk (Dunnett's control test, p < 0.05). Refer to Table 2 for abbreviations. Subjects are identified by their WM lesion load: low = 2.6 cc, med = 8.6 cc, high = 24 cc. Note that with the exception of SIENAx, which performed worse than the experts on all slices, most of the methods' DSCs differed from those of the experts only on the intermediate slice, where segmentation of dGM was also required. Only the MBC did not differ from the experts on this slice.

3.4.3 AUTOMATED RESULTS BY GREY MATTER COMPARTMENT

For all of the automated segmentation methods, the mean cGM volumes are similar both to each other (ANOVA, $F_{4,10} = 2.08$, $p = 0.16$) and to the STAPLE reference standard (Dunnett's test, all $p > 0.30$) (see Figure 3.2). However, this was not true for the mean dGM volumes, which differed amongst automated segmentation techniques (ANOVA, $F_{5,12} = 16.93$, $p < 0.0001$), with the volumes from SIENAx and SPM being smaller than that of the reference standard ($p = 0.0003$ and $p = 0.0259$, respectively, indicated by the asterisks in Figure 3.2 and clearly evident in Figure 3.4). The DSCs for dGM also showed significant variation (ANOVA, $F_{5,12} = 37.67$, $p < 0.0001$) (Figure 3.3). SIENAx and SPM dGM segmentations showed the least amount of overlap with the STAPLE reference standards, with mean DSCs of 0.25 and 0.51, respectively, indicating results with very little agreement. On the other hand, FIRST and Freesurfer showed a high degree of overlap with the STAPLE reference standards, with mean DSCs of 0.87 and 0.85, respectively. Figure 3.4 shows the results of the segmentations of the intermediate slice from a single subject (WM lesion load 8.4cc) by all methods, clearly illustrating the variability and inaccuracies that exist in the segmentation of dGM.

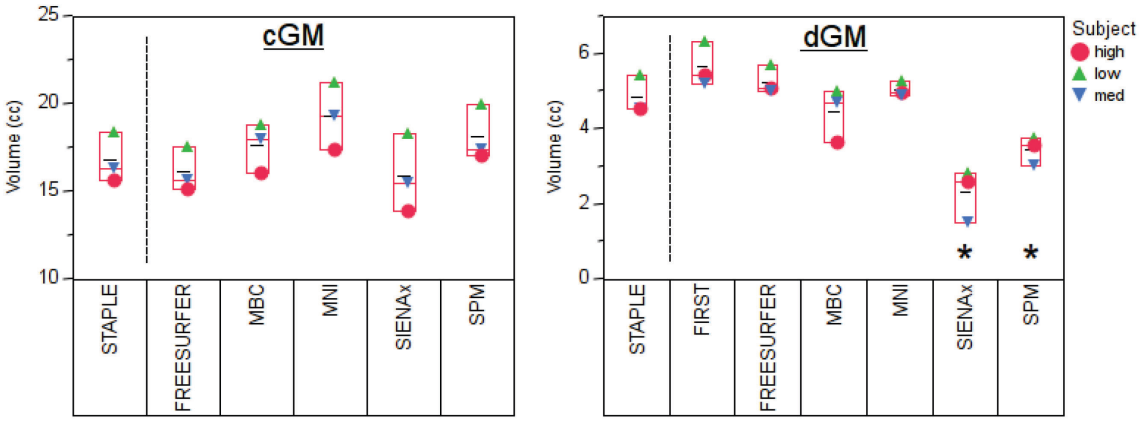

Figure 3.2 - Volumes in cubic centimeters (cc) of segmented cortical grey matter (cGM, left) and deep grey matter (dGM, right) for each automated method as well as for the STAPLE reference standard. Methods that differed significantly from the STAPLE reference standard marked with an asterisk (Dunnett's control test, p < 0.05). Refer to Table 2 for abbreviations. Subjects are identified by their WM lesion load: low = 2.6 cc, med = 8.6 cc, high = 24 cc. Note that all the methods had similar volumes to the STAPLE reference standard when segmenting cGM, and that in the case of dGM segmentation, the volumes from SIENAx and SPM were drastically lower.

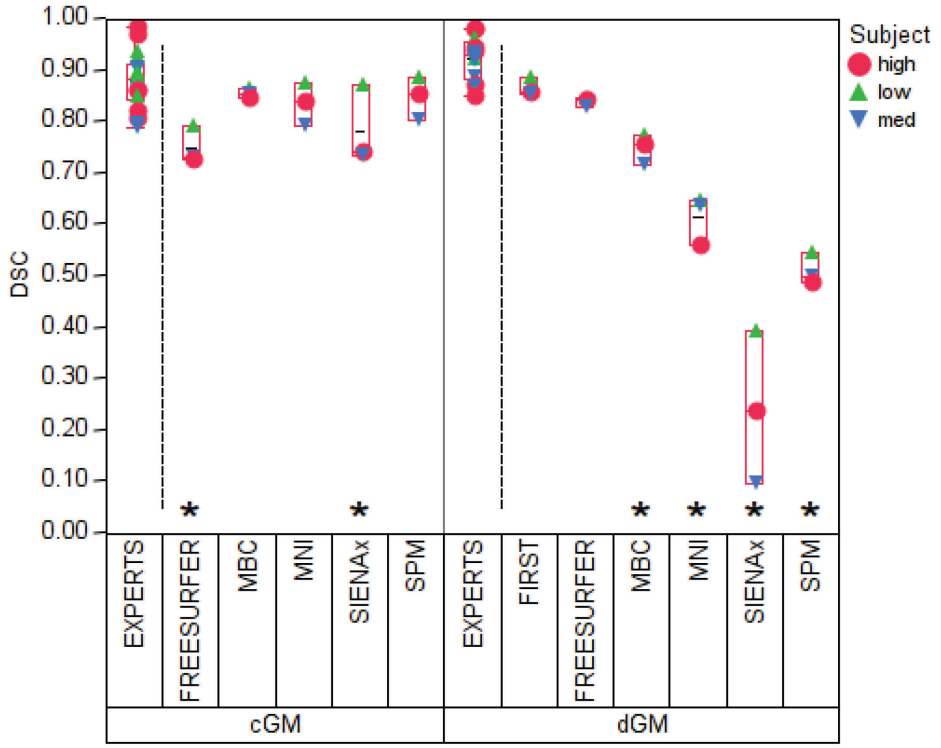

Figure 3.3 - Dice similarity coefficients (DSCs) of the experts' and each automated method's segmentation of grey matter versus those of the STAPLE reference standard, separated by cortical grey matter (cGM, left) and deep grey matter (dGM, right). Methods that differed significantly from the experts are marked with an asterisk (Dunnett's control test, $p < 0.05$). Refer to Table 2 for abbreviations. Subjects are identified by their WM lesion load: low = 2.6 cc, med = 8.6 cc, high = 24 cc. Note the increased number of methods unable to perform as well as the experts on dGM compared to cGM, as well as, the low DSCs for SIENAx and SPM, which indicate very little agreement with the experts' dGM segmentations.

3.5 DISCUSSION

This study aimed to evaluate some of the most commonly used automated techniques for measuring GM atrophy on MRI data typically acquired in clinical trials. While previous studies have touched upon possible pitfalls of some of these techniques [239, 241, 244, 341], to the best of our knowledge, ours is the first to explore the problem at the root of every technique, namely, GM segmentation.

Undoubtedly, the accurate segmentation of GM is difficult to achieve. The cortex is a thin layer of GM surrounded by WM on one side and CSF on the other, both of which produce partial volume effects that confound its delineation. (The small number of subjects in this study did not allow the evaluation of the different techniques with

respect to partial volume effects. This will be the subject of future work.) Even among expert readers, a discernable amount of variability is found in the segmentation of GM. Here, this variability is evidenced by the inter-reader DSCs, with a mean of 0.80 for total GM from every slice (Table 3.1). This mean value highlights the difficulty in obtaining or estimating a known true segmentation for clinical data. We considered using a synthetic image from an MRI simulator or a phantom that would include the imaging system characteristics, but such phantoms do not reproduce the full range of imaging artifacts (partial volume, intensity inhomogeneity, noise, etc.). Moreover, phantoms typically do not reproduce the normal and pathological anatomical variability observed in clinical data, and their use would lower the clinical relevance of this study. Instead, by using STAPLE to estimate and minimize the error in those areas of GM segmentation where the experts disagreed, we created an optimal linear combination of the experts' manual segmentations in order to establish a more accurate GM segmentation reference standard.

In terms of the overall GM segmentation, the performance evaluations of the automated techniques seem comparable to, though not as good as, the expert segmentations (Table 3.2). That is, for many of the evaluation metrics, the automated techniques perform almost as well as, and in some cases, the same as, the expert manual segmentations. Nevertheless, judging the techniques by considering only the performance evaluations for total GM would be misleading.

In contrast to the abovementioned difficulties in segmenting cGM, dGM structures are affected by partial volume to a much smaller degree and only in the outermost layers of the structures. Thus, it is not surprising that the experts' manual segmentations showed significant improvement ($p = 0.01$) in their mean DSCs, increasing from 0.88 to 0.93 for cGM to dGM, respectively. However, the reverse was found for the automated methods: the automated segmentation of dGM is much less accurate than that of cGM. For instance, the automated techniques tended to misclassify large portions of dGM as WM (Figure 3.4). Furthermore, FIRST and Freesurfer were the only automated methods that

produced relatively accurate segmentations of dGM. Given that these particular two methods were designed for subcortical segmentation and explicitly label individual subcortical structures in their optimization routines, this result is both consistent and reassuring. Conversely, the other automated methods were lacking in their ability to produce accurate segmentations of dGM structures, as evidenced by their mean DSCs: 0.25, 0.51, 0.62, and 0.75 for SIENAx, SPM, MNI, and MBC, respectively. Of particular concern are the low results for SIENAx and SPM, as they are the predominant methods employed in the literature and, in many instances, are the bases for conclusions drawn regarding the degree of atrophy in dGM structures [246, 261, 262, 342-345].

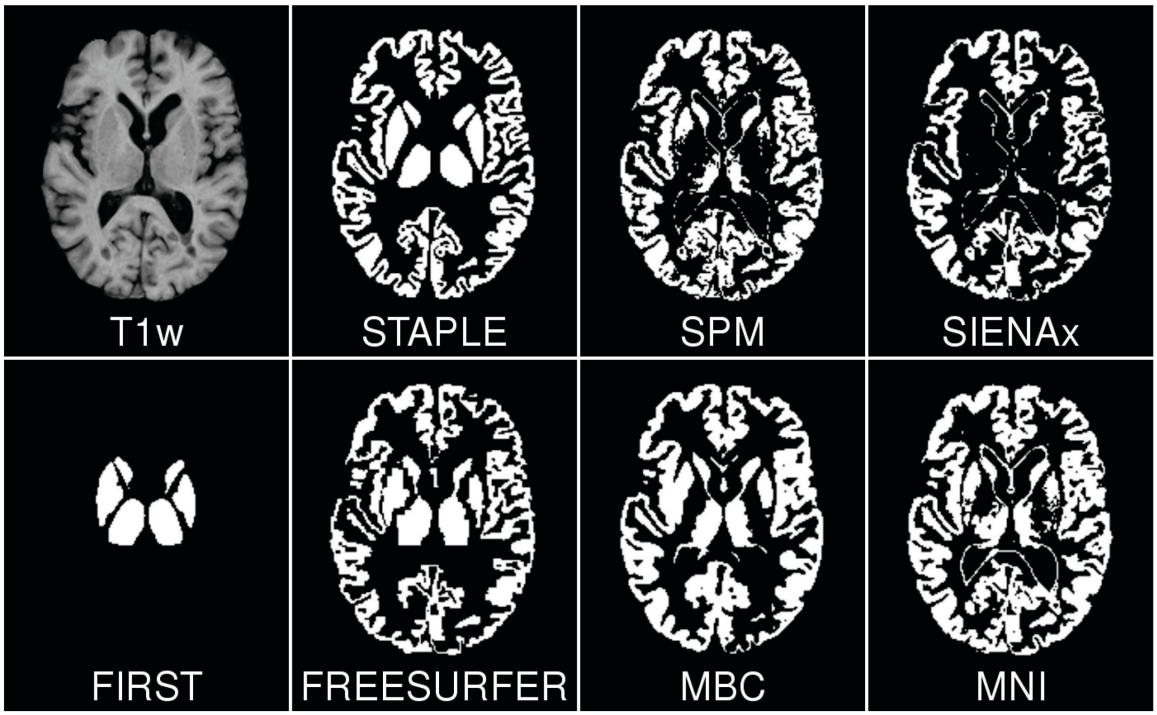

Figure 3.4 - Grey matter segmentations and the T1w image of the intermediate slice for a single subject. Note the accuracy of the STAPLE reference standard, which is a linear combination of the segmentations by our six experts, and some of the gross inaccuracies evident in the dGM segmentation by SPM and SIENAx in particular. Recall that FIRST only provides dGM segmentations. Refer to Table 3.2 for abbreviations.

A number of methods exist to help address the variable accuracy of GM segmentation by automated techniques. Some groups have used a high threshold on the probabilistic segmentations of GM so that the VBM analysis only includes voxels with a particular certainty of being GM [261, 262, 346], while others have explicitly masked the dGM structures in order to report only cGM volumes [347, 348]. However, most have been solely concerned with addressing the misclassification of WM lesions, which are most frequently misclassified as GM [349]. A common solution to this problem has been to simply mask the WM lesions so they are included in the volume of WM and not of GM. In another approach [350], the WM lesions are filled with an intensity similar to WM prior to the classification even though, depending on the size of the lesions and the intensity used to fill them, the degree of misclassification may be greater than if nothing had been done at all [351]. Still, despite the variability of the techniques used to address this problem, every attempt is worthwhile given that typical WM lesion misclassifications (as measured with SPM) are on the order of 0.56% of brain parenchyma [244], a significant number considering the rate of atrophy we are trying to detect in patients with MS is on the order of 1% per year [352, 353]. Even more concerning are the gross misclassifications of dGM shown in this study that account for approximately 3.8% of brain parenchyma in the case of SPM, again, an unacceptably high number.

While a high degree of accuracy may be the primary goal of the automated segmentation techniques examined in the paper, accuracy is not necessarily the only way to gauge a tool designed to measure atrophy. The automated techniques described above can be used to detect atrophy by assessing differences between measurements taken at different timepoints. Thus, although accuracy is certainly important, it does not necessarily translate to the reproducibility of the technique and its sensitivity in detecting change. That is, if a technique is inaccurate or has a bias, as long as it is consistent in this bias, it may still be able to achieve acceptable results by way of being reproducible. The potential danger in this approach, however, is if this bias is not

consistent between the two measurements, in which case the results obtained will reflect the change in the bias, rather than the genuine differences one is trying to detect. A more precise measurement can be obtained when serial scans from an individual are accurately registered and the volume changes are derived directly [230]. Two examples of techniques that make use of this principle and are used for detecting whole brain atrophy are the brain boundary shift integral [231] and SIENA [232], both of which would be preferable to cross-sectional techniques for the reasons mentioned above.

Understanding the shortcomings of the various methodologies that are available is essential both to the end user who has to first select the appropriate methodology and then interpret the results, and to the development community who we hope will use the information presented in this paper as a springboard for newer, more accurate algorithms that address these biases and inaccuracies.

3.6 CONCLUSION

In summary, we evaluated the GM segmentations of several commonly used automated techniques for the detection of atrophy in MS. Results demonstrate that, although the algorithms perform similarly to manual segmentations of cortical GM, severe shortcomings exist in the segmentation of deep GM structures. Such misclassifications are of particular importance in studies on MS given that their magnitude can be more than four times the annual rate of atrophy. In general, given the specificity of tools being developed for image processing, one should consider the specific purpose, strengths, weaknesses, and appropriateness of the tool for a particular task.

3.7 ACKNOWLEDGMENTS

This research was supported in part by the NSERC/MITACS Industrial Postgraduate Scholarship Program and the endMS society of Canada. We thank the expert readers for their manual segmentations: Tao Li, David Araujo, Xu Liu, and Ling Han. We also thank Simon Warfield for his discussion and assistance with STAPLE. STAPLE is supported in part by NIH R01 RR021885 from the National Center for Research Resources and by an

award from the Neuroscience Blueprint I/C through R01 EB008015 from the National Institute of Biomedical Imaging and Bioengineering.

Chapter 4

A PROPOSED METHOD FOR IMAGING SUBPIAL DEMYELINATION

The previous chapter focused on a quantitative evaluation of methods used in the assessment of GM pathology in patients with MS. The manuscript presented in this chapter introduces a novel surface-based technique developed for the detection and quantification of cGM pathology, specifically subpial demyelination. As was presented in Chapter 2, a number of studies have been looking to fully describe this pathology using ultra-high-field imaging (i.e., ≥ 7 T), often combined with highly specific immunohistochemical staining. The primary aim of this manuscript was not to provide the same degree of insight into the disease pathology, but rather to assess the sensitivity and feasibility of the method to detect differences in the cGM of subjects with MS. Importantly, this was done using images typically obtained in a clinical trial setting, thus giving the methodology practical relevance.

Earlier forms of this work were published as conference proceedings at the 2009 and 2010 International Society for Magnetic Resonance in Medicine (ISMRM) conferences [354, 355]. On both occasions, the work was selected for oral presentation, alluding not only to the community's interest in the work but also to the recognition of its originality and relevance to the field. The current version of the manuscript has been submitted to *Human Brain Mapping* and is currently under review.

SURFACE-BASED ANALYSIS REVEALS REGIONS OF REDUCED CORTICAL MAGNETIZATION TRANSFER RATIO IN PATIENTS WITH MULTIPLE SCLEROSIS: A PROPOSED METHOD FOR IMAGING SUBPIAL DEMYELINATION

MISHKIN DERAKHSHAN[1], ZOGRAFOS CARAMANOS[1], SRIDAR NARAYANAN[1], DOUGLAS L. ARNOLD[1,2], AND D. LOUIS COLLINS[1]

[1]McConnell Brain Imaging Centre, Montreal Neurological Institute, Montreal, Canada H3A 2B4
[2]NeuroRx Research, 3575 Parc Avenue, Suite #5322, Montreal, Canada H2X 4B3

Human Brain Mapping (In Submission: HBM-12-0884)

4.1 ABSTRACT

The in vivo detection of subpial cortical grey matter lesions in multiple sclerosis is challenging. We quantified the spatial extent of subpial decreases in the magnetization transfer ratio (MTR) of cortical grey matter in subjects with multiple sclerosis, as such reductions may indicate regions of cortical demyelination.

We exploited the unique geometry of cortical lesions by using two-dimensional parametric surface models of the cortex instead of traditional three-dimensional voxel-wise analyses. MTR images were mapped onto intermediate surfaces between the pial and white matter surfaces and were used to compute differences between secondary-progressive MS (n = 12), relapsing-remitting MS (n = 12), and normal control (n = 12) groups as well as between each individual patient and the normal controls.

We identified large regions of significantly reduced cortical MTR in secondary-progressive patients when compared with normal controls. We also identified large regions of reduced cortical MTR in 11 individual patients (8 secondary-progressive, 3 relapsing-remitting).

The secondary-progressive patients showed larger areas of abnormally low MTR compared with relapsing-remitting patients both at the group level and on an individual basis. The spatial distributions of abnormal MTR preferentially involved cingulate

cortex, insula, and the depths of sulci, in agreement with pathological descriptions of subpial grey matter lesion distribution. These findings suggest that our method is a plausible in vivo imaging technique for quantifying subpial cortical demyelinating lesions in patients with multiple sclerosis and, furthermore, can be applied at the typical clinical field strength of 1.5 T.

4.2 INTRODUCTION

Although multiple sclerosis (MS) has traditionally been viewed as a disease of the white matter (WM), the involvement of grey matter (GM) was noted over a hundred years ago in some early pathological studies of chronic MS cases [27, 29]. In these case studies, researchers noted that demyelination affected the cortex at least as much as the WM and, in some cases, even more. Almost a century later, with advances in immunohistochemical staining techniques sensitive to myelin [32, 39, 330], the MS research community is only beginning to appreciate how extensive GM damage can be [32, 33, 59].

Several schemas have emerged for classifying cortical GM lesions in MS. For simplicity, we use the taxonomy defined by Trapp et al. [26] and enumerate three types of cortical lesions: Type I lesions are leukocortical areas of demyelination that involve both subcortical WM and cortex. Type II lesions are small, perivascular, and purely intracortical and do not significantly contribute to cortical lesion load. Type III lesions are characterized by subpial demyelination that follows the shape of the cortical mantle, often extending over multiple gyri and often stopping at cortical layer III or IV [26]. Although the exact distribution of each type of lesion has varied by study, subpial demyelinating lesions are the most prevalent, accounting for half or more of the cortical lesions seen on histology [33, 54] and up to 67% of the total demyelinated cortical area [33]. In addition, these lesions have been found to be more extensive in progressive cases of MS than in cases of relapsing-remitting (RR) MS [39].

Despite the prevalence of lesions in cortical GM (cGM) in MS, conventional magnetic resonance imaging (MRI) is not able to detect most cGM pathology. Newer imaging techniques such as double inversion recovery (DIR), in which the signal from both the WM and the cerebrospinal fluid (CSF) is suppressed, have increased the detection of cGM pathology compared with more conventional sequences [119], but this increase has been limited to juxtacortical and purely intracortical lesions (Types I and II) [119, 120, 191]. The difficulty with using conventional imaging to detect subpial cortical demyelination (Type III) is that these lesions do not appear to be associated with an appreciable influx of inflammatory cells or edema and consequently show little alteration of T2 or T1 relaxation times [32, 38, 54]. Increasing the signal-to-noise ratio has been purported to increase the sensitivity of conventional MRI to detect cGM pathology; however, as with the DIR sequences, this improvement was limited to juxtacortical and intracortical lesions and has not significantly helped in the detection of Type III subpial demyelinations[121, 125]. Scanning at ultra-high field (e.g., ≥ 7 T) has been able to identify Type III lesions ex vivo [52, 126, 136]; however, even at such high field strengths, these lesions may [127] or may not [124, 356] be visualized in vivo. In addition, the extremely limited access to such scanners, combined with technical challenges (i.e., B0 and B1 field inhomogeneities, higher energy deposition, etc.), makes ultra-high-field MRI for the detection of subpial demyelination clinically impractical at the present time.

Microscopic changes seen in Type III cortical lesions include microglial activation, axonal transection, and apoptotic neurons [32]; importantly, however, compared with WM lesions, these subpial cortical lesions in patients with long-standing MS show a significantly less pronounced inflammatory response (lack of T-cell and B-cell infiltration, microglial activation, and astrogliosis) [32, 38]. That is not to say that all subpial demyelinating lesions are devoid of inflammation as a recent study of patients with early MS found T-cell infiltrates and myelin-laden macrophages [41]; still, the predominant feature of the Type III cortical lesion is demyelination. This characteristic

suggests the potential usefulness of studying such lesions using magnetization transfer imaging, a methodology that, in the present context, has two very important strengths: first, as we will describe below, it is sensitive to myelin content; second, it can be easily acquired on standard clinical MRI scanners.

Magnetization transfer imaging, specifically, the magnetization transfer ratio (MTR), has been shown to be sensitive to changes in myelin content in WM [90-95], a marker of intrinsic GM damage [357, 358], and significant decreases and increases in MTR have been shown to be associated with de- and remyelination, respectively, of WM lesions on postmortem histology [90, 91]. Though MTR may not be a purely specific marker for myelin as it may be affected by axonal loss and, to a lesser degree, inflammation, it does appear to be weighted toward myelination [92, 95, 99], and in a recent postmortem 9.4 T imaging study of cortical lesions [52], MTR was strongly correlated to the intensity of myelin basic protein staining (r = .52, p = .02). Based on these findings, we sought to determine whether cortical MTR has the potential to provide a measure of subpial demyelination.

Macroscopically, subpial demyelinating lesions are characterized as extending over multiple gyri, following the shape of the cortical mantle [33], and preferentially affecting the outermost layers of the cortex [26, 39]. Therefore, to increase sensitivity, we exploited this unique geometry and performed our analyses on two-dimensional (2D) parametric surface models of the cortex instead of using the traditional three-dimensional (3D) voxel-wise analyses. These surface models were reconstructed at a subvoxel resolution that enabled the quantification of MRI signal(s) (in this case MTR) at given depth(s) of the cortex [217, 223]. Blurring along these surface models also avoids the problems associated with 3D voxel-wise blurring (e.g., the signal from the voxels of two separate gyri can be erroneously blurred together as they are considered adjoining in 3D space because of the particular folding pattern but, in reality, are not neighbours along the cortical surface). Indeed, the superior power and precision of surface-based techniques have been most notably demonstrated in cortical thickness analyses [228]. When

combined with the difficulties of 3D voxel-based GM segmentation methods [359], a surface-based approach seems intrinsically better suited to the analysis of the cerebral cortex or, more specifically in this case, subpial demyelination that follows the geometry of the cortical mantle.

As our appreciation for the importance of cGM pathology grows, reliable imaging methods are essential to accurately measure and analyze cGM pathology in MS. In this report, we present a novel surface-based method developed by our group [354, 355] to quantify the extent of subpial decreases in MTR that may correspond to regions of cortical demyelination. Our method was applied to group data from patients with MS and normal controls, as well as to data from individuals to identify areas that differed from the control group. Importantly, these differences were detected using conventional-resolution images such as those typically obtained in clinical practice.

4.3 METHODS

4.3.1 SUBJECTS

Magnetic resonance scans of healthy controls and patients with RR and secondary-progressive (SP) MS recruited from the MS Clinic at the Montreal Neurological Institute and Hospital for a previous study were analyzed in this retrospective analysis. Of the possible 26 SP MS patients in the original study, 14 were included based on availability of MTR data and adequate scan quality. The SP group was further limited to 12 subjects as we were unable to obtain accurate cortical reconstructions for 2 of the subjects. Further selection was then performed to match all groups for population size, sex, and mean age, given that age-related changes in the MTR of both the GM [360-362] and WM [361, 362] of normal controls (NC) have been demonstrated. Demographics for the 12 patients with SP MS, 12 patients with RR MS, and 12 NC are presented in Table 4.1. Patients were not being treated with disease-modifying drugs at the time of the scan.

Local ethics board approval and informed consent from each participant were obtained.

Table 4.1 – Subject information. All values are mean [range]. DD: Disease duration, EDSS: Kurtzke's Expanded Disability Status Scale, T2LV: manually identified white matter lesion volume on the T2-weighted scan, NBV: Normalized brain volume output using SIENAx, cThx: Global cortical thickness, NC: Normal Controls, SP: Secondary-Progressive, RR: Relapsing-Remitting.

	M/F	Age (yrs)	DD (yrs)	EDSS	NBV (cc)	T2LV (cc)	cThx (mm)
NC	4/8	44 [28-60]	-	-	1510.6 [1397.6-1685.6]	-	2.43 [2.27-2.56]
RR	4/8	45 [30-59]	13 [1-33]	2.5 [1.0-4.0]	1491.2 [1299.6-1661.2]	7.6 [0.8-27.3]	2.42 [2.29-2.53]
SP	4/8	47 [30-62]	14 [3-26]	5.7 [3.5-8.0]	1461.7 [1335.0-1581.7]	33.1 [4.0-80.1]	2.25 [1.89-2.44]

4.3.2 MRI ACQUISITION

Subjects were scanned on a 1.5 T Philips ACS II scanner (Philips Medical Systems, Best, the Netherlands). Oblique axial T1-weighted (T1w) images were acquired parallel to the antero-posterior commissural line using a 3D spoiled gradient-recalled echo sequence (TR = 35 ms, TE = 10 ms, 256×256 matrix, 250 mm field-of-view, 60 partitions, 3 mm partition thickness, 1 signal average, voxel size = 0.98x0.98x3.0 mm^3). Images were acquired without and with a 1.2 ms on-resonance, bipolar (1-2'-1) magnetization transfer pulse (20 μT RF field strength) placed just before each slice-selective excitation. To calculate the MTR for each patient, the magnetization transfer image volume acquired with the saturation pulse (Sat) was first linearly registered (mritoself, McConnell Brain Imaging Centre [363]) to the volume without the saturation pulse (NoSat), and the MTR image volume was then calculated as 100 × (NoSat – Sat)/NoSat. In order to remove outliers caused by noise and possible data discretization errors, the MTR image was clamped so that all values were within the theoretical limits of 0 and 100.

4.3.3 MRI PROCESSING

Cortical reconstruction was performed using the T1w image volume without the magnetization saturation pulse (NoSat) as input to the FreeSurfer image analysis suite (v4.0.5, http://surfer.nmr.mgh.harvard.edu) [217, 223]. We visually checked the cortical reconstruction of all patients and controls, and manual corrections (inclusion of WM control points, corrections for WM lesions including Type I juxtacortical lesions, and pial

surface edits) were performed as necessary to ensure accurate surfaces. As mentioned above, two patients were excluded due to inaccurate surfaces that could not be corrected. From the generated pial and WM surfaces, intermediate surfaces were created by travelling 25%, 50%, or 75% of the distance along the vector linking a vertex on the WM surface to its corresponding vertex on the pial surface. See Figure 4.1 for examples of the surface extractions.

The information from the MTR image volume was interpolated onto every vertex on each surface. As Lerch et al. demonstrated that 2D geodesic blurring yields less error and bias than conventional 3D blurring for surface analyses [228], the MTR data was blurred along the surface with a 10 mm full-width-at-half-maximum geodesic kernel. The kernel size was chosen to preferentially detect the large subpial Type III cortical lesions over the small, punctate Type II lesions [26, 37, 54].

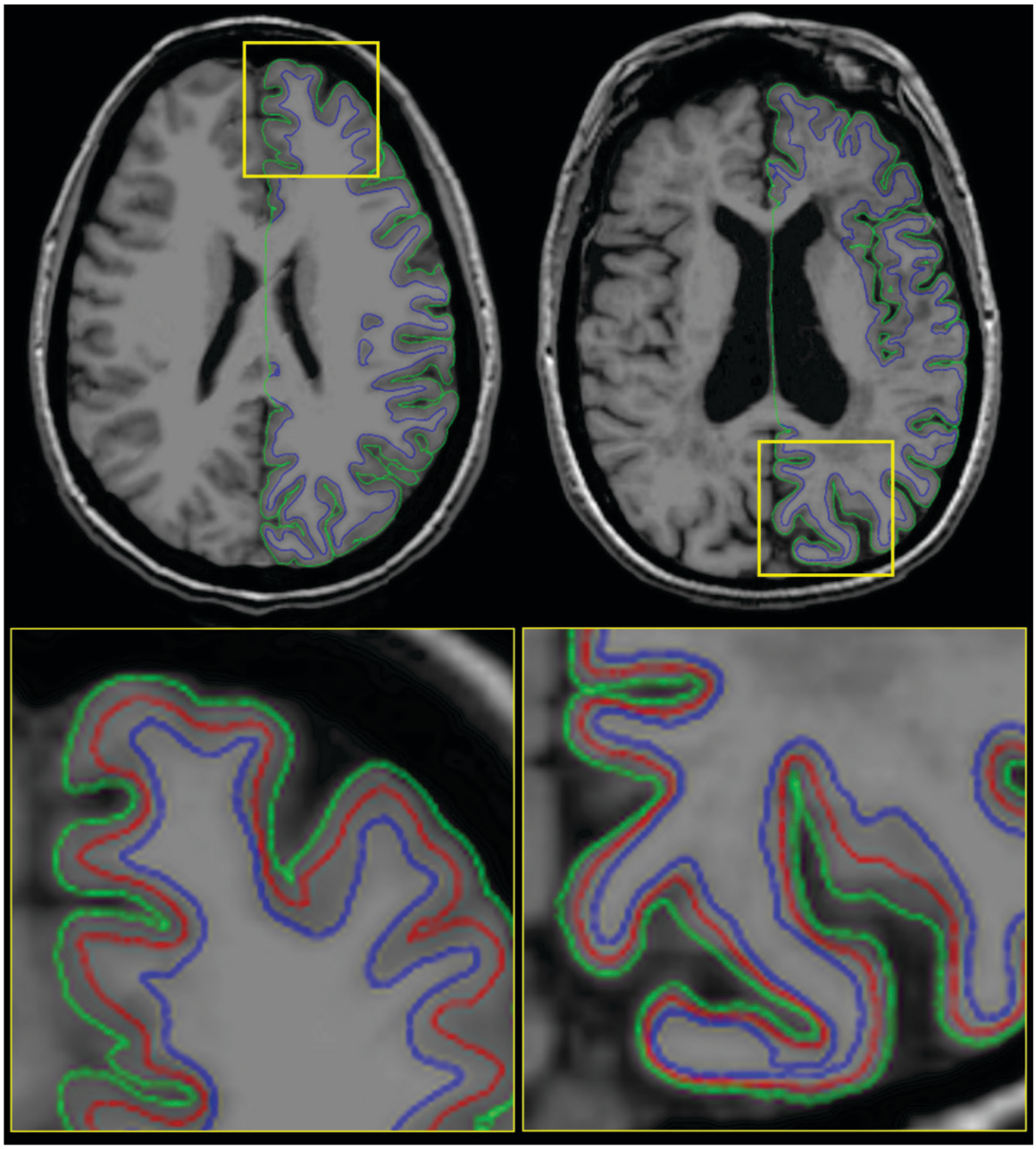

Figure 4.1 – Examples of the cortical surface extractions for a normal control (left) and an atrophic MS patient (right). Top: axial view showing the extractions of the right hemisphere; bottom: zoomed-in axial view of the yellow squares showing the pial (green), white matter (blue), and the intermediate surfaces at 50% depth (red). Subjects were selected at random as an accurate representation of the data, and not just the best possible slice of the best possible subject.

4.3.4 STATISTICAL ANALYSIS

To facilitate group comparisons, the cortical models were nonlinearly registered to a spherical atlas that uses individual cortical folding patterns to match cortical geometry across subjects [224], after which a general linear model (GLM) was performed to test for MTR differences between groups. The GLM was run at each vertex of each hemisphere, producing *t*-statistics that were then thresholded for significance and corrected for multiple comparisons using a false discovery rate (FDR) of 0.05. A mask of the cortex (lh.cortex.label and rh.cortex.label provided by FreeSurfer) was used to ensure that the FDR was not erroneously influenced by the noncortical, central regions of the brain (e.g., corpus callosum, third ventricle, and diencephalon) that are by default included on the medial surface of each hemisphere (see the noncoloured region in the bottom three views of Figure 4.5A). Statistical analyses were run using the mni.cortical.statistics library (courtesy of Jason Lerch) for R[364].

The above procedure was followed for each group comparison (SP vs. NC, RR vs. NC) and for each individual patient compared with the group of NC (e.g., Subject1 vs. NC, Subject2 vs. NC, etc.).

4.3.5 SENSITIVITY ANALYSIS

In order to assess the sensitivity of our proposed method to cortical MTR changes, we calculated the amount of MTR decrease that would be necessary to be detected using two methods: (1) direct calculation of a minimum detectable difference map and (2) simulated decreases in MTR.

Sensitivity analysis – Direct calculation

As indicated in the statistical analysis section above, the magnitude of the MTR difference at each vertex is tested for significance using a standard two-tailed *t*-test. However, one could directly calculate the minimum MTR difference required at each vertex to be deemed significant by rearranging the formula for a *t*-test, given that we know the cortical MTR values for each control subject and can look up the significant *t*-

ratio threshold in a *t*-table (using our known degrees of freedom). Although this method provides us with a minimum detectable difference map, we are unable to correct for multiple comparisons via FDR using this model because the *t*-ratio is the same for each vertex; FDR correction requires a variance in the to-be-corrected *t*-ratios. To overcome this limitation and obtain a more accurate map of the minimum detectable differences, we used simulated data (described below) that took FDR correction into account.

Sensitivity analysis – Simulations

Six datasets were created with a simulated decrease in MTR as follows: First, the average MTR value at each vertex, along each intermediate surface, was calculated for the group of 12 NC. These MTR values were then uniformly decreased by 1 to 6 MTR percentage units (p.u.) below the average value to create six new simulated datasets. For example, the intermediate surface (50% depth) of simulated dataset number 1 would be the same as the average of the 12 NC's intermediate surface MTR values, decreased by 1 MTR p.u. at every vertex, dataset 2 would have each value decreased by 2 MTR p.u., and so forth. Each of the six simulated datasets was then treated as an individual subject and run though the identical surface-based analysis method described above (surface-based registration, blurring, and statistical analysis), where each vertex of the decreased MTR surfaces was compared to the group of 12 NC and FDR corrected.

4.4 RESULTS

4.4.1 GROUP DIFFERENCES

Figure 4.2 shows the average cortical MTR maps for each group, with clearly evident decreases from the NC to the RR and SP groups. The groupwise mean MTR value in percentage units and the groupwise mean of the standard deviations over the entire intermediate surface were 25.49 (1.41), 24.24 (2.60), and 23.80 (2.29) for the NC, RR, and SP groups, respectively. The mean MTR value for each group differed significantly from those for other groups ($p < .0001$), as evidenced by a Tukey-Kramer HSD test.

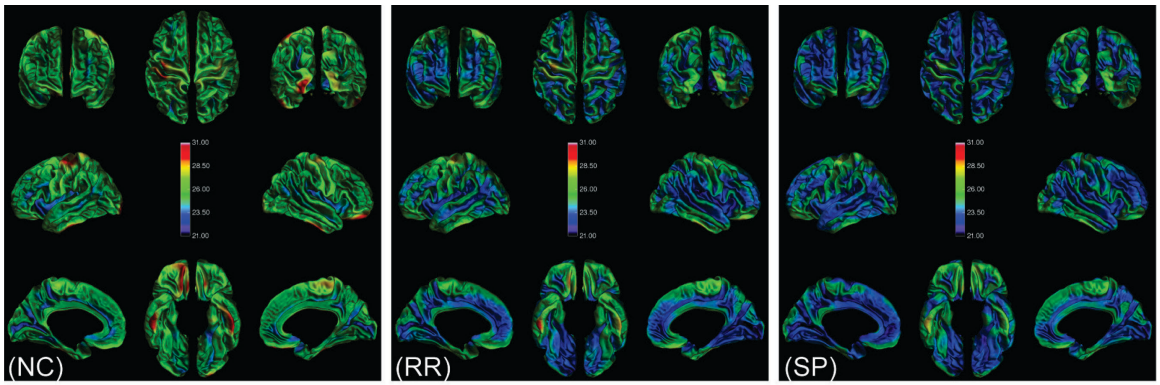

Figure 4.2 – Average (n = 12) maps of cortical magnetization transfer ratio (MTR) for each group. From left to right, there is a visually discernible decline in MTR going from healthy normal controls (NC) to the relapsing-remitting (RR) and secondary-progressive (SP) groups. The colour bars in each panel show the identical windowing used for displaypurposes going from 21 to 31 MTR percentage units (p.u.).The eight views presented in each panel starting from the upper left and going clockwise are: front, top, back, right hemisphere temporal, right hemisphere medial, bottom, left hemisphere medial, and left hemisphere temporal.

Also observed in Figure 4.2 are regions where the MTR appeared higher than that of its surroundings. This is best illustrated in the NC group where the occipital pole, motor cortex, and inferior surface of the brain appear brighter, but it can also be seen to a lesser extent in the RR and SP groups. The observation of relatively high MTR in occipital and motor cortex is consistent with the known increased density of myelin in these regions [365].

The highlighted areas in Figure 4.3 map the spatial distribution of significantly low cortical MTR (after FDR correction) in the SP group compared with the NC group. The SP group shows large contiguous areas of low MTR covering approximately 34% of the cortex. The RR group as a whole showed no such areas, likely because of more sparsely distributed areas of low focal cortical MTR [39] and spatial variability in their location across individuals, rather than as an indication that all RR patients have normal cortical MTR.

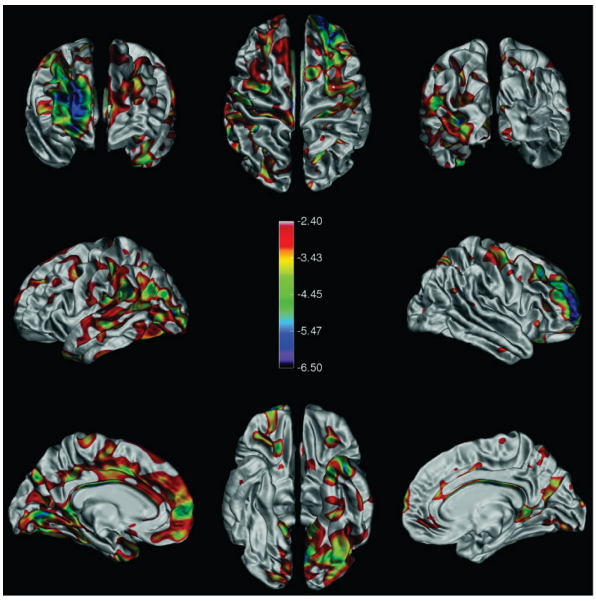

Figure 4.3 – Significant differences in cortical magnetization transfer ratio (MTR) between the secondary-progressive (SP) and normal control (NC) groups. The colour bar shows the false-discovery-rate-corrected *t*-statistics, such that any highlighted area on the surface represents a statistically significant decrease in MTR between groups. The relapsing-remitting (RR) cohort showed no such significant differences.

4.4.2 Individual Differences

Each of the 24 subjects with MS (12 RR and 12 SP) was compared with the NC group to reveal differences at the individual level. Figure 4.4 shows the 8 SP and 3 RR subjects for whom a localized significant decrease in MTR was detected, as well as a more detailed view of the median subject in each cohort in terms of area of decreased MTR detected.

4.4.3 Sensitivity analysis – Direct calculation

The map showing the minimum MTR difference required at each vertex in order to detect a significant change on an individual basis is shown in Figure 4.5A. The mean (standard deviation) decrease in cortical MTR p.u. was 2.63 (0.84).

4.4.4 Sensitivity analysis – Simulations

The first three simulated subjects with a global decrease of 3 or fewer MTR p.u. did not show any vertices being detected as having significantly low MTR after FDR correction. A global decrease of 4 p.u. resulted in 74.2% of the vertices being detected as having low MTR and is shown in Figure 4.5B. Further decreases of 5 p.u. and 6 p.u. resulted in 94.0% and 98.1% of vertices being detected, respectively.

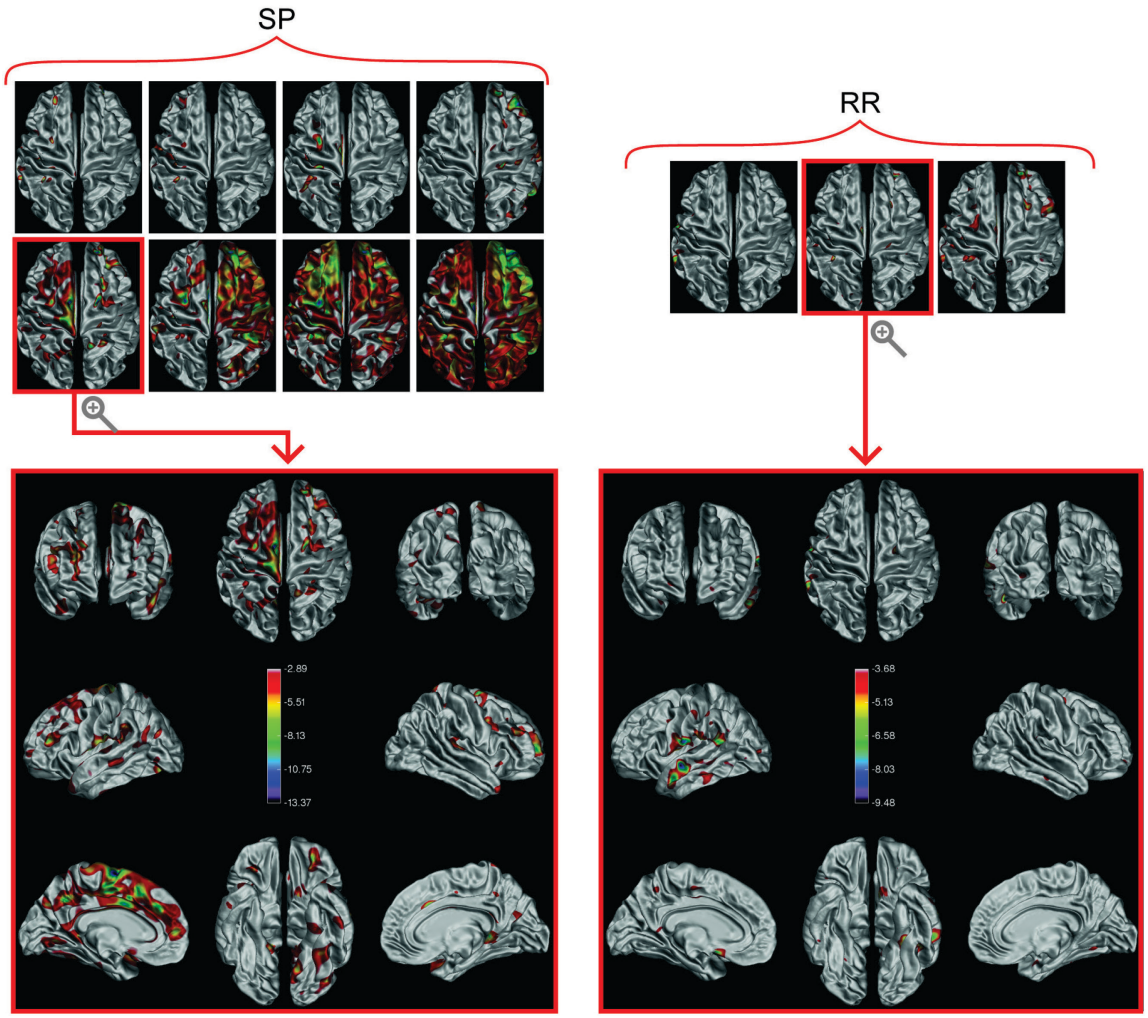

Figure 4.4 – Significant differences in cortical magnetization transfer ratio (MTR) between individuals and the normal control (NC) group. On the left is a montage of the top views of the cortical surfaces of the 8/12 individual secondary-progressive (SP) subjects for whom a significant difference from the NC group was detected. The 3/12 individual relapsing-remitting (RR) subjects are shown on the right. The median subject (in terms of affected surface area) is presented with all surface views visible below each respective montage. The colour bar shows the false-discovery-rate-corrected *t*-statistics, such that any highlighted area on the surface represents a statistically significant decrease in MTR between the individual and the NC group. Notice the range of affected surface area detected progressing from small clusters to almost complete coverage in the SP cohort.

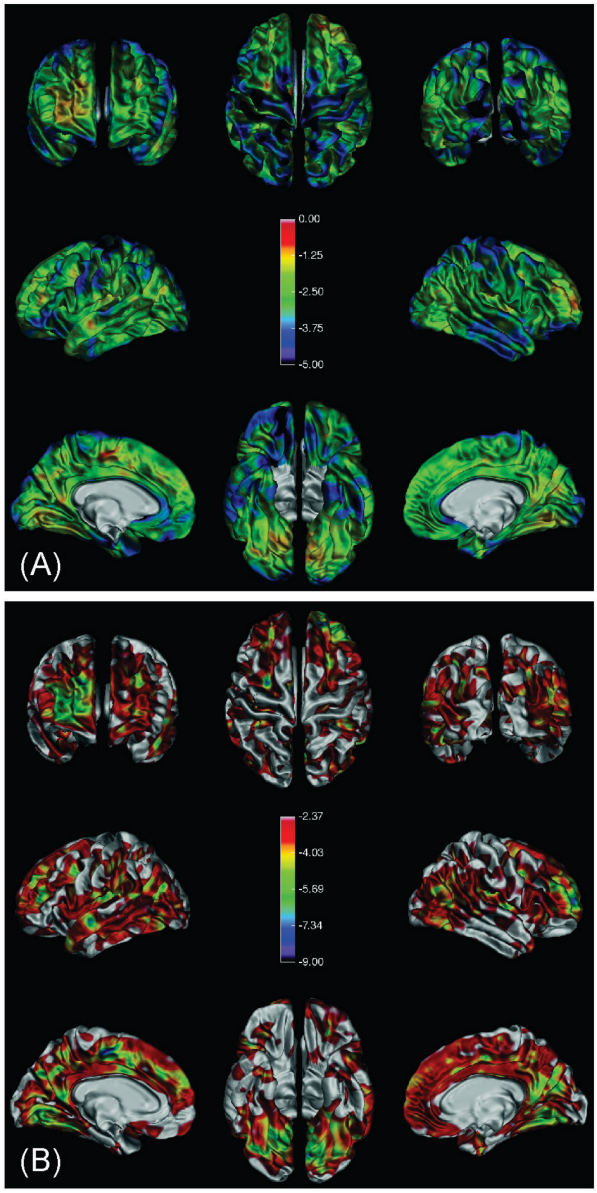

Figure 4.5 – (A) The spatial distribution of the calculated theoretical decrease in magnetization transfer ratio(MTR) percentage units (p.u.) needed at each point on the surface in order to detect a significant difference (uncorrected) from the normal control (NC) group. The colour bar ranges from 0 MTR p.u. (in red) to -5 MTR p.u. (in purple).(B) The spatial distribution of significant differences (false-discovery-rate-corrected *t*-ratios) detected using a simulated subject dataset for which each point on the surface was 4 MTR p.u. lower than the average NC value at that vertex. Note that the highlighted areas represent a 74.2% detection of the surface area. Simulations with global decreases of 3 MTR p.u. or less resulted in no significant detections (not shown), whereas increasing the simulated decrease to 6 MTR p.u. gave almost complete (98.1%) detection.

4.5 DISCUSSION

The complex geometry and thinness of the cortex induce variable amounts of partial volume at the cortical boundaries, and the pathology we are trying to detect (subpial demyelination) is often thinner than the largest voxel dimension of our MTR images. These factors severely limit the sensitivity of traditional 3D voxel-based methods to localized changes in MTR. By using a 2D surface that follows the cortical mantle, we have developed a new analytical method that is better suited to the task of detecting regional MTR abnormalities in the cortex that are of the shape expected for subpial demyelinating lesions. In addition to matching the profile of these lesions, our surface-based method has the advantage of not diluting our signal of interest. In a traditional 3D voxel-wise approach, the focal MTR decrease elicited by Type III cortical lesions would be diluted by information perpendicular to the plane of interest at the blurring stage of the analysis. Conversely, in our surface-based blurring scheme, we use only the information from adjacent vertices at the same depth of cortex. Although we present this technique in the context of detecting subpial demyelination using MTR images, the method can be applied to any imaging modality (e.g., T2-relaxometry data, double-inversion recovery, etc.) to look for aberrances in the cortex of an individual (or population) and possible correlations to cognitive or behavioural data.

Understanding cGM pathology in MS is of great interest to the field, and a number of studies have been looking to fully describe this pathology using ultra-high-field imaging, often combined with highly specific immunohistochemical staining. The primary aim of this study was not to provide the same degree of insight into the disease pathology, but rather to assess the sensitivity and feasibility of a new technique to detect differences in the cGM of subjects with MS, and to do so using images easily acquired on standard clinical MRI scanners, which—once validated with postmortem data—would make this approach feasible for use as a measure of subpial demyelination in clinical trials of patients with MS—trials that will most likely still be acquiring data on 1.5 T to 3.0 T scanners in the foreseeable future.

4.5.1 GROUP DIFFERENCES

To compare the MS patients with the NC group, we assessed the spatial distribution of the areas of decreased cGM MTR by thresholding for significant differences using a false discovery rate of 0.05. Compared with the NC group, only the SP group showed large, contiguous areas of significantly decreased MTR. The highlighted areas in Figure 4.3 indicate the regions where the decreases in cGM MTR over the entire group were large enough to be considered significant when compared with the NC group. That the RR group did not show large group differences in regional MTR does not mean that there were no regional decreases in individuals, but rather that the extent and spatial distribution of any such decreases were not consistent enough between RR subjects to be detected across the entire group.

As can be seen in Table 1, the RR cohort in this study is somewhat atypical. By design, the RR subjects were matched for age to the SP and NC groups to control for the age-related changes that have been seen in the MTR of the NC group [360-362]. Without age matching, the detected differences in MTR may simply have been the result of normal aging. Consequently, in this study, we presented an RR group with an uncharacteristically elevated age, as well as a mean disease duration that does not differ from the SP group as would be expected. While this limits the generalization of the disease-related findings, it does not take away from the primary aim of this study.

4.5.2 INDIVIDUAL DIFFERENCES

When individual subjects were compared with the NC group, 8/12 SP patients and 3/12 RR patients showed large, contiguous areas of decreased cortical MTR, with spatial distributions consistent with the Type III subpial GM lesions described pathologically [33]. An examination of the median subject (in terms of affected area) in each group indicated a much larger affected area in the median SP subject than in the median RR subject. This higher frequency of patients with a detectable difference in the SP group than in the RR group, and generally greater extent of the differences seen, is consistent

with pathological observations that the later stages of MS are characterized by an increase in cortical subpial demyelinating lesions [366].

4.5.3 Cortical Layers

In addition to carrying out our analysis on the intermediate (50%) surface, we also looked at parallel surfaces at differing depths (25% and 75%) between the pial and WM surfaces. We observed that, qualitatively, in both group comparisons and in those individuals who displayed large areas of decreased cGM MTR, the amount of affected surface area increased closer to the pial surface and decreased toward the WM surface. These observations are consistent with what has been seen on postmortem histology, namely, that subpial demyelination preferentially affects the outermost layers of the cortex, often stopping at cytoarchitectonic layer III or IV [26, 39].

4.5.4 Sensitivity Analysis

In order to assess the sensitivity of our technique, we generated the theoretical map of minimum detectable differences shown in Figure 4.5A. The map varies spatially and reflects the variability of MTR in the NC group (i.e., changes in MTR were most easily detected in areas of the cortex where there was little variation in the MTR of our control group). This method provides the MTR difference needed at each vertex to achieve significance; however, it does not correct for multiple comparisons, as is actually done in this sort of analysis. The simulations shown in Figure 4.5B do correct for multiple comparisons and also show that, with this technique, acquisition protocol, and number of controls, the minimum detectable decrease of MTR is somewhere between 3 and 4 p.u. We expect that higher resolution acquisitions would reduce variability due to partial volume effects and increase sensitivity (i.e., reducing the minimum MTR change that can be detected).

4.5.5 VOXEL GEOMETRY

Because our study aimed to analyze data that is typically used in clinical research, we did not employ high field strengths or long scanning times to obtain high-resolution images. The T1w images input into FreeSurfer were 0.98x0.98x3.0 mm^3, which is indeed suboptimal in comparison to the 1 mm isotropic recommended resolution; however, we feel that we were able to obtain precise surface extractions (see Figure 4.1 for examples) for the following reasons: (i) FreeSurfer relies heavily on WM and GM contrast to identify brain surfaces;(ii) our T1w sequences were suited to providing this contrast; and (iii) each surface was carefully visually examined and corrected as needed. The effects of voxel size and anisotropy on the reliability of FreeSurfer have been previously investigated [367], and the reliability of the morphometric measures has been reported to be generally high and largely unaffected by differences in voxel geometry (reliability was also largely unaffected by parallel acceleration factors or the use of high-bandwidth multi-echo techniques). Larger, anisotropic voxel sizes did, however, result in a significant measurement bias—the cortical thickness was higher than that from a sequence with smaller isotropic voxels—to which our study may be subject. However, all the subjects in our study were acquired with the same sequence parameters and therefore all experienced the same amount of said bias; thus, tests looking for differences between groups or individuals would not be affected.

4.5.6 BLURRING KERNELS

We investigated the effect of 2D versus 3D blurring by examining how the type and size of blurring kernel affected the variance in the cortical MTR of our cohorts (Figure 4.6). The 3D analysis was done by obtaining the median MTR values of each subject within a mask of the cerebral cortex (provided by the volumetric stream of FreeSurfer), while the 2D analysis was performed as previously described—by obtaining the mean MTR value along the intermediate surface and using surface-based blurring. For each method, and for each of the 0, 2, 4, 6, 8 and 10 mm kernel sizes tried, an ANOVA was performed to reveal how much of the variance in cGM MTR values was explained by the three cohorts (i.e., NC, RR, SP) (higher r^2 means the groups are better separated). We found that the r^2 value remained high and the p-value remained low regardless of the blurring kernel used for the 2D analysis, but that for the 3D analysis, there was much variation depending on the blurring kernel size. Intuitively, this makes sense because the median cGM MTR value in the 3D case was altered by signal outside the cortex, whereas in the 2D scenario, the mean values remained consistent, as one would expect, because the cGM signal was blurred by signal within the cortex. Also worth noting is that, for all the 2D analysis blurring kernels (0-10 mm), a Tukey-Kramer HSD test revealed that the normal control cGM MTR values were significantly different from those for the RR and SP groups, but for the 3D analysis method, this was only the case for smaller kernels (2-8 mm). (There was actually no separation of the groups where there was no blurring (0 mm has $p > 0.05$).) In the 3D analysis, the r^2 value did plateau to a value slightly higher than that in the 2D analysis for the larger blurring kernels, but as the kernel size is increased to 10 mm and higher, only the SP group became significantly different from the NC group. The dependence of the 3D analysis on technicalities such as the blurring kernel size—which varies immensely study by study—combined with the consistent result of the 2D analysis, suggests that the 2D analysis is better suited to investigating abnormalities within the cortex.

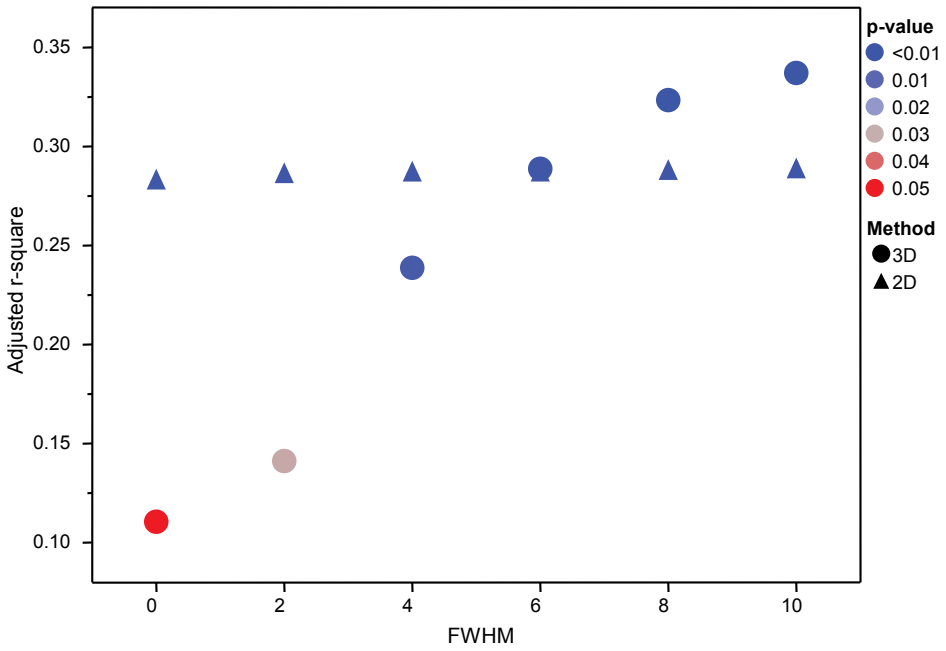

Figure 4.6 – Adjusted r^2 values at each of the full-width-at-half-maximum (FWHM) blurring kernel sizes for a 2D and a 3D analysis looking for group differences in the cGM MTR values. The r^2 of the ANOVA reveals how much of the variance in cGM MTR values is explained by the three groups (NC, RR, SP) (higher r^2 means the groups are better separated), while the colour of each datapoint represents the associated *p*-value.

4.5.7 Partial Volume Effects

As with any finite sampling MRI acquisition, we are burdened by partial volume effects whereby the signal represented in a single voxel may consist of contributions from more than one type of tissue. In particular, the cGM voxels in the MTR image will be a mixture of pure GM tissue and the CSF and WM tissues that abut the cortex. This is of concern because subjects with more brain atrophy have thinner cortices and larger sulci; thus, the cGM voxels in these subjects are more likely to have some CSF contamination that could lower MTR values. Unfortunately, the quantitative effects of partial volume on cortical surface MTR is not well described; however, by looking at the cortical thickness and MTR in the cortex of a normal control subject, we can see if more partial volume is associated with a lower MTR value, as would be evidenced by an expected positive correlation. We found no such correlation ($r^2 < .01$), and thus we are confident that the subpial decreases in MTR we detect with this method are biologically driven and not artifactual.

4.6 CONCLUSIONS

In this study, we have presented a novel surface-based method for the detection of local subpial decreases in MTR. This new approach represents an important advancement for the detection of putative Type III cortical lesions in vivo, as conventional methods to date have had only limited success. Our method exploits what is known about both the spatial organization and the microscopic composition of these lesions and produces results that are very consistent with postmortem pathological studies. Our method identifies focal, cortical pathology that increases in the later stages of disease, affects the outermost layers of the cortex more than the deepest layers, and preferentially involves the same regions that have been identified to be preferentially involved in histopathological studies (e.g., cingulate cortex, insula, and the depths of sulci) [33, 60, 368]. Importantly, our method seems to allow us to do this using data that can be acquired with conventional clinical scanners, which makes this approach feasible for use in clinical trials interested in measuring cortical pathology in patients with MS. For all of these reasons, we believe that surface-based MTR analysis is an extremely promising approach for the detection of Type III cortical demyelinating lesions in vivo that merits further development and validation.

4.7 ACKNOWLEDGEMENTS

This work was supported by grants from the Canadian Institutes of Health Research (CIHR, MOP-84367 and MOP-84360) and from the Multiple Sclerosis Society of Canada (MSSC).

Chapter 5

CORTICAL NEUROIMAGING CORRELATES OF COGNITIVE DYSFUNCTION

This chapter builds on the findings presented in Chapter 3 and the method for detecting focal cortical damage introduced in Chapter 4: We investigated the most appropriate ways of measuring focal, regional, and global cortical pathology in patients with MS and, more importantly, explored the relationship between the imaging findings and the subjects' cognitive performance.

The rationale behind the primary hypotheses explored in this manuscript resulted from a preliminary study of cortical thickness correlates of cognitive function that was previously published as a conference proceeding at the 2011 American Academy of Neurology conference [369]. The current version of the manuscript is in preparation for submission to *NeuroImage*.

WIDESPREAD CORTICAL PATHOLOGY MEASURED WITH MAGNETIZATION TRANSFER RATIO BEST PREDICTS COGNITIVE IMPAIRMENT IN PATIENTS WITH MULTIPLE SCLEROSIS

MISHKIN DERAKHSHAN[1], GABRIEL LEONARD[1], SRIDAR NARAYANAN[1], ZOGRAFOS CARAMANOS[1], D. LOUIS COLLINS[1] AND DOUGLAS L. ARNOLD[1,2],

[1]McConnell Brain Imaging Centre, Montreal Neurological Institute and Hospital, Montreal, Canada, H3A 2B4
[2]NeuroRx Research, 3575 Parc Avenue, Suite #5322, Montreal, Canada H2X 4B3

NeuroImage (in preparation)

5.1 ABSTRACT

As the appreciation grows for cortical grey matter pathology in patients with multiple sclerosis (MS), so does the need for imaging methods that capture said pathology and for understanding the relationship of those methods to the clinical outcome of the patient. In this paper, we investigate focal, regional, and global measures of cortical magnetization transfer ratio (MTR) as well as cortical thickness, and model their relationship to a measure of processing speed from the Wechsler Adult Intelligence Scale (3rd edition), the Tower of London task, the Paced Auditory Serial Addition Test (PASAT), and two specific hand tapping tasks. Results demonstrated that, when controlling for global effects, widespread pathology captured with a global measure of cortical MTR is best for predicting cognitive abilities. As a secondary aim, the study also evaluated the ability (effect size) of a number of other magnetic resonance imaging metrics and neuropsychological tests to discriminate between normal controls and cognitively impaired patients with MS.

5.2 INTRODUCTION

Multiple sclerosis (MS), an immune-mediated disease of the human central nervous system, is characterized pathologically by varying degrees of inflammation, demyelination, axonal and neuronal degeneration, and gliosis. The clinical manifestations also vary immensely between individuals, with cognitive impairment (CI), a disabling and poorly managed consequence of MS, being a symptom present in

roughly half of patients with MS (varying from 35% to 65% depending on the specific cohort studied [247, 269, 370-372]). The seminal papers of the early 1990s by Rao and colleagues [269, 286] underscored CI as an important contributor to disability in patients with MS and, with the presentation of their proposed Brief Repeatable Neuropsychological Battery of tests (BRNB), have been, for better or worse, pivotally influential in nearly all the subsequent studies on CI in patients with MS.

The potential domains for CI in patients with MS are understandably as heterogeneous as the disease pathology itself. Still, one of the core deficits reported is the slowing of information processing speed [276, 278, 373-376], with other key domains including impaired memory, decreased visual/spatial processing, and impaired (higher-order) executive function [301]. As cognitive ability is not readily observable in routine neurological examinations, a wide range of standardized neuropsychological tests have been used to provide an assessment of domain-specific and general CI. The Wechsler Adult Intelligence Scale (WAIS-III) [377] is one such well-established neuropsychological test that provides a measure of global intelligence as well as subindices of verbal and nonverbal abilities (with the latter including an index of processing speed). Another test, and perhaps one of the most used in MS cognitive research, is the Paced Auditory Serial Addition Test (PASAT) [295, 378], which measures information processing speed as well as memory functions [376] and is reported to be both reliable and sensitive to the early detection of CI in patients with MS [379-381]. Higher-order problem-solving and executive planning abilities can be assessed with the Tower of London (TOL) [382, 383], and although it is not traditionally used in MS research, the TOL has been widely studied and found to be a robust assessment of executive function in the study of other neurological disorders (e.g., Parkinson's, schizophrenia, epilepsy, and mild cognitive impairment) [384-390].

These tests and others have become the de facto gold standard for assessing CI in patients with MS, though there has been growing interest in identifying markers on magnetic resonance (MR) imaging that can be used to quantify CI. Indeed, in patients

with MS, an association between CI and white matter (WM) lesion burden is well established [302, 391], though better correlations have been achieved by measuring whole brain atrophy [302-306] or the atrophy of specific brain regions such as the deep grey matter (GM) [306-308], corpus callosum [309, 310], cerebral cortex [246, 247, 311], mesial temporal lobes [312, 313], and other subregions of the cerebral cortex [254, 314]. In fact, so many different MR measures of global and/or regional damage have been correlated with CI in patients with MS that, unfortunately, the literature has become diluted and difficult to navigate. Particularly challenging to interpret are studies that look for correlations between a cognitive index and a multitude of regions without correcting for multiple comparisons or controlling for a global effect in order to assess whether the observed regional damage is not simply driven by more widespread pathology.

In this cross-sectional study, we examined the relationship between several specific neuropsychological tests and advanced MR imaging markers. Our primary hypothesis was that frontal lobe pathology as assessed by cortical thickness and cortical magnetization transfer ratio (MTR) measurements would predict cognitive deficits visible on the WAIS-III subindex of processing speed, as well as on the TOL and PASAT. Moreover, we hypothesized that, for the more specialized tapping tasks that were included in the battery, regional motor or premotor cortical pathology would best predict performance. As a secondary aim of the paper, we sought to evaluate the effect size of the various MR imaging markers and neuropsychological tests, in the hopes that the results of this pilot study could be used to identify the metrics that best characterize CI in patients with MS.

5.3 METHODS

5.3.1 SUBJECTS

A total of 26 patients with clinically definite MS were recruited from the MS clinic at the Montreal Neurological Institute and Hospital. To be included in the study, patients had to display some degree of cognitive disability as indicated by a score of one or greater on the cerebral functional system subscale of Kurtzke's Expanded Disability Status Scale (EDSS). Patients about to start a new disease modifying therapy were excluded from the study. Fifteen healthy volunteers were then recruited to serve as the normal control (NC) group, and were selected to be matched on age, gender and education level (Table 5.1). Parking costs were covered for all participants, and the healthy volunteers received $50 compensation for their participation. The written consent of each participant was obtained along with local ethics board approval for the study.

Table 5.1 – Subject demographics.

	NC (n = 15)	MS (n = 26)	*F* (df)	*p*
Age	38.73 ± 11.37	44.11 ± 10.49	2.35 (1,39)	.13
Years of Education	15.73 ± 3.13	14.88 ± 3.22	0.68 (1,39)	.42
Gender (F/M)	19/7	10/5	-	.67

Data are presented as mean ± standard deviation unless otherwise indicated. Data analyzed using ANOVA or chi-squared tests as appropriate. NC = normal controls; MS = patients with multiple sclerosis.

5.3.2 MRI ACQUISITION

All participants were scanned on a 3.0 T MAGNETOM Trio TIM scanner (Siemens, Erlangen, Germany). A 3D fast low-angle shot (FLASH) gradient-recalled echo (GRE) sequence [TR=20ms, TE=5ms, FA=27°, 256x256 matrix, 256mm field of view, 1mm slice thickness, 1 signal average, voxel size= 1x1x1mm^3] provided 1mm isotropic T1-weighted (T1w) anatomical MR images. A turbo spin-echo sequence [TR = 2,100 ms, TE = 16 ms and 80 ms, voxel size = 1x1x3 mm^3] provided T2-weighted (T2w) and PD-weighted (PDw) sequences. A pair of 3D FLASH sequences [TR = 33ms, TE = 3.81ms, FA = 10°], with geometry matched to that of the high-resolution T1w sequence, was acquired with

and without an off-resonance magnetization transfer pulse [Gaussian pulse, offset frequency=1,200 Hz, FA = 500°, pulse length = 9,984 μs, bandwidth = 192 Hz]; MTR maps were created from these images after co-registration as voxel-wise percent difference maps.

5.3.3 NEUROPSYCHOLOGICAL TESTING

All participants completed an extensive neuropsychological assessment within an average of nine days of their MR examination (median = two days). Testing was administered in English or French according to the preferred language of the participant. In order to avoid fatigue the testing was split between two sessions lasting approximately two hours each and on average seven days apart (median = two days apart). Testing was administered by trained psychometricians or neuropsychologists under the supervision of a senior neuropsychologist (GL), and the examiners were blinded to the MR findings and not told in advance about the clinical status of the participant. The battery was selected to measure a wide range of cognitive abilities and included the WAIS-III [377], Wechsler Memory Scale (third edition; WMS-III) [392], PASAT [295, 378], Grooved Pegboard Test (GPT; Lafayette Instrument, Lafayette, USA), Woodcock-Johnson III math fluency exam (MATH) [393], TOL Drexel University edition [383], Wisconsin Card Sorting Test (WCST) [394], Boston Naming Test (BNT) [395], Delis-Kaplan Executive Function System colour word interference test and trail making test (D-KEFS) [396], Leonard Tapping Tasks (LTT) [397], and the Controlled Oral Word Association Test with letters FAS (COWAT). The testing order and the breakdown per testing session is presented as supplementary data in section 5.7.

The WAIS-III [377] provides an assessment of global cognitive functioning yielding a Full-Scale Intelligence rating (FSIQ), which is the composite of verbal (VIQ) and nonverbal (PIQ) abilities. Both the VIQ and PIQ comprise subindices, and of particular interest in our study is the Processing Speed Index, a subindex of the PIQ that comprises the Digit-Symbol Coding and Symbol-Search subtests. Briefly, in the Symbol-Search subtest, participants are first shown two symbols, and then asked whether either of them

is present amongst a series of five adjoining symbols. In the Digit-Symbol Coding test, participants are first shown a digit-symbol key, and are then asked to transcribe the symbol associated with each digit as rapidly as possible for a total of two minutes. Both subtests assess processing speed and are very similar to the frequently used Symbol-Digit Modalities Test (SDMT), which is an oral version of the digit-symbol coding test [294].

Auditory information processing speed was also assessed with the Rao adaptations of the PASAT. Subjects were required to monitor audiotaped digits played at 2 and 3 second intervals while adding each consecutive digit to the preceding one during a continuous stream of 60 stimuli.

Executive function and conceptual reasoning were assessed using the well-described WCST. We reported the number of categories achieved, total number of errors, the total number of perseverative errors, and the total time to complete the task. Higher-order problem-solving and executive planning abilities were assessed with the three-peg Drexel University edition of the TOL. We recorded the total number of correct moves, the total number of moves, as well as the total problem solving time.

A number of language and verbal fluency tests were also administered including the D-KEFS [396], BNT [395], and the FAS phonemic verbal fluency test (COWAT). The D-KEFS included the trail making tests (i.e., visual scanning, number sequencing, letter sequencing, number-letter switching, and motor speed) and the STROOP tests, where we recorded the total time to complete for the color-naming, word reading, and inhibition tasks. The total phonemic FAS score was used for the COWAT as well as the total verbal fluency score.

The WMS-III provides indices for both immediate and delayed visual and auditory memory. We also reported the indices for general memory, working memory, and immediate memory.

Finally, a Grooved Pegboard Test (GPT) for manual dexterity as well as a series of tapping tasks (Leonard Tapping Tasks; LTT) were administered to determine motor coordination. For the GPT, participants had to insert key-like pegs into holes by turning them until they fit in the randomly oriented holes. We recorded the time to insert and remove the pegs for each hand. The LTT in this study used Leonard's computerized version of the Thurston tapping tasks [398], which have demonstrated deficits on complex bimanual tapping in patients with frontal lobe lesions sparing the precentral gyrus [399]. The tapping tasks involved (i) spatially ordered unimanual tapping where the participant has to tap a stylus on a plate with four numbered quadrants in sequential order for 30 seconds; (ii) bimanual balanced tapping where the participant has to tap in the same spatial location on each plate with each hand (i.e., the hands moved in phase) for 30 seconds; (iii) bimanual unbalanced tapping where the quadrants to be tapped with each hand are in different spatial locations (i.e., the movements of the two hands were out-of-phase) for 30 seconds; and, finally, (iv) a simple unimanual rapid tapping motor task where the participant has to tap a single quadrant as fast as possible for 15 seconds [397].

5.3.4 MRI Analyses

Cortical Thickness

Cortical thickness was estimated on the 1 mm isotropic T1w anatomical MR images using v5.1.0 of FreeSurfer [217, 223]. In addition to the WM and pial surfaces that are generated for each hemisphere, a parcellation of the cerebral cortex [225] was also used to define surface-based regions of interest (ROIs). Specifically, we combined the parcellation as described in Desikan et al. [226] to create ROIs of the frontal, occipital, temporal and parietal lobes as well as ROIs of the primary motor strip and the premotor cortex (Figure 5.1).

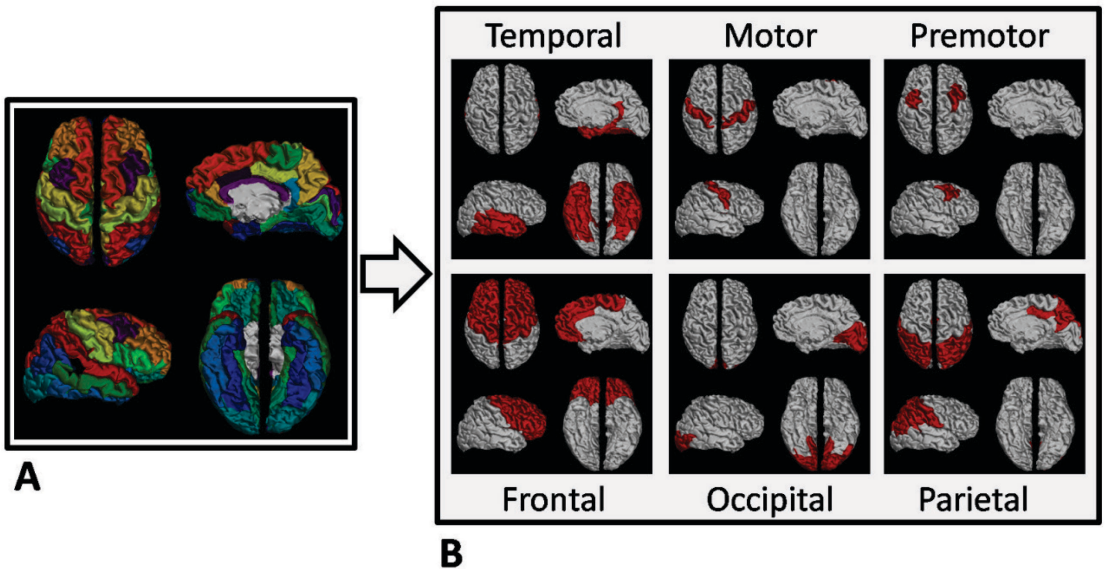

Figure 5.1 – The cortical regions of interest examined in the study. (A) The output of the FreeSurfer cortical parcellations displayed on the pial surface for one of the subjects in the study. Each colour represents one of the different parcellations described in Desikan et al. (B) We combined select parcellations to create ROIs of the (i) frontal lobe (which, using the FreeSurfer nomenclature, comprised these gyri: caudal anterior cingulate, caudal middle frontal, frontal pole, lateral orbitofrontal, medial orbitofrontal, paracentral pars opercularis, pars orbitalis, pars triangularis, precentral, rostral anterior cingulate, rostral middle frontal, superior frontal); (ii) occipital lobe (cuneus, lateraloccipital, lingual,pericalcarine); (iii) temporal lobe (bankssts, entorhinal, fusiform, inferior temporal, isthmus cingulate, middle temporal, parahippocampal, superior temporal, temporal pole, transverse temporal); (iv) parietal lobe (inferiorparietal,postcentral, posteriorcingulate,precuneus, superiorparietal,supramarginal); (v) primary motor strip (precentral); and (vi) premotor cortex (caudal middle frontal).

All surfaces were visually inspected, and manual edits were made where appropriate (i.e., fixing skull stripping issues, filling "holes" in the WM segmentation, setting WM control points, etc.). Of the 26 brains whose surfaces were extracted, five required some form of manual intervention (four subjects and one control).

Cortical thickness values were blurred with a geodesic diffusion kernel (i.e., surface-based blurring) [228], and mean thickness values were reported for the ROI analysis. While there was no a priori knowledge about the size of kernel that would be appropriate (i.e., filter matching theory suggests that the kernel size should be the same size as the target of analysis, which in this case would be the regional size of cortical atrophy), blurring was still performed in order to (i) reduce the noise in the

measurement, thus increasing the signal-to-noise ratio; (ii) increase the validity of statistical tests by rendering the data more normally distributed and reducing the number of comparisons that need be statistically controlled for (i.e., blurring increases the interdependence of the neighbouring vertices); and (iii) reduce the impact of imperfect anatomical alignment when comparing the same vertices between cortices [228]. Several kernel sizes were investigated, but as they had little impact on the results and did not change the interpretation of our findings, we used a modestly sized 10 mm full-width-at-half-maximum (FWHM) kernel for all analyses.

Cortical MTR

To calculate the MTR for each patient, the magnetization transfer image acquired with the saturation pulse (Sat) was first linearly registered (mritoself [363]) to the volume without the saturation pulse (NoSat), and the MTR map was then calculated as 100 * (NoSat – Sat)/NoSat. In order to remove outliers caused by noise and possible data discretization errors, the MTR image was clamped so that all values were within the theoretical limits of 0 and 100.

The cortical surfaces extracted from FreeSurfer were also used to create an intermediate surface lying halfway between the WM and pial boundaries. A projection of the MTR values registered to the T1w anatomical image was then loaded onto this intermediate surface, and as above, the mean 10 mm FWHM surface-blurred MTR values were reported for the ROI analysis.

Surface-Based Cortical MTR Abnormalities

An estimate of focal cortical damage was obtained using the methods previously described in [400]. Briefly, using surface-based analyses, statistically significant regions of abnormally low MTR that resemble Type III subpial demyelinating cortical lesions were automatically identified, and the total surface area of these irregularities as well as their regional distribution was recorded for each subject.

Volumetrics

Normalized whole-brain, peripheral GM, and WM volumes were obtained using v2.6 of SIENAx [202], distributed as part of v4.1.7 of FSL [200, 201]. While SIENAx is widely used and provides accurate measures of whole-brain and peripheral GM volume, specialized software was used to help ensure the accurate quantification of deep GM structures [359]. Normalized measures of caudate nuclei, thalamus, and putamen were obtained via the volumetric stream of FreeSurfer [208], while normalized hippocampal volumes were obtained using the even further specialized label fusion-based template library scheme described in [401]. Importantly, the same scaling factor used for the normalization of head size in SIENAx was applied to the deep GM volumes obtained via the other methods to ensure consistency.

5.3.5 Statistical Analyses

In order to test the primary hypothesis, each of the five neuropsychological metrics (WAIS-III Processing Speed Index, TOL total problem-solving time, PASAT time, LTT unimanual right-handed rapid tapping score, LTT bimanual unbalanced tapping score) was modeled by cortical thickness or cortical MTR values within combinations of selected ROIs. While the strength of the fit of the model is explained with the r^2 and associated p-value, the appropriateness of one hypothesized model over another was assessed with the second-order Akaike Information Criterion (AICc) [402]. As the number of parameters varies in each model, the use of a fixed level of significance (i.e., p-value) for the comparison of models would be inappropriate because it does not take into account the increase in the variability of the estimates when the number of parameters is increased from model to model [402]. The AICc was preferable to the first-order AIC because of the sample size of our study and the number of parameters in the hypothesized models [403, 404]. Since AICc values are difficult to interpret beyond just the fact that lower is better, we also calculated the more intuitive Akaike weights [403, 405]. Defined as the difference between the AICc of the current model being evaluated and the model with the lowest AICc (termed delta AICc), relative to the sum of the delta

AICc for all candidate models, the Akaike weighting is straightforward to interpret as it indicates the probability that the model is the best among the whole set of candidate models. For example, a model with an Akaike weight of 0.75 indicates that, given the data, it has a 75% probability of being the best of the candidate models considered [404].

Though we used the framework of a general linear model implemented in R (R project for statistical computing [364], v2.14.1 on a 64-bit Linux machine running Ubuntu 12.04), the varying terminology for statistical tests used within the literature prompted us to cross-validate our tests with other statistical software packages whenever possible (i.e., JMP v8.0 (SAS Institute Inc., Cary, USA) and SPSS v20.0.0 (IBM Corp., Armonk, USA) on a Windows XP machine). For example, a model with two main effects modeled in R as $y \sim x_1 + x_2$ produces an r^2 and associated p-value for the x_1 term that can be interpreted as the strength of the linear relationship between y and x_1, while controlling for the effects of x_2. That same r^2 value can be referred to as a main effect, or leveraged r^2 if one uses the general linear model analysis in JMP, or as a partial r^2 using either the partial correlation analysis or the multilinear regression analysis menus in SPSS.

For the WAIS-III Processing Speed Index, because we hypothesized that focal damage to the frontal cortex would explain the variations in the neuropsychological metric, we first tested the relationship of this index to the mean cortical thickness and mean MTR of the frontal lobe. To determine whether the effects were truly driven by the frontal lobe and not by just a global change in cortical thickness or MTR, we also tested the strength of the correlation to global measures of cortical thickness and MTR. Furthermore, we looked at the partial r^2 of frontal lobe thickness and MTR while controlling for nonfrontal thickness and MTR. In total, six models were tested for the WAIS-III Processing Speed Index, and the same six models were also used for the PASAT and TOL.

For the two LTTs, we hypothesized the involvement of the specialized motor cortices (rather than the more general frontal lobe) and thus used slightly different models. In

the case of the unimanual right-handed rapid LTT, we hypothesized that the left-hemispheric motor strip would best explain the variations in the subjects' performance. As above, we modeled the relationship between the task and (i) the mean cortical thickness and MTR of our specific ROI; (ii) a global measure of cortical thickness; and (ii) our specific ROI while controlling for a global effect. In the case of the LTT bimanual unbalanced tapping tasks, we hypothesized the ROI involved would be the premotor cortex.

For the secondary analysis, the effect size between the NC and patients with MS for each neuropsychological test and MRI metric was computed using Cohen's *d*.

5.4 RESULTS

5.4.1 PRIMARY HYPOTHESES

The models, and their associated r^2 and *p*-values, as well as their AICc and Akaike weights can be found in Table 5.2. As there were multiple models tested for each of the primary hypotheses tested, all *p*-values presented were Bonferroni-corrected.

Estimates of focal cortical damage using the surface-based MTR technique revealed putative cortical lesions in six of the patients with MS, with only three of them having an affected surface area greater than 1% of the total cortical surface area. Because the sample size of those presenting with focal cortical damage was so small, we did not include the component as a covariate in any of the models.

For the WAIS-III Processing Speed Index, a significant correlation was found with both the frontal lobe cortical thickness ($r^2 = .21$, $p = 0.03$) and frontal lobe MTR ($r^2 = .29$, $p < 0.01$); however, these relationships were no longer present after controlling for the effect of nonfrontal thickness and MTR, which suggests that the regional relationship is actually being driven by a global effect. Indeed, when the global cortical MTR value was regressed against the Processing Speed Index, there was a significant relationship ($r^2 = .32$, $p < 0.01$), and the Akaike weight for that model was the highest at 0.60 (indicating it has a 60% chance of being the best of the six models tested).

For the PASAT, none of the six models tested were found to be significant, and the correlations were weak at best ($r^2 < .07$).

The TOL correlated with global cortical MTR values ($r^2 = .27$, $p = 0.01$), and though the frontal MTR values also correlated ($r^2 = .26$, $p = 0.01$), when the global effect was controlled for, this relationship disappeared. None of the correlations with cortical thickness were found to be significant after correction for multiple comparisons.

The unimanual right-handed rapid LTT was found to correlate with the left-hemispheric motor cortex thickness ($r^2 = .23$, $p < 0.05$), though the strength and significance of the correlation diminished when global cortical thickness was controlled for. Interestingly, the Akaike weights indicated that the left-hemispheric motor MTR value was the best candidate model (0.57), even though the model was not significant after multiple comparison correction ($r^2 = .20$, $p = 0.11$).

The final neuropsychological test that was part of the primary hypothesis was the bimanual unbalanced LTT. Though the MTR in the premotor cortex was the model with the best Akaike weight of 0.56 ($r^2 = .29$, $p = 0.02$), this effect was no longer present after controlling for nonpremotor cortical MTR values. When regressed against cortical thickness, this task was the only one that exhibited a true regional effect; the correlation with just the premotor cortical thickness value was moderate ($r^2 = .19$, $p = 0.10$), but increased significantly once the nonpremotor cortical thickness values were controlled for ($r^2 = .30$, $p = 0.01$).

Table 5.2 – Correlations between the neuropsychological metrics and cortical MTR and thickness values within selected ROIs

Test	ROI	r^2	p	AICc	Akaike weight
WAIS-III processing speed subindex	frontal MTR	0.29	**0.01**	288.38	0.25
	frontal MTR (controlling for nonfrontal MTR)	0.02	2.88	289.43	0.15
	global MTR	0.32	**0.00**	286.64	**0.60**
	frontal THX	0.21	0.03	316.14	0.00
	frontal THX (controlling for nonfrontal THX)	0.14	0.15	316.77	0.00
	global THX	0.16	0.09	318.16	0.00
PASAT	frontal MTR	0.06	1.02	243.84	0.41
	frontal MTR (controlling for nonfrontal MTR)	0.01	4.21	246.23	0.12
	global MTR	0.07	0.87	243.57	**0.47**
	frontal THX	0.00	5.99	269.81	0.00
	frontal THX (controlling for nonfrontal THX)	0.01	3.26	271.88	0.00
	global THX	0.00	5.13	269.78	0.00
TOL	frontal MTR	0.26	**0.01**	279.58	0.39
	frontal MTR (controlling for nonfrontal MTR)	0.03	1.86	281.69	0.14
	global MTR	0.27	**0.01**	279.18	**0.48**
	frontal THX	0.16	0.08	306.68	0.00
	frontal THX (controlling for nonfrontal THX)	0.03	1.85	309.19	0.00
	global THX	0.15	0.10	307.13	0.00
Right-handed rapid LTT	left-hemispheric (LH) motor MTR	0.20	0.11	239.19	**0.57**
	LH motor MTR (controlling for LH nonmotor MTR)	0.08	1.05	241.61	0.17
	LH global MTR	0.15	0.26	240.79	0.26
	LH motor THX	0.23	**0.05**	254.13	0.00
	LH motor THX (controlling for LH nonmotor THX)	0.14	0.31	256.12	0.00
	LH global THX	0.14	0.25	257.30	0.00
Bimanual unbalanced LTT	premotor MTR	0.29	**0.02**	228.45	**0.56**
	premotor MTR (controlling for nonpremotor MTR)	0.05	1.71	231.17	0.14
	global MTR	0.26	**0.04**	229.68	0.30
	premotor THX	0.19	0.11	247.41	0.00
	premotor THX (controlling for nonpremotor THX)	0.30	**0.01**	243.83	0.00
	global THX	0.07	1.03	251.50	0.00

Correlations were considered significant at a corrected p-value of 0.05 (bold).
WAIS-III = Wechsler Adult Intelligence Scale; PASAT = Paced Auditory Serial Addition Test; TOL = Tower of London; LTT = Leonard Tapping Test; ROI = Region of Interest; MTR = Magnetization Transfer Ratio; THX = Cortical Thickness; AICc = second-order Akaike Information Criterion.

5.4.2 SECONDARY AIM

The absolute effect size of each of the neuropsychological tests is presented in Figure 5.2 and those for the MR metrics can be viewed in Figure 5.3.

The effect sizes for each neuropsychological test as a whole were similar, though the TOL, D-KEFS verbal test, and WCST seem limited in their ability to distinguish between our two cohorts of subjects. By dividing each test into its components (subtests), it became clear that some subtests had much larger effect sizes than others. The WAIS-III, for example, contains several subindices, but the Processing Speed Index, comprising the Symbol-Search test and the Digit-Symbol Coding test, had the largest effect size (Cohen's $d = 1.52$). In turn, one can see the contribution of each of those tests; the Symbol-Search test had the largest effect size of any of the subtests (Cohen's $d = 1.74$), while the Digit-Symbol Coding test had only a moderate effect size but would still be considered capable of distinguishing between the two cohorts, unlike, for example, the Working Memory Index from the WAIS-III (Cohen's $d = 0.17$) or the BNT (Cohen's $d = 0.05$).

For the MR metrics, the surface-based MTR values of the cortex had the largest effect sizes, with the temporal lobe being the largest (Cohen's $d = 1.64$). Volume-based measures of MTR (i.e., within selected volumetric masks) were not as effective as the surface-based measures, with whole-brain MTR having the smallest effect size (Cohen's $d = 0.19$) of any of the MR metrics.

The cortical thickness measurements produced fairly similar results (Cohen's $d \approx 0.80$), with the exception of the premotor cortex, which had a superior ability to distinguish between the two groups (Cohen's $d = 1.14$).

Finally, volumetric measurements of normalized deep GM and other brain structures also produced, on average, large effect sizes (Cohen's $d > 0.80$), with the volume of the thalamus being a notable metric for detecting differences between the NC and patients with MS (Cohen's $d = 1.37$).

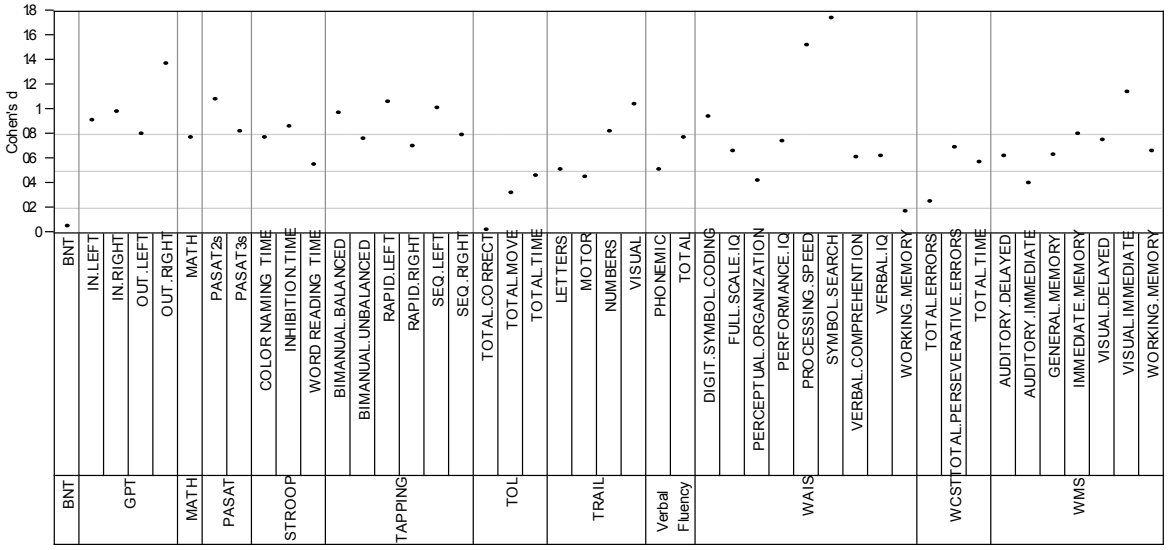

Figure 5.2 – The effect size (absolute Cohen's d) for each of the tests that were part of the neuropsychological battery. Dotted lines representing small (0.2) , medium (0.5) and large (0.8) effect sizes are added for reference. BNT = Boston Naming Test; GPT = Grooved Pegboard test; PASAT = Paced Auditory Serial Addition Test; TOL = Tower of London; WAIS-III = Wechsler Adult Intelligence scale; WCST = Wisconsin Card Sorting Task; WMS-III = Wechsler Memory Scale.

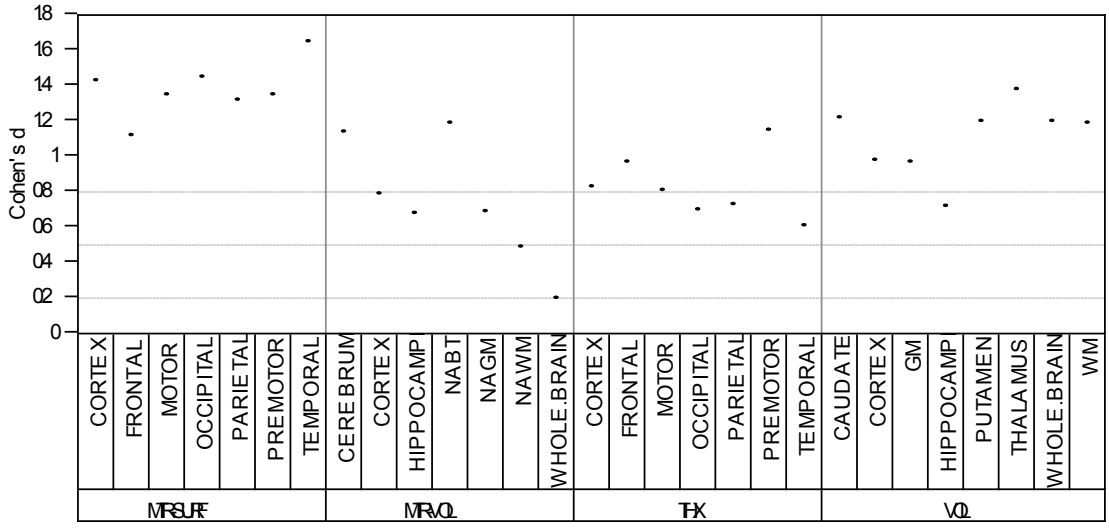

Figure 5.3 – The effect size (absolute Cohen's d) for each of the MR metrics tested. Dotted lines representing small (0.2) , medium (0.5) and large (0.8) effect sizes are added for reference. MTR-SURF = Magnetization Transfer Ratio measured along the surface; MTR-VOL = MTR measured within a selected volume; THX = cortical thickness; VOL = anatomical volume; NABT = normal appearing brain tissue; NAGM = normal appearing grey matter; NAWM = normal appearing white matter; GM = grey matter; WM = white matter.

5.5 DISCUSSION

5.5.1 PRIMARY HYPOTHESES

Our primary aim was to evaluate the relationship between select neuropsychological measures and MR metrics of cortical damage. We found for nearly all of the tests that, while at first glance regional measures of cortical damage (i.e., MTR or cortical thickness) appeared to be predictors of cognitive ability, once the corresponding global measures were controlled for, this relationship was no longer present. This finding, combined with the significant correlations with the global measures, leads us to believe that performance on certain tests is not simply the result of regional cortical damage, but rather of a potentially much broader and more widespread pathology.

Our attempt at an in-depth investigation of the relationship between focal cortical damage (assessed via putative cortical lesions identified using our surface-based MTR technique) and CI was undermined by our sample size, or more specifically, the low number of patients with secondary-progressive MS included in this study (three out of 26 patients). Pathological observations indicate that the later stages of the disease are characterized by an increase in the frequency and extent of subpial demyelinating lesions [39], and our results (i.e., that only 3/23 patients with relapsing-remitting MS displayed focal cortical MTR abnormalities, compared with 2/3 secondary-progressive MS subjects), are consistent with those pathological findings. That our total cohort studied had a wide range of CI with the majority of them showing no focal cortical damage leads us to believe that such damage is not the cause for CI in patients with MS; thus, we focused our efforts on (broader) regional and global measures of cortical damage. Our results were of course limited by the MR signal, contrast and resolution available at the clinically meaningful field strength of 3.0 T, and thus future research at 7.0 T and/or with additional modalities, may be used to investigate the nature of this relationship further. Future studies would also benefit from having a cohort with a normal distribution of subjects with focal cortical damage.

A possible explanation for the strong relationship we observed between CI and global measures of cortical pathology is that the neuropsychological tests may not be as focal or regionally specific as many purport them to be. Indeed, many of the tests are multifaceted and involve complex networks.

Identified early on as one of the cognitive domains in which patients with MS are deficient [375], processing speed has been heavily studied in MS. Most often measured with the SDMT, mental processing speed can also be measured with the PASAT. Although reliable and modestly sensitive in MS [380], the PASAT requires mathematical skills, is susceptible to practice effects [406], and has been shown to involve different anatomical and functional areas compared with the SDMT [407]. Our results showed no significant correlations between the PASAT and any of the measures of cortical damage that we tested. We did not perform the SDMT as part of our neuropsychological battery; however, the WAIS-III Processing Speed Index was used as a measure of mental processing speed. Previous studies have shown robust correlations between measures of processing speed (SDMT) with specific ROIs (e.g., putamen and thalamic volumes, caudate atrophy, frontal and parietal lobe thinning [247, 306, 323, 407, 408]). Our investigations, though focused on cerebral cortical damage, revealed two important findings: (i) that measures of processing speed correlated with global and not regional cortical damage and (ii) that reduced cortical MTR was a better predictor of disability than measures of cortical thickness.

The correlation with global cortical damage was also found with the TOL and unimanual LTT (Table 5.2), whereas regional cortical damage was observed only in the case of the bimanual unbalanced LTT.

As for MTR being a better predictor of disability than measures of cortical thickness, this held true regardless of the neuropsychological test. The AICc and the Akaike weights indicated that, for each test, the best model involved cortical MTR—and not cortical

thickness—to the extent that all cortical thickness models tested had a less than 1% chance of being the best predictor.

5.5.2 SECONDARY AIM

According to Cohen's original interpretation, effect sizes of 0.2, 0.5 and 0.8 were considered small, medium, and large, respectively. Alternatively, effect sizes can also be interpreted as the percent of nonoverlap of the distribution of the two groups. An effect size of 0 indicates that the two groups overlap completely, that is, 0% nonoverlap. An effect size of 0.8 translates to 47.4% nonoverlap, while an effect size of 2.0 can be interpreted as 81.8% nonoverlap. A full table of nonoverlap percentage equivalents to Cohen's *d* effect size can be found in [409].

Neuropsychological Metrics

To our surprise, the TOL was among the tests that displayed the most overlap between the NC and our patients with MS. We had chosen the TOL as one of our primary hypotheses following promising results from our 11-subject pilot study that showed a strong relationship between the TOL and frontal lobe cortical thickness [369], though it is clear now that this correlation was driven by the small sample size and/or global cortical damage that was not controlled for.

Apart from the TOL, BNT, and WAIS-III Working Memory Index, the majority of the tests had moderate to large effect sizes and would be considered adequate for detecting CI in patients with MS compared with controls. The Symbol-Search test, a subtest that makes up the WAIS-III Processing Speed Index, stood out as being particularly sensitive in distinguishing between controls and patients with MS. This result comes as no surprise as mental processing speed has already been well established as one of, if not the primary, cognitive domain in which patients with MS are deficient compared with NC. As this was a cross-sectional study we cannot currently comment on the evolution of CI in patients with MS, but as we acquire the longitudinal data for this ongoing study, it will be interesting to see how the patients' cognitive profiles are affected by time.

MR Imaging Metrics

Of the MR metrics tested, the surface-based MTR markers had the largest effect size. For example, the MTR along the entire cerebral cortex was found to have 68% nonoverlap between groups (Cohen's $d = 1.41$), whereas cerebral cortical thickness only displayed 47% of nonoverlap (Cohen's $d = 0.78$), suggesting that surface-based MTR metrics may be preferable to cortical thickness for measuring disability in patients with MS.

Volume-based measures of MTR on the whole produced the same degrees of nonoverlap as the cortical thickness measurements; however, there was a large variation depending on the mask that was used to calculate the MTR. Perhaps most striking is the difference between whole-brain and cerebral measures of MTR, a difference attributed to the inclusion of cerebellar tissue, which we believe adds noise and thus dilutes the aberrant signal of interest. We also observed that volumetric measures of MTR within the cerebral cortex produced a smaller effect size than surface-based measures of MTR within the cortex, a finding which we attribute to several factors: (i) we believe that cortical surface extraction provides a better segmentation of cerebral cortex than voxel-based volumetric segmentation; (ii) for our surface-based measures, we used an intermediate surface between the pial and WM surfaces, thus ensuring that minimal inclusion of data points that were not exclusively cortical GM; and (iii) surface-based blurring was used to increase the signal-to-noise ratio, whereas for the volumetric measurements, blurring was ill-advised since the cortical mantle is a complex and thin folded ribbon surrounded by cerebrospinal fluid on one side and WM on the other; that is, volume-based blurring would have diluted our signal of interest with unwanted data adjacent to the cortex.

Volumetric measurements of different brain structures also displayed a large effect size. Our findings support the literature regarding basal ganglia atrophy and CI involvement in patients with MS [61, 306, 308, 313, 321], with normalized thalamic volumes showing 67% nonoverlap (Cohen's $d = 1.37$). Also worth noting is that the effect size produced with volumetric measures of normalized peripheral GM produced via SIENAx was

slightly larger than cortical thickness measures, leading us to believe that the quantification of cortical GM volume using SIENAx is at least as good as the more complex and highly time-consuming surface extraction method of FreeSurfer for the purpose of differentiating patients with MS from controls.

5.6 CONCLUSIONS

The results of the primary aim of this paper demonstrate that performance on certain tests designed to measure CI is not simply the result of focal or regional cortical damage, but rather of a much broader and more widespread pathology. Furthermore, cortical MTR values proved to be a better predictor of CI in patients with MS than did measures of cortical thickness.

The results of the secondary aim of this paper demonstrate that a wide range of neuropsychological tests are adequate to distinguish patients with MS from normal controls, with one particular test, the WAIS-III symbol search, outperforming the others. As was the case with the primary hypotheses, surface-based MTR metrics also outperformed measures of cortical thickness in their ability to distinguish between the two cohorts studied.

APPENDIX 5A

Supplementary Table 1 – Neuropsychological testing order.

Day 1	Day 2
WAIS-III Information and Orientation	Leonard Tapping Tasks
Handedness	Boston Naming Test
Picture Completion	WMS-III List Learning Immediate
Vocabulary	Trails
Stories Immediate	Visual - Reproduction Immediate
Faces Immediate	STROOP
Paired Associates Immediate	Tower of London II
Digit Symbol	Picture Arrangement
Symbol Search	Card Sorting
Family Pictures Immediate	Comprehension
Letter/Number	PASAT
Spatial Span	Object Assembly
Mental Control	Word Fluency
Digit Span	Math Fluency
Similarities	WMS-III List Learning Recall
Block Design	WMS-III List Learning Recognition
Arithmetic	Grooved Pegboard Test
Stories Recall	Visual Reproduction Recall
Stories Recognition	Visual Reproduction Recognition
Faces Recall	
Paired Associates Recall	
Paired Associates Recognition	
Family Pictures	
Matrix Reasoning	
WMS-III information and Orientation	

Notes: Each testing day lasted approximately two and half hours.

Chapter 6

DISCUSSION

The main goal of this dissertation was to evaluate and develop image processing methods for the in vivo quantification of GM pathology in patients with MS. First, we explored the validity of the most common techniques that have been used to draw conclusions about GM atrophy in patients with MS, identifying their strengths and weaknesses to help shape the way studies will be interpreted in the future. Second, we developed and evaluated the sensitivity of a novel surface-based technique that allowed for the identification of a more focused type of cortical GM pathology, namely, putative areas of subpial demyelination. Finally, we used the developed surface-based framework to investigate the relationship of our uniquely identified cortical pathology to cognitive impairment in patients with MS.

This chapter presents a discussion of each of the three parts of this thesis, summarizing the main findings, discussing the strengths and weaknesses, highlighting the clinical relevance, novelty, and scientific interest of each manuscript, as well as suggesting future improvements and applications. For the sake of brevity, the discussion will focus on those aspects that were not already stressed in the manuscripts themselves.

6.1 AUTOMATED GREY MATTER SEGMENTATION METHODS

The focus of the study presented in Chapter 3 was to explore the validity and variability of some of the freely available automated methods currently used in the literature to segment and estimate GM atrophy in patients with MS. Results demonstrated that, although the algorithms perform similarly to manual segmentation of cGM, there are severe shortcomings as far as the segmentation of dGM structures.

One of the strengths of the study was in the selection of the GM segmentation that was used as the ground truth, or reference standard. STAPLE [339] provided an optimal linear combination of the experts' manual segmentations that was used as the reference standard and is preferable to simple voting rules where the resultant segmentation varies greatly depending on the number of votes required (see Figure 6.1). As mentioned in the manuscript, a phantom or MR imaging simulator is a viable alternative—and in fact has the advantage that there is an a priori established ground truth segmentation; however, using such a method would not have reproduced the full range of imaging artifacts, nor would it have resulted in the anatomical or pathological variability observed in clinical data. In other words, by using the typical images obtained in a clinical trial setting, the study was not simply a scientific excursion, but had real-world clinical relevance.

Another strength of the manuscript was that the study design afforded the use of straightforward statistical tests to quantitatively assess and support our qualitative observations. While this may seem trivial, it allows the results to be easily interpretable, and without the need for complex assumptions about multiple comparisons, model parameters, degrees of freedom, and so forth, the conclusions drawn from the study are not readily contentious.

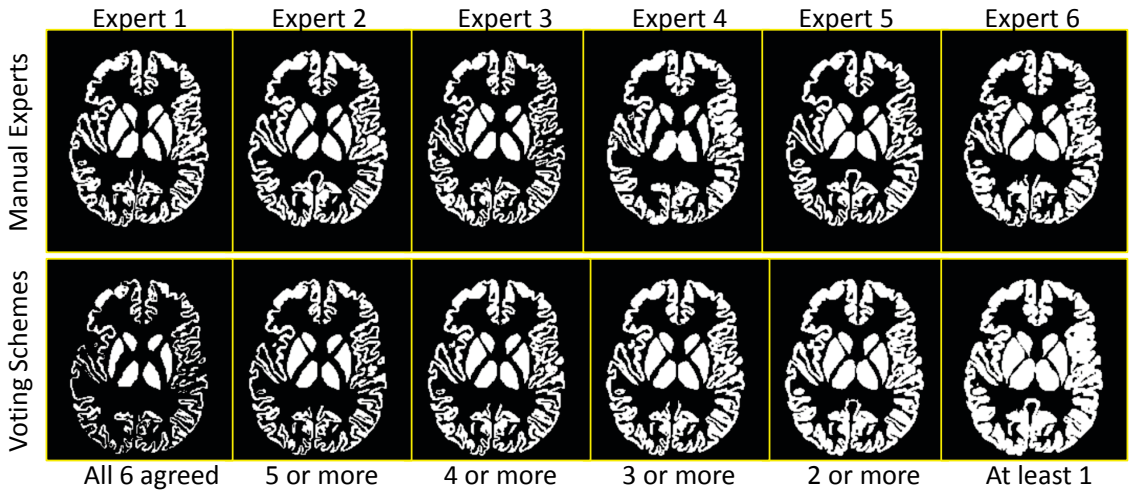

Figure 6.1 – GM segmentation used in the study. Top row: The manual GM segmentations from the six experts. Bottom row: The resulting reference gold standard segmentation that results from the different voting schemes. Note how the amount of GM included changes dramatically depending on the voting scheme.

A final strength of the study is the already-alluded-to clinical relevance of the manuscript. The automated methods tested were not chosen at random, but after a review of the most commonly employed methods that are being used both in academic studies and in large-scale clinical trials. The impact of the results of these trials has profound effects not only for the pharmaceutical companies that have invested millions of dollars in the trials, but more importantly, on the lives of the patients the studies were designed to help. With such matters resting on the results of these trials, one gains an appreciation for the need for an accurate GM segmentation in order to assess neurodegeneration. To put our findings into perspective, the rate of atrophy that disease-modifying therapies are trying to reverse is on the order of 1% per year in patients with MS [352, 353], and the gross misclassification of dGM we illustrated in the manuscript accounts for approximately 3.8%.

Perhaps the most glaring weakness of the manuscript is the low number of subjects (n = 3) and slices (n = 3) used for validation, a point that was mentioned therein. Under ideal circumstances, we would have of course dramatically increased the number of subjects, and instead of simply having a single subject with low, medium, and high

lesion load, we would have had a normally distributed sample.[liii] This would have allowed us to comment properly on the impact of lesion load on the GM segmentation results, a topic of much interest and active research [241, 350, 351, 410]. In actuality, we did have access to appropriate images. Where we were limited, however, was with the tedious and time-consuming act of careful, fully manual segmentation of the GM of the cerebrum. Recall that we solicited six experts (i.e., radiologists, neuroradiologists, and neurologists), all trained in manual segmentation. Given that the time of these experts is limited, it was just not feasible to have as large a dataset as we would have liked.[liv] In fact, to address this lack of time, our first design of the study used only three experts to generate our STAPLE reference standard. However, after an initial analysis, a number of reasons prompted us to increase the number of experts. First, we noticed a high degree of variation among the readers' GM segmentations (evidenced by the high standard deviation in the inter-reader Dice similarity coefficients). Second, this high variability combined with the choice of only three segmentations made it difficult for STAPLE to create an optimal linear combination of the segmentations. Since the comparisons of the automated techniques would be made to this single reference standard, we wanted to be confident in its accuracy. After discussions with the author of STAPLE, it was concluded that our reference standard would benefit greatly from the addition of more expert readers.

Regarding the low number of slices, we chose slices that were representative of the anatomical and artifactual challenges. While a full segmentation of the cerebrum would have been ideal, we believe the addition of more slices would not have changed the major findings of this paper. Furthermore, given the gross degree of error in the dGM

[liii] As this was a retrospective study, we were fortunate enough to have access to a large dataset of MR images of brains of patients with MS. The decision to use scans displaying varying degrees of lesion burden was again fueled by our quest for clinical relevance. While healthy brains from normal controls would have still revealed much insight into the performance of the automated methods, the results would not be directly beneficial to the MS imaging community. But, as is discussed above, the limitation we faced was not related to the availability of data.

[liv] While we could have easily subjugated a number of undergraduate students to perform the time-consuming and tedious task of labelling the GM on every slice of multiple brains, we felt it was more important, especially when trying to establish a gold standard, to have fewer segmentations but more confidence in their accuracy.

segmentation that is readily visible in Figure 3.4, even a single slice allows for an adequate (albeit not comprehensive) comparison.

The other large pitfall of the paper is that it only examined cross-sectional methods for measuring GM volume, which is not synonymous with atrophy. The manuscript includes a discussion of how accuracy does not necessarily translate to the reproducibility of the technique, nor to its sensitivity in detecting change. Longitudinal methods where serial scans from an individual are accurately registered and the volume changes are derived directly (e.g., SIENA [232], brain boundary shift integral [231]) are in fact preferable to repeated cross-sectional measurements. However, the current limitation of such techniques is that they only provide a measurement of whole-brain atrophy. What would be ideal, and is the subject of future work, is a method that combines the sensitivity of the longitudinal methods while still maintaining the tissue classification features of the cross-sectional measurements to provide a more accurate and precise measurement of tissue-specific atrophy. At the 2012 International Society for Magnetic Resonance in Medicine (ISMRM) conference, we proposed such a method [411]. By combining SIENA and SIENAx, we arrived at a technique that (i) when tested with a simulated atrophy dataset, had reduced variance and increased sensitivity over its repeated cross-sectional measurement counterpart; (ii) when tested with a scan-rescan dataset, had significantly improved reproducibility over simple cross-sectional measurements; (iii) was equally applicable to MR images of 1.5 mm or 3 mm slice thickness; and (iv) when applied to a multi-centre clinical trial dataset, had slightly lower variances that trended towards significance. For reference, the poster is reproduced in Appendix 6A.

While the method is presented using SIENA and SIENAx, the technique is not limited to these specific tools. Any method for tissue-specific segmentation can be used—and combined with any method for detecting a longitudinal change. In fact, future work that is already underway is looking not only at using other automated segmentation methods besides SIENAx, but also using the Jacobian of an optimized nonlinear

registration to determine the longitudinal change instead of the percent brain volume change reported from SIENA.

Finally, it should be noted that, as the manuscript was an evaluation of existing methods and did not introduce a new method, in terms of developed methodology, the novelty of the paper was minimal. What it lacked in original methods, however, it gained in scientific importance and in the novelty of its findings. FSL is one of, if not the most, widely used software library used within the neuroimaging community. It now includes a separate segmentation of peripheral GM in addition to whole-brain GM as a direct influence of this work.[lv] While having the manuscript referenced a thousand times would have been more advantageous from an academic standpoint, the result of this paper is that thousands[lvi] of users feel its impact every time they use SIENAx.

[lv] Following the presentation of the preliminary findings at MICCAI 2009, I was able to discuss the results, and the changes that would be required to the SIENAx pipeline, with Mark Jenkinson from FSL. The following release of FSL contained an estimate of peripheral GM where the questionable segmentation of the dGM had been masked out as discussed.

[lvi] A gross estimate based solely on the number of users on the FSL mailing list

6.2 A PROPOSED METHOD FOR IMAGING SUBPIAL DEMYELINATION

In contrast with the manuscript presented in Chapter 3, the novelty of the method introduced in Chapter 4 was one of the strengths of the manuscript. With our oral presentations at the 2009 [354] and 2010 [355] ISMRM conferences, we were the first group to present an in vivo image processing technique that was capable of detecting changes that resembled the subpial demyelination seen in the postmortem pathological assessments of the brains of patients with MS. Until then, and even until very recently, the focus had been on the development of new sequences and novel contrast mechanisms in order to visualize cortical lesions, and even then, the success of such efforts was limited to the visualization of mixed WM/GM (Type I) or purely intracortical (Type II) lesions.

Moreover, although the method presented in Chapter 4 is presented in the context of detecting subpial demyelination using MTR images, the method is highly generalizable and can be applied to any imaging modality to look for aberrances in the cortex of an individual (or population). In fact, our ISMRM abstract has been referenced for similar work that has already been published using T2* imaging at 7 T [137].

The main finding of the paper, that there were decreased regions of subpial MTR in patients with MS, was unparalleled and of significant interest to the clinical community. In addition, we helped confirm the existence of focal cortical pathology in patients with MS, as well as underscored the importance of using MTR, a somewhat nonconventional sequence, for the detection of such pathology. The other chief finding of the paper—that on both the group and individual level, cortical decreases in MTR are more apparent in patients with the progressive forms of the disease than in the relapsing-remitting cohort—confirmed via in vivo imaging what the pathologists had reported via postmortem immunohistochemical analysis (see Chapter 2), and is also clinically relevant as it helps define who is affected by such pathology.

Concerning postmortem histological analysis, it is the lack of this form of validation that is the major weakness of the manuscript. Without the combined imaging and postmortem validation, the decreases in cortical MTR that we detected cannot be unequivocally referred to as subpial demyelination.[lvii] In addition to this absence of definite pathologic specificity, we are unable to provide a true sensitivity analysis. In its place, however, the manuscript did provide a form of sensitivity analysis in that we presented two methods for determining the minimum amount of decrease in MTR signal that would be required for detection. While this is far from the ideal validation, it did validate the practical applicability of the method.[lviii] Still, postmortem validation is probably the single most important area on which to focus the future work related to this manuscript. Unfortunately, we were unable to carry out such work under the scope of this thesis due to the complications[lix] encountered in obtaining and acquiring postmortem images and histological data that can be used in the same framework as the presented method.

Further research should also be devoted to modifying the current method so that it is not reliant on a normative group for comparison. For many studies, and especially for clinical trials, it would be extremely costly and impractical to have an additional arm of normal controls to be used just for imaging purposes. In addition, the current method has not been validated across multiple MR scanners, and we actually hypothesized that the method would not work well as there are observable differences in MTR across scanners (recall that MTR is only a semi-quantitative technique). At the root of this issue is the variability in the MTR image itself, though this can possibly be overcome using an intensity normalization method (another area of active research).[lx]

[lvii] The astute reader will have noticed the careful wording in the manuscript, which states that the pathology imaged corresponded to "putative" areas of subpial demyelination and that the method was presented as a "possible" technique for imaging subpial demyelination.

[lviii] For example, if the minimum amount of decrease in MTR was on the order of a 10 p.u. change, then the technique, however sophisticated, would be impractical since this degree of difference (or change) in MTR is rarely seen in practice.

[lix] For an example, see the discussion of the limitation of cross-scanner MTR normalization in the following paragraph.

[lx] One current form of normalization is a two-point tissue normalization in which the WM and GM MTR values from a single subject on one MR scanner are mapped to the corresponding values obtained for that same subject on a different scanner. This mapping of MTR values could then be used for any subsequent subject being scanned on the second scanner. Other normalization methods

Another limitation of the method is that it has currently only been used for cross-sectional analyses. Of course, the method can be applied successively at multiple timepoints to provide an assessment of the evolution of the cortical pathology, but as was already discussed for the previous manuscript, specialized longitudinal methods are often preferable since they are more precise. The development of such a method is another worthy focus for future research.

include piecewise linear techniques that do not rely on tissue segmentation for their samples, but rather use predefined quantities from the histogram (e.g.,deciles, quartiles, median value, etc.) as the samples that are to be mapped from one subject to another.

6.3 CORTICAL NEUROIMAGING CORRELATES OF COGNITIVE DYSFUNCTION

The manuscript constituting Chapter 5 contained several important findings. First, regarding the primary hypotheses tested—that specific cognitive tasks are related to specific focal or regional cortical damage—it was found that widespread global cortical damage was primarily responsible for the declines in cognitive performance. From the scientific standpoint, these results are interesting as they imply that there may be a compensatory form of cortical reorganization that takes place to circumvent focal or regional damage. The reasoning being that the cognitive tests are purported to be specific to certain areas of the brain, yet we found that damage in those areas alone did not result in cognitive decline (i.e., the brain had rewired itself in order to maintain functionality); but, once the degree of damage had become widespread, the cognitive decline was inevitable (i.e., there were no more possibilities left for the brain to rewire itself).

Moreover, it was found that surface-based measures of MTR were better predictors of cognitive impairment than were measures of cortical thickness. If one accepts MTR as an adequate surrogate for myelin, then the results also suggest that cognitive impairment is driven by cortical demyelination, rather than cortical atrophy. From the clinical standpoint, the results also highlight the importance of using MTR as a standard sequence in clinical trials because it provides complementary information to measurements that are performed on a standard anatomical T1w MR image.

The secondary findings of the paper—that certain neuropsychological tests and MR metrics are better suited for detecting differences between normal controls and patients with MS—were not revolutionary, given that several of these studies have already been performed, and in many cases, with larger sample sizes (see section 2.4.3). Still, in terms of the neuropsychological tests themselves, we added several key discoveries to the field. First, the Symbol-Search test, which is easy to administer and exists in multiple forms to facilitate longitudinal analyses and to eliminate practice effects, is highly

sensitive to identifying patients with MS who are cognitively impaired. Second, the Tower of London test, a measure of higher-order executive function that we hypothesized would be an appropriate test but had yet to be studied in patients with MS, turned out to be highly insufficient for measuring their cognitive deficiencies. Findings such as these are invaluable to the clinical practitioners who are trying to screen for cognitive impairment in their patients and need to decide which tests to administer. However, it should be noted that, as most of the other tests we looked at displayed an adequate effect size, numerous tests can be used to screen patients. Furthermore, some (i.e., neuropsychologists) would even argue that such a wide variety of tests are not only possible, but necessary for screening purposes since the cognitive impairment profile of patients with MS is exceptionally broad. Such extensive testing, though, is often in direct conflict with the aim of the clinical neurologist who seeks the minimum set of tests required to screen for cognitive impairment, allowing for rapid administration and minimal taxation of the patient.

The manuscript also reported on the effect sizes for various MR metrics. Our findings indicated that, of the metrics tested, the surface-based MTR markers had the largest effect sizes—indirectly validating the surface-based framework[lxi] presented in the manuscript in Chapter 4. Unfortunately, we were not able to fully assess the role of focal cortical damage via the putative areas of subpial demyelination that we detected using the method of that manuscript because of our subject demographics.[lxii] While we were able to show that patients *without* focal cortical demyelinating lesions still exhibited cognitive impairment, future studies would benefit from a having a cohort with a normal distribution of subjects *with* focal cortical damage (i.e., a higher proportion of subjects with progressive stages of the disease than was included in the present study) for a more comprehensive analysis.

[lxi] The framework developed in the manuscript presented in Chapter 4 can be used not only to obtain a measurement of putative areas of subpial demyelination but, more generally, to provide a surface-based MTR measurement that is superior to a volume-based measurement.

[lxii] While it does not take away from the conclusions of this specific manuscript, it would have been a nice addition to the overall dissertation to have included an in-depth analysis of focal cortical pathology.

Oddly enough, the statistical analyses presented in this paper can simultaneously be considered a strength and weakness of the manuscript. When assessing the relationship between regional areas of cortical pathology (i.e., MTR or cortical thickness) and the specific neuropsychological tests (i.e., PASAT, WAIS-III Processing Speed Index, etc.), the most straightforward analysis is a simple univariate regression. The problem with such a method, despite being so commonly used in the literature, is that it does not take into account any global effects. Put more formally, there is a problem of collinearity between the covariates. By including a term of global cortical pathology into the models tested, we accounted for this problem and looked for any regional effects that were above and beyond whatever global effects were present. Had we not done so, the conclusions of the paper would have been quite the opposite—that specific regional damage was responsible for the performance on specific neuropsychological tasks—which in fact is what much of the previous literature has concluded.

At the same time, the statistical tests used can be considered a weakness of the paper simply because they are not as straightforward or as widely used and, in many instances, there may be (many) alternative statistical tests. For example, in years past, *r*-values and their corresponding *p*-values reigned supreme for evaluating the fit of a model. Today, alternatives for model evaluation such as the Akaike, Bayesian, Deviance, Focused, and Hannan-Quinn information criteria are also used, as are measures of the simplicity of the model.[lxiii] In addition, the method of model construction (i.e., deciding which components should or should not go into a model) can be based on theory, hypotheses, or on automated methods such as stepwise linear regression. Yet, even stepwise linear regression can be performed in a forward-wise manner (i.e., adding each term to the model one by one), a reverse manner (i.e., removing them), or a mixed manner, and the decision for inclusion/exclusion can be based on the *p*-value, or any of the information criterion already mentioned above. In short, there is no shortage of methods available for model evaluation. And, since the theories on which method

[lxiii] Simpler models are (currently) assumed to be better.

should be used are ever-changing, this abundance unfortunately leads to readily contentious conclusions simply as a result of the choice of statistical analysis method.[lxiv]

That we only reported cross-sectional data in this paper can also be considered a weakness. Often, longitudinal data (i.e., how a patient changes over time) is of greater interest to the clinician. Luckily, the study is ongoing, and many of the subjects are reaching the two-year mark when they will return for follow-up scanning and neuropsychological testing. This longitudinal insight into the progression of cognitive impairment and the associated MR characteristics that best predict such progression will prove invaluable. As a result, it will be possible to develop an MR screening technique to identify those patients at the highest risk of developing cognitive problems in the hope that this early detection can lead to earlier treatment and, thus, a delay in disease progression.

With such a wealth of data having been (and still being) acquired, the project will go on to fuel many studies of cognition and MS. Even though the cognitive measures tested were just too numerous for us to examine each in depth within the current manuscript, we have already begun to look into some other specific tasks. For example, postmortem immunohistochemical studies have shown that hippocampal demyelination is both frequent and extensive in patients with MS. Yet, the imaging characteristics of hippocampal demyelination and, more importantly, their relationship to memory impairment are still to be studied. We recently submitted an abstract to the 2013 ISMRM conference wherein we investigated hippocampal atrophy (assessed by volume) and hippocampal demyelination (assessed via MTR) and their relationship to memory impairment [412]. Our data corroborated the postmortem findings of hippocampal involvement in patients with MS and extended them by showing that the pathology most relevant for negative cognitive effects is driven by demyelination rather than neuronal loss. We also found that working memory and, to a lesser extent, visual immediate memory are linked to hippocampal MTR and that whatever contribution

[lxiv] Our choices for the statistical analysis are enumerated within the manuscript itself and so are not reproduced here.

hippocampal volume has to memory performance is the result of global atrophy and not hippocampal specific volume loss. Though still ongoing, this work has thus far yielded many scientifically interesting results and will be an area of active research in the future. A copy of the abstract has been included for reference in Appendix 6B.

6.4 OVERARCHING CONCLUSIONS

This dissertation was concerned with in vivo imaging methods for quantifying GM pathology in patients with MS. Chapter 3 revealed the strengths and weaknesses of the currently available techniques, identifying which methods are best suited for specific measurements. Chapter 4 introduced a novel surface-based technique for imaging putative areas of subpial demyelination, the most common form of cortical lesion seen on postmortem analysis that had yet to be captured via in vivo imaging techniques. Chapter 5 combined the knowledge gleaned from Chapter 3 and the new methodology introduced in Chapter 4 to investigate to the best of our ability GM pathology in patients with MS, and then looked at the relationship between these imaging characteristics and the patients' cognitive performance. Importantly, each manuscript accomplished its aims while still being clinically feasible, using images typically obtained in the clinical setting and at the accessible field strengths of 1.5 T and 3 T.

To appreciate the contribution of this dissertation to the imaging community and to our understanding of MS, consider the state of WM lesion identification at its infancy. First, the most conspicuous hyperintense T2w lesions were observable. The research community then charged forward with methods to better identify these lesions, and more importantly, they sought to investigate the relationship of these lesions to the clinical outcome of their patients. The result? They were found wanting. And so the quest to image more specific WM lesion pathology began. Gadolinium-enhancing lesions and T1w hypointense lesions were then introduced into the equation in the hope that these markers of disease activity and disease burden would explain patient outcome. Despite providing great insight into the disease, they were also found wanting. As new modalities became available and more sophisticated techniques were developed, new in vivo measures of pathology (e.g., WM atrophy, NAWM integrity, GM atrophy, etc.) were explored. Now we find ourselves a full century after the presence of different GM lesions had been established in the disease—and subsequently forgotten—trying to image more than just the GM volume that had been focused on so

intently over the last decade. The contribution of this thesis, namely, surface-based measures of cortical MTR, is simply the next step in this evolution of knowledge. Even with these new methodologies, we will, once again without a doubt, find ourselves wanting; however, the insight gained from this thesis and its influence on the course of future studies will play an invaluable role in shaping what we will come to learn about this devastating disease.

APPENDIX 6A

Poster #2210 presented at the International Society for Magnetic Resonance in Medicine's 20th annual meeting (ISMRM 2012) [411].

Combining SIENA and SIENAx for improved quantification of grey and white matter atrophy

Mishkin Derakhshan, Sridar Narayanan, D. Louis Collins, and Douglas L. Arnold
Montreal Neurological Institute, McGill University, Montreal, QC, Canada

cGM = cortical grey matter, **GM** = grey matter, **WM** = white matter, **WB** = whole brain, **PBVC** = percent brain volume change, **TP1** = timepoint 1

Summary and Conclusions

• For tissue specific atrophy, we proposed a method that combines the precision of SIENA with the tissue classification feature of SIENAx

• Compared with repeated cross-sectional measures (i.e., SIENAx at each timepoint), our method:
  • Is more accurate (see Test 1)
  • has better reproducibility (see Test 2)
  • is applicable to clinically typical (1x1x3 mm^3) T1w images (see Test 3)

The Proposed Method: SIENA + SIENAx

GM (or cGM, or WM) atrophy is calculated as:

$$GM\ atrophy_{TP1\to TP2} = 100\% \times \frac{GM\ volume_{TP2} - GM\ volume_{TP1}}{GM\ volume_{TP1}}$$

= calculated volume
= output of SIENA
= output of SIENAx (normalized volumes)

where

$$GM\ volume_{TP1} = SIENAx\ GM\ volume_{TP1}$$

$$GM\ volume_{TP2} = SIENAx\ WB\ volume_{TP1} \times \left(1 + \frac{SIENA\ PBVC_{TP1\to TP2}}{100\%}\right) \times \frac{SIENAx\ GM\ volume_{TP2}}{SIENAx\ WB\ volume_{TP2}}$$

Test 1 – Simulated Atrophy Dataset (N=10)

To evaluate the accuracy of each technique, a known amount of atrophy was simulated by creating an "atrophied" version of a T1w scan (TE=11ms, TR=30, FA=30°, 1.5T SIEMENS Sonata Vision) by replacing the brain with a scaled version of itself. Only the brain was scaled as skull size is used as a normalization factor within the method. A global scaling of ~86% was achieved using a linear transform with a 0.95 scaling factor in each of the x, y, and z dimensions.

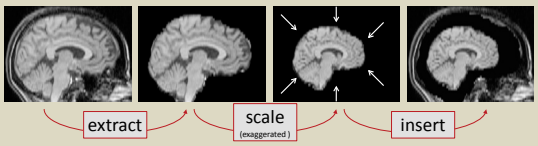

Figure 1 – The original image (leftmost) is compared with the simulated atrophy image (rightmost).
Note: scaling is exaggerated for visualization purposes

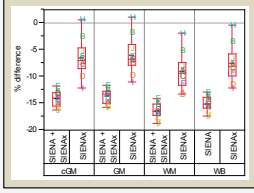

Figure 2 – The percent difference between a regular and simulated-atrophic brain. Note how our proposed method (**SIENA + SIENAx**) has reduced variance and is closer to the theoretical value (-14%) than repeated cross-sectional measurements (**SIENAx**).

Test 2 – Scan-Rescan Dataset (N=10)

A scan-rescan dataset of 10 healthy controls was used to assess reproducibility. For each scan-rescan pair, the subject was scanned, removed, repositioned, and then immediately rescanned.

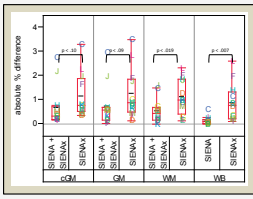

Figure 3 – The absolute percent difference between the scan-rescan pair for each tissue type. Note the reduced error for our proposed method (**SIENA + SIENAx**) compared with repeated cross-sectional measurements (**SIENAx**).

Test 3 – Slice Thickness

For each of the datasets (Test 1, Test 2, and Application), there were no observable quantitative or qualitative improvements to using a 1.5mm slice thickness versus 3mm (in-plane resolution: 0.98mm x 0.98mm) for any of the tissue volumes (cGM, GM, WM, and WB).

Application – Multi-Centre Clinical Trial Dataset (N=59)

We applied the method to a clinically relevant dataset with an expected but unknown degree of atrophy, comprising subjects with relapsing-remitting multiple sclerosis scanned 3 years apart. Compared with repeated cross-sectional measurements, our method had slightly lower variances that trended towards significance. We found a mean WB atrophy of -1.45% (SD=1.13), driven more by GM atrophy (mean=-2.14%, SD=2.42%) than by WM atrophy (mean=-0.81%, SD=1.87%).

APPENDIX 6B

Abstract #2384 submitted to the International Society for Magnetic Resonance in Medicine's 21st annual meeting (ISMRM 2013) [412].

HIPPOCAMPAL MAGNETIZATION TRANSFER RATIO AND NOT HIPPOCAMPAL ATROPHY BEST EXPLAINS MEMORY DYSFUNCTION IN PATIENTS WITH MULTIPLE SCLEROSIS

MISHKIN DERAKHSHAN[1], GABRIEL LEONARD[1], DAVID ARAUJO[1], D. LOUIS COLLINS[1], DOUGLAS L. ARNOLD[1], AND SRIDAR NARAYANAN[1]

[1]Montreal Neurological Institute, McGill University, Montreal, QC, Canada

BACKGROUND

Postmortem immunohistochemical studies have shown that hippocampal (HC) demyelination is frequent and extensive in patients with multiple sclerosis (MS) and that, while there is a significant decrease in synaptic density, the demyelinated HC have minimal neuronal loss[1,2]. Recent MRI studies have correlated the number of HC lesions[3], as well as HC atrophy[4], with memory dysfunction (i.e., verbal and visuospatial learning/immediate memory) in patients with MS.

PURPOSE

To the best of our knowledge, no studies have looked at the relationship between MRI correlates of HC demyelination and memory dysfunction in patients with MS. Thus, we proposed to (i)confirm the postmortem findings of HC demyelination in vivo using measures of magnetization transfer ratio (MTR), a marker of demyelination, and (ii)investigate the relationship of both HC demyelination and neuronal loss (atrophy) with memory dysfunction in patients with MS.

METHODS

SUBJECTS

A total of 26 patients with clinically definite MS were recruited from the MS clinic at the Montreal Neurological Hospital. To be included in the study, patients had to display some degree of cognitive disability as indicated by a score of one or greater on the cerebral functional system subscale of Kurtzke's Expanded Disability Status Scale (EDSS). Fifteen healthy volunteers were recruited based on age, gender, and education level to serve as the normal control (NC) group.

IMAGE ACQUISITION

All participants were scanned on a 3T MAGNETOM Trio TIM scanner (Siemens, Erlangen, Germany). A 3Dfast low-angle shot (FLASH) gradient-recalled echo (GRE) sequence [TR=20ms, TE=5ms, FA=27, 256x256 matrix, 256mm field of view, 1mm slice thickness, 1 signal average, voxel size= 1x1x1mm3] provided 1mm isotropic T1-weighted (T1w) anatomical MRI.A pair of 3D FLASH sequences [TR=33ms, TE=3.81ms, FA=10], with geometry matched to that of the high-resolution T1w sequence, was acquired with and without an off-resonance magnetization transfer pulse [Gaussian pulse, offset frequency=1,200Hz, FA=500°, pulse length=9,984μs, bandwidth=192Hz]; MTR maps were created from these images after co-registration as voxel-wise percent difference maps.

IMAGE PROCESSING

Automated HC labels created using the label fusion-based template library scheme[5] were manually corrected as needed by a single trained neuroradiologist (DA). Cortical grey matter (cGM) maps were obtained via SIENAx[7]. All volumes were normalized using the same normalization factor obtained from SIENAx. The mean MTR values inside the HC and cGM labels were calculated.

NEUROPSYCHOLOGICAL TESTING

All participants completed a detailed neuropsychological assessment within an average of nine days of their MR examination (median = two days). Included in the battery was the Wechsler Memory Scale (third edition; WMS-III), which provides indices of immediate, general/delayed, and working memory. Since HC pathology has a reported effect on immediate memory, we also included visual and auditory immediate subindices that make up the immediate memory index in the analysis.

STATISTICS

Differences between NC and MS groups for the four MR metrics and five memory indices were calculated with a one-tailed t-test and are accompanied by Cohen's d as a measure of effect size. Univariate analysis of both HC volumes and HC MTR values was performed with each of the five memory indices examined. To assess the independent contributions of each, partial r2 values were obtained for each component using a multivariate analysis. Further stepwise linear regression models were explored in an attempt to identify the best imaging correlate of memory impairment. Bonferroni multiple-comparison correction was applied where appropriate.

RESULTS

Left and right HC volumes and MTR values were combined as no significant differences in lateralization were found ($F=1.04$, $p=.31$, and $F=2.51$, $p=.12$, respectively). With the exception of the auditory immediate subindex, medium to large effect sizes ($0.61 <$ Cohen's $d < 1.11$) were observed for all MR and memory metrics between NC and MS groups. After correction for multiple comparisons, the only differences that remained significant were in the normalized cGM volume ($p=.03$) and the visual immediate memory subindex ($p<.01$).

Univariate correlations revealed a relationship between HC MTR and working memory ($r^2=.22$, $p=.04$) as well as between visual immediate memory and both HC MTR ($r^2=.32$, $p<.01$) and HC volume ($r^2=.29$, $p<.01$). To assess whether the relationships truly reflected

localized HC damage and not global damage, cGM MTR and volume were used as covariates. We found that the independent contributions for HC volume disappeared entirely, whereas localized HC MTR still predicted working memory performance (partial r^2=.23, p=.04). Two approaches were used to assess the independent contributions of HC MTR and HC volume to the relationships with the memory findings. In the first analysis, each was assessed independently in a model where it was found that both HC MTR and volume contribute independently to visual immediate memory performance(partial r^2 =.22, p=.04, and partial r^2=.23, p=.04, respectively). Though not significant after Bonferroni correction, there was still a trend for the independent contribution of HC MTR to working memory performance (partial r^2=.19, p=.08). The second analysis used stepwise linear regression to predict the best models for each memory index. We found that working memory was best explained with HC MTR and normalized cGM volume; delayed memory with HC MTR; immediate memory with HC MTR and HC volume; auditory immediate memory with HC MTR only; and visual immediate memory with HC MTR and normalized cGM and HC volumes. In all cases, HC MTR was the first parameter to enter the model, thus accounting for the most variance, and always remained in the final model.

DISCUSSION AND CONCLUSIONS

Our data corroborate the postmortem findings of HC involvement in patients with MS and extend these by showing that the pathology most relevant for negative cognitive effects is driven by demyelination (as measured with MTR) rather than neuronal loss (as measured with HC volume). We found that working memory and, to a lesser extent, visual immediate memory are linked to HC MTR and that whatever contribution HC volume has to memory performance is the result of global atrophy and not HC specific volume loss. We found that visuospatial memory is affected by HC pathology, which supports previous findings[3], but that auditory immediate memory was preserved, in contrast to other studies[4,8]. These results highlight the importance of including measures

related to demyelination as well as atrophy for assessing cognitive dysfunction in patients with MS.

REFERENCES

1. Dutta, R., et al., Ann Neurol, 2011. 69(3): p. 445-54.
2. Geurts, J.J., et al., J Neuropathol Exp Neurol, 2007. 66(9): p. 819-27.
3. Roosendaal, S.D., et al., Mult Scler, 2009. 15(6): p. 708-14.
4. Sicotte, N.L., et al., Brain, 2008. 131(Pt 4): p. 1134-41.
5. Collins, D.L. and J.C. Pruessner, Neuroimage, 2010. 52(4): p. 1355-66.
6. Francis, S.J., M.Sc. Thesis, 2004, McGill University: Montreal.
7. Smith, S.M., et al., Neuroimage, 2002. 17(1): p. 479-89.
8. Fuentes, A., et al., J Int Neuropsychol Soc, 2012. 18(3): p. 471-80.

REFERENCES

1. CTV.ca News. *W5 investigates intriguing new theory about MS*. 2009 [cited 2012 July 23]; Available from: http://www.ctvnews.ca/w5-investigates-intriguing-new-theory-about-ms-1.456316.
2. McAlpine, D. and A. Compston, *McAlpine's multiple sclerosis*. 4th ed2006, Edinburgh ; Philadelphia, PA: Churchill Livingstone/Elsevier. x, 982 p.
3. Murray, T.J., *Multiple sclerosis : the history of a disease*2005, New York: Demos Medical Pub. xi, 580 p.
4. Carswell, R., *Pathological anatomy. : Illustrations of the elementary forms of disease*1838, London: Longman, Orme, Brown, Green and Longman. 218 p., 48 col. pl.
5. Compston, A. and A. Coles, *Multiple sclerosis.* Lancet, 2002. **359**(9313): p. 1221-31.
6. Multiple Sclerosis Society of Canada. *About MS*. [cited 2012 July 24]; Available from: http://mssociety.ca/en/information/default.htm.
7. Beck, C.A., et al., *Regional variation of multiple sclerosis prevalence in Canada.* Mult Scler, 2005. **11**(5): p. 516-9.
8. Orton, S.M., et al., *Sex ratio of multiple sclerosis in Canada: a longitudinal study.* Lancet Neurol, 2006. **5**(11): p. 932-6.
9. Poser, C.M. and V.V. Brinar, *Diagnostic criteria for multiple sclerosis: an historical review.* Clin Neurol Neurosurg, 2004. **106**(3): p. 147-58.
10. Sadovnick, A.D., et al., *Age of onset in concordant twins and other relative pairs with multiple sclerosis.* Am J Epidemiol, 2009. **170**(3): p. 289-96.
11. Banwell, B., et al., *Clinical features and viral serologies in children with multiple sclerosis: a multinational observational study.* Lancet Neurol, 2007. **6**(9): p. 773-81.
12. Polliack, M.L., Y. Barak, and A. Achiron, *Late-onset multiple sclerosis.* J Am Geriatr Soc, 2001. **49**(2): p. 168-71.
13. Ebers, G.C., et al., *A population-based study of multiple sclerosis in twins.* N Engl J Med, 1986. **315**(26): p. 1638-42.
14. Willer, C.J., et al., *Twin concordance and sibling recurrence rates in multiple sclerosis.* Proc Natl Acad Sci U S A, 2003. **100**(22): p. 12877-82.
15. Kurtzke, J.F., *Multiple sclerosis in time and space--geographic clues to cause.* J Neurovirol, 2000. **6 Suppl 2**: p. S134-40.
16. Kennedy, J., et al., *Age at onset of multiple sclerosis may be influenced by place of residence during childhood rather than ancestry.* Neuroepidemiology, 2006. **26**(3): p. 162-7.
17. Ebers, G.C., et al., *Parent-of-origin effect in multiple sclerosis: observations in half-siblings.* Lancet, 2004. **363**(9423): p. 1773-4.
18. Willer, C.J., et al., *Timing of birth and risk of multiple sclerosis: population based study.* BMJ, 2005. **330**(7483): p. 120.
19. Ho, S.L., L. Alappat, and A.B. Awad, *Vitamin d and multiple sclerosis.* Crit Rev Food Sci Nutr, 2012. **52**(11): p. 980-7.
20. Lublin, F.D. and S.C. Reingold, *Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis.* Neurology, 1996. **46**(4): p. 907-11.

21. Azevedo, F.A., et al., *Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain.* J Comp Neurol, 2009. **513**(5): p. 532-41.

22. *http://en.wikipedia.org/wiki/Neuron*. Wikipedia [cited 2012 Nov 03].

23. Raine, C.S., L. Scheinberg, and J.M. Waltz, *Multiple sclerosis. Oligodendrocyte survival and proliferation in an active established lesion.* Lab Invest, 1981. **45**(6): p. 534-46.

24. Dawson, J.W., *The Histology of Disseminated Sclerosis.* Transactions of the Royal Society of Edinburgh, 1914. **50**(03): p. 517-740.

25. Trapp, B.D., et al., *Axonal transection in the lesions of multiple sclerosis.* N Engl J Med, 1998. **338**(5): p. 278-85.

26. Trapp, B.D. and K.A. Nave, *Multiple sclerosis: an immune or neurodegenerative disorder?* Annu Rev Neurosci, 2008. **31**: p. 247-69.

27. Taylor, E.W., *Zur pathologischen Anatomie der multiplen Sklerose.* Journal of Neurology, 1894. **5**(1): p. 1-26.

28. Sander, M., *Hirnrindenbefunde bei multipler Sklerose.* European Neurology, 1898. **4**(6): p. 427-436.

29. Dinkler, M., *Zur Kasuistik der multiplen Herdsklerose des Gehirns und Rückenmarks.* Journal of Neurology, 1904. **26**(3): p. 233-247.

30. Brownell, B. and J.T. Hughes, *The distribution of plaques in the cerebrum in multiple sclerosis.* J Neurol Neurosurg Psychiatry, 1962. **25**: p. 315-20.

31. Lumsden, C.E., *The neuropathology of multiple sclerosis*, in *Handbook of Clinical Neurology. Multiple Sclerosis and Other Demyelinating Diseases*, P.J. Vinken and G.W. Bruyn, Editors. 1970: Amsterdam. p. 217-309.

32. Peterson, J.W., et al., *Transected neurites, apoptotic neurons, and reduced inflammation in cortical multiple sclerosis lesions.* Ann Neurol, 2001. **50**(3): p. 389-400.

33. Bo, L., et al., *Subpial demyelination in the cerebral cortex of multiple sclerosis patients.* J Neuropathol Exp Neurol, 2003. **62**(7): p. 723-32.

34. Peters, A. and E.G. Jones, *Cerebral cortex*1984, New York: Plenum Press. v.

35. Vogt, O., *Die myeloarchitektonische Felderung des menschlichen Stirnhirns.* J Psychol Neurol, 1910. **15**: p. 221.

36. Trapp, B.D., R. Ransohoff, and R. Rudick, *Axonal pathology in multiple sclerosis: relationship to neurologic disability.* Curr Opin Neurol, 1999. **12**(3): p. 295-302.

37. Kidd, D., et al., *Cortical lesions in multiple sclerosis.* Brain, 1999. **122 (Pt 1)**: p. 17-26.

38. Bo, L., et al., *Intracortical multiple sclerosis lesions are not associated with increased lymphocyte infiltration.* Mult Scler, 2003. **9**(4): p. 323-31.

39. Kutzelnigg, A., et al., *Cortical demyelination and diffuse white matter injury in multiple sclerosis.* Brain, 2005. **128**(Pt 11): p. 2705-12.

40. Serafini, B., et al., *Detection of ectopic B-cell follicles with germinal centers in the meninges of patients with secondary progressive multiple sclerosis.* Brain Pathol, 2004. **14**(2): p. 164-74.

41. Lucchinetti, C.F., et al., *Inflammatory cortical demyelination in early multiple sclerosis.* N Engl J Med, 2011. **365**(23): p. 2188-97.

42. Magliozzi, R., et al., *Meningeal B-cell follicles in secondary progressive multiple sclerosis associate with early onset of disease and severe cortical pathology.* Brain, 2007. **130**(Pt 4): p. 1089-104.

43. Popescu, B.F. and C.F. Lucchinetti, *Meningeal and cortical grey matter pathology in multiple sclerosis.* BMC Neurol, 2012. **12**: p. 11.

44. Serafini, B., et al., *Dendritic cells in multiple sclerosis lesions: maturation stage, myelin uptake, and interaction with proliferating T cells.* J Neuropathol Exp Neurol, 2006. **65**(2): p. 124-41.

45. Cudrici, C., et al., *Dendritic cells are abundant in non-lesional gray matter in multiple sclerosis.* Exp Mol Pathol, 2007. **83**(2): p. 198-206.

46. Ludewig, B., et al., *Dendritic cells in autoimmune diseases.* Curr Opin Immunol, 2001. **13**(6): p. 657-62.

47. *http://en.wikipedia.org/wiki/Meninges*. [cited 2012 Nov 04].

48. Vercellino, M., et al., *Grey matter pathology in multiple sclerosis.* J Neuropathol Exp Neurol, 2005. **64**(12): p. 1101-7.

49. Wegner, C., et al., *Neocortical neuronal, synaptic, and glial loss in multiple sclerosis.* Neurology, 2006. **67**(6): p. 960-7.

50. Gray, E., et al., *Elevated activity and microglial expression of myeloperoxidase in demyelinated cerebral cortex in multiple sclerosis.* Brain Pathol, 2008. **18**(1): p. 86-95.

51. Howell, O.W., et al., *Meningeal inflammation is widespread and linked to cortical pathology in multiple sclerosis.* Brain, 2011. **134**(Pt 9): p. 2755-71.

52. Schmierer, K., et al., *High field (9.4 Tesla) magnetic resonance imaging of cortical grey matter lesions in multiple sclerosis.* Brain, 2010. **133**(Pt 3): p. 858-67.

53. Popescu, B.F., et al., *A case of multiple sclerosis presenting with inflammatory cortical demyelination.* Neurology, 2011. **76**(20): p. 1705-10.

54. Geurts, J.J., et al., *Cortical lesions in multiple sclerosis: combined postmortem MR imaging and histopathology.* AJNR Am J Neuroradiol, 2005. **26**(3): p. 572-7.

55. Allen, I.V. and S.R. McKeown, *A histological, histochemical and biochemical study of the macroscopically normal white matter in multiple sclerosis.* J Neurol Sci, 1979. **41**(1): p. 81-91.

56. Seewann, A., et al., *Diffusely abnormal white matter in chronic multiple sclerosis: imaging and histopathologic analysis.* Arch Neurol, 2009. **66**(5): p. 601-9.

57. Evangelou, N., et al., *Quantitative pathological evidence for axonal loss in normal appearing white matter in multiple sclerosis.* Ann Neurol, 2000. **47**(3): p. 391-5.

58. Bjartmar, C., et al., *Axonal loss in normal-appearing white matter in a patient with acute MS.* Neurology, 2001. **57**(7): p. 1248-52.

59. Kutzelnigg, A., et al., *Widespread demyelination in the cerebellar cortex in multiple sclerosis.* Brain Pathol, 2007. **17**(1): p. 38-44.

60. Gilmore, C.P., et al., *Regional variations in the extent and pattern of grey matter demyelination in multiple sclerosis: a comparison between the cerebral cortex, cerebellar cortex, deep grey matter nuclei and the spinal cord.* J Neurol Neurosurg Psychiatry, 2009. **80**(2): p. 182-7.

61. Cifelli, A., et al., *Thalamic neurodegeneration in multiple sclerosis.* Ann Neurol, 2002. **52**(5): p. 650-3.

62. Nishimura, D.G., *Principles of Magnetic Resonance Imaging*1996: Stanford University.

63. Lauterbur, P.C., *Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance.* Nature, 1973. **242**(5394): p. 190-191.

64. Henkelman, R.M., G.J. Stanisz, and S.J. Graham, *Magnetization transfer in MRI: a review.* NMR Biomed, 2001. **14**(2): p. 57-64.

65. Hu, B.S., et al., *Pulsed saturation transfer contrast.* Magn Reson Med, 1992. **26**(2): p. 231-40.

66. Sled, J.G. and G.B. Pike, *Quantitative imaging of magnetization transfer exchange and relaxation properties in vivo using MRI.* Magn Reson Med, 2001. **46**(5): p. 923-31.

67. Harrison, R., M.J. Bronskill, and R.M. Henkelman, *Magnetization transfer and T2 relaxation components in tissue.* Magn Reson Med, 1995. **33**(4): p. 490-6.

68. Koenig, S.H., et al., *Magnetic cross-relaxation among protons in protein solutions.* Biochemistry, 1978. **17**(20): p. 4348-58.

69. Zhong, J.H., J.C. Gore, and I.M. Armitage, *Quantitative studies of hydrodynamic effects and cross-relaxation in protein solutions and tissues with proton and deuteron longitudinal relaxation times.* Magn Reson Med, 1990. **13**(2): p. 192-203.

70. Brady, S.T., et al., *Basic neurochemistry : principles of molecular, cellular and medical neurobiology*. 8th ed2012, Amsterdam ; Boston: Elsevier Academic Press. xxiv, 1096 p.

71. Cole, W.C., A.D. LeBlanc, and S.G. Jhingran, *The origin of biexponential T2 relaxation in muscle water.* Magn Reson Med, 1993. **29**(1): p. 19-24.

72. MacKay, A., et al., *Insights into brain microstructure from the T2 distribution.* Magn Reson Imaging, 2006. **24**(4): p. 515-25.

73. Menon, R.S. and P.S. Allen, *Application of continuous relaxation time distributions to the fitting of data from model systems and excised tissue.* Magn Reson Med, 1991. **20**(2): p. 214-27.

74. Menon, R.S., M.S. Rusinko, and P.S. Allen, *Proton relaxation studies of water compartmentalization in a model neurological system.* Magn Reson Med, 1992. **28**(2): p. 264-74.

75. MacKay, A., et al., *In vivo visualization of myelin water in brain by magnetic resonance.* Magn Reson Med, 1994. **31**(6): p. 673-7.

76. Stewart, W.A., et al., *Spin-spin relaxation in experimental allergic encephalomyelitis. Analysis of CPMG data using a non-linear least squares method and linear inverse theory.* Magn Reson Med, 1993. **29**(6): p. 767-75.

77. Does, M.D. and R.E. Snyder, *T2 relaxation of peripheral nerve measured in vivo.* Magn Reson Imaging, 1995. **13**(4): p. 575-80.

78. Wachowicz, K. and R.E. Snyder, *Assignment of the T(2) components of amphibian peripheral nerve to their microanatomical compartments.* Magn Reson Med, 2002. **47**(2): p. 239-45.

79. Schumacher, G.A., et al., *Problems of experimental trials of therapy in multiple sclerosis: Report by the panel on the evaluation of experimental trials of therapy in multiple sclerosis.* Annals of the New York Academy of Sciences, 1965. **122**(1): p. 552-568.

80. Poser, C.M., et al., *New diagnostic criteria for multiple sclerosis: guidelines for research protocols.* Ann Neurol, 1983. **13**(3): p. 227-31.

81. McDonald, W.I., et al., *Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis.* Ann Neurol, 2001. **50**(1): p. 121-7.

82. Polman, C.H., et al., *Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria.* Ann Neurol, 2011. **69**(2): p. 292-302.

83. Ormerod, I.E., et al., *The role of NMR imaging in the assessment of multiple sclerosis and isolated neurological lesions. A quantitative study.* Brain, 1987. **110 (Pt 6)**: p. 1579-616.

84. Filippi, M., et al., *Association between pathological and MRI findings in multiple sclerosis.* Lancet Neurol, 2012. **11**(4): p. 349-60.

85. McFarland, H.F., et al., *Using gadolinium-enhanced magnetic resonance imaging lesions to monitor disease activity in multiple sclerosis.* Ann Neurol, 1992. **32**(6): p. 758-66.

86. Sormani, M.P., et al., *Statistical power of MRI monitored trials in multiple sclerosis: new data and comparison with previous results.* J Neurol Neurosurg Psychiatry, 1999. **66**(4): p. 465-9.

87. Hasan, K.M., et al., *Multi-modal quantitative MRI investigation of brain tissue neurodegeneration in multiple sclerosis.* J Magn Reson Imaging, 2012. **35**(6): p. 1300-11.

88. Narayanan, S., et al., *Axonal injury in the cerebral normal-appearing white matter of patients with multiple sclerosis is related to concurrent demyelination in lesions but not to concurrent demyelination in normal-appearing white matter.* Neuroimage, 2006. **29**(2): p. 637-42.

89. Barkhof, F., G.B. Karas, and M.A. van Walderveen, *T1 hypointensities and axonal loss.* Neuroimaging Clin N Am, 2000. **10**(4): p. 739-52 ,ix.

90. Chen, J.T., et al., *Magnetization transfer ratio evolution with demyelination and remyelination in multiple sclerosis lesions.* Ann Neurol, 2008. **63**(2): p. 254-62.

91. Chen, J.T., et al., *Voxel-based analysis of the evolution of magnetization transfer ratio to quantify remyelination and demyelination with histopathological validation in a multiple sclerosis lesion.* Neuroimage, 2007. **36**(4): p. 1152-8.

92. Dousset, V., et al., *Experimental allergic encephalomyelitis and multiple sclerosis: lesion characterization with magnetization transfer imaging.* Radiology, 1992. **182**(2): p. 483-91.

93. Giacomini, P.S., et al., *Measuring demyelination and remyelination in acute multiple sclerosis lesion voxels.* Arch Neurol, 2009. **66**(3): p. 375-81.

94. Pike, G.B., et al., *Multiple sclerosis: magnetization transfer MR imaging of white matter before lesion appearance on T2-weighted images.* Radiology, 2000. **215**(3): p. 824-30.

95. Schmierer, K., et al., *Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain.* Ann Neurol, 2004. **56**(3): p. 407-15.

96. Brochet, B. and V. Dousset, *Pathological correlates of magnetization transfer imaging abnormalities in animal models and humans with multiple sclerosis.* Neurology, 1999. **53**(5 Suppl 3): p. S12-7.

97. Schmierer, K., et al., *Quantitative magnetic resonance of postmortem multiple sclerosis brain before and after fixation.* Magn Reson Med, 2008. **59**(2): p. 268-77.

98. Vavasour, I.M., et al., *Is the magnetization transfer ratio a marker for myelin in multiple sclerosis?* J Magn Reson Imaging, 2011. **33**(3): p. 713-8.

99. Schmierer, K., et al., *Quantitative magnetization transfer imaging in postmortem multiple sclerosis brain.* J Magn Reson Imaging, 2007. **26**(1): p. 41-51.

100. Barkhof, F., et al., *Remyelinated lesions in multiple sclerosis: magnetic resonance image appearance.* Arch Neurol, 2003. **60**(8): p. 1073-81.

101. Miller, D.H., F. Barkhof, and J.J. Nauta, *Gadolinium enhancement increases the sensitivity of MRI in detecting disease activity in multiple sclerosis.* Brain, 1993. **116 (Pt 5)**: p. 1077-94.

102. Thompson, A.J., et al., *Major differences in the dynamics of primary and secondary progressive multiple sclerosis.* Ann Neurol, 1991. **29**(1): p. 53-62.

103. Willoughby, E.W., et al., *Serial magnetic resonance scanning in multiple sclerosis: a second prospective study in relapsing patients.* Ann Neurol, 1989. **25**(1): p. 43-9.

104. Kappos, L., et al., *Predictive value of gadolinium-enhanced magnetic resonance imaging for relapse rate and changes in disability or impairment in multiple sclerosis: a meta-analysis. Gadolinium MRI Meta-analysis Group.* Lancet, 1999. **353**(9157): p. 964-9.

105. Cleveland Clinic. Fox, R.J. *Multiple Sclerosis*. Disease Management Project 2010 August 1, 2010 [cited 2012 August 17]; Available from:

http://www.clevelandclinicmeded.com/medicalpubs/diseasemanagement/neurology/multiple_sclerosis/.

106. Fox, R.J., et al., *Multiple sclerosis: advances in understanding, diagnosing, and treating the underlying disease.* Cleve Clin J Med, 2006. **73**(1): p. 91-102.

107. Kurtzke, J.F., *Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS).* Neurology, 1983. **33**(11): p. 1444-52.

108. Kurtzke, J.F., *Natural history and clinical outcome measures for multiple sclerosis studies. Why at the present time does EDSS scale remain a preferred outcome measure to evaluate disease evolution?* Neurol Sci, 2000. **21**(6): p. 339-41.

109. Goodin, D.S., *Magnetic resonance imaging as a surrogate outcome measure of disability in multiple sclerosis: have we been overly harsh in our assessment?* Ann Neurol, 2006. **59**(4): p. 597-605.

110. Hobart, J., J. Freeman, and A. Thompson, *Kurtzke scales revisited: the application of psychometric methods to clinical intuition.* Brain, 2000. **123 (Pt 5)**: p. 1027-40.

111. Sharrack, B., et al., *The psychometric properties of clinical rating scales used in multiple sclerosis.* Brain, 1999. **122 (Pt 1)**: p. 141-59.

112. Barkhof, F., *MRI in multiple sclerosis: correlation with expanded disability status scale (EDSS).* Mult Scler, 1999. **5**(4): p. 283-6.

113. Zivadinov, R. and T.P. Leist, *Clinical-magnetic resonance imaging correlations in multiple sclerosis.* J Neuroimaging, 2005. **15**(4 Suppl): p. 10S-21S.

114. Caramanos, Z., et al., *Large, nonplateauing relationship between clinical disability and cerebral white matter lesion load in patients with multiple sclerosis.* Arch Neurol, 2012. **69**(1): p. 89-95.

115. Truyen, L., et al., *Accumulation of hypointense lesions ("black holes") on T1 spin-echo MRI correlates with disease progression in multiple sclerosis.* Neurology, 1996. **47**(6): p. 1469-76.

116. Bakshi, R., et al., *Fluid-attenuated inversion recovery magnetic resonance imaging detects cortical and juxtacortical multiple sclerosis lesions.* Arch Neurol, 2001. **58**(5): p. 742-8.

117. Boggild, M.D., et al., *Cortical plaques visualised by fluid-attenuated inversion recovery imaging in relapsing multiple sclerosis.* Neuroradiology, 1996. **38 Suppl 1**: p. S10-3.

118. Newcombe, J., et al., *Histopathology of multiple sclerosis lesions detected by magnetic resonance imaging in unfixed postmortem central nervous system tissue.* Brain, 1991. **114 (Pt 2)**: p. 1013-23.

119. Geurts, J.J., et al., *Intracortical lesions in multiple sclerosis: improved detection with 3D double inversion-recovery MR imaging.* Radiology, 2005. **236**(1): p. 254-60.

120. Pouwels, P.J., et al., *Human gray matter: feasibility of single-slab 3D double inversion-recovery high-spatial-resolution MR imaging.* Radiology, 2006. **241**(3): p. 873-9.

121. Bagnato, F., et al., *In vivo detection of cortical plaques by MR imaging in patients with multiple sclerosis.* AJNR Am J Neuroradiol, 2006. **27**(10): p. 2161-7.

122. Bluestein, K.T., et al., *Detecting cortical lesions in multiple sclerosis at 7 T using white matter signal attenuation.* Magn Reson Imaging, 2012. **30**(7): p. 907-15.

123. de Graaf, W.L., et al., *Clinical application of multi-contrast 7-T MR imaging in multiple sclerosis: increased lesion detection compared to 3 T confined to grey matter.* Eur Radiol, 2012.

124. de Graaf, W.L., et al., *Lesion detection at seven Tesla in multiple sclerosis using magnetisation prepared 3D-FLAIR and 3D-DIR.* Eur Radiol, 2012. **22**(1): p. 221-31.

125. Geurts, J.J., et al., *Does high-field MR imaging improve cortical lesion detection in multiple sclerosis?* J Neurol, 2008. **255**(2): p. 183-91.

126. Kangarlu, A., et al., *Cerebral cortical lesions in multiple sclerosis detected by MR imaging at 8 Tesla.* AJNR Am J Neuroradiol, 2007. **28**(2): p. 262-6.

127. Mainero, C., et al., *In vivo imaging of cortical pathology in multiple sclerosis using ultra-high field MRI.* Neurology, 2009. **73**(12): p. 941-8.

128. Nielsen, A.S., et al., *Focal cortical lesion detection in multiple sclerosis: 3 Tesla DIR versus 7 Tesla FLASH-T2.* J Magn Reson Imaging, 2012. **35**(3): p. 537-42.

129. Tallantyre, E.C., et al., *3 Tesla and 7 Tesla MRI of multiple sclerosis cortical lesions.* J Magn Reson Imaging, 2010. **32**(4): p. 971-7.

130. Seewann, A., et al., *Postmortem verification of MS cortical lesion detection with 3D DIR.* Neurology, 2012. **78**(5): p. 302-8.

131. Ciccarelli, O. and J.T. Chen, *MS cortical lesions on double inversion recovery MRI: few but true.* Neurology, 2012. **78**(5): p. 296-7.

132. Abduljalil, A.M. and P.M. Robitaille, *Macroscopic susceptibility in ultra high field MRI.* J Comput Assist Tomogr, 1999. **23**(6): p. 832-41.

133. Cremillieux, Y., S. Ding, and J.F. Dunn, *High-resolution in vivo measurements of transverse relaxation times in rats at 7 Tesla.* Magn Reson Med, 1998. **39**(2): p. 285-90.

134. Novak, V., et al., *High-resolution ultrahigh-field MRI of stroke.* Magn Reson Imaging, 2005. **23**(4): p. 539-48.

135. Hammond, K.E., et al., *Development of a robust method for generating 7.0 T multichannel phase images of the brain with application to normal volunteers and patients with neurological diseases.* Neuroimage, 2008. **39**(4): p. 1682-92.

136. Pitt, D., et al., *Imaging cortical lesions in multiple sclerosis with ultra-high-field magnetic resonance imaging.* Arch Neurol, 2010. **67**(7): p. 812-8.

137. Cohen-Adad, J., et al., *In vivo evidence of disseminated subpial T2* signal changes in multiple sclerosis at 7 T: a surface-based analysis.* Neuroimage, 2011. **57**(1): p. 55-62.

138. Cohen-Adad, J., et al., *T(2)* mapping and B(0) orientation-dependence at 7 T reveal cyto- and myeloarchitecture organization of the human cortex.* Neuroimage, 2012. **60**(2): p. 1006-14.

139. Cercignani, M., et al., *Magnetisation transfer ratio and mean diffusivity of normal appearing white and grey matter from patients with multiple sclerosis.* J Neurol Neurosurg Psychiatry, 2001. **70**(3): p. 311-7.

140. Filippi, M., et al., *A magnetization transfer imaging study of normal-appearing white matter in multiple sclerosis.* Neurology, 1995. **45**(3 Pt 1): p. 478-82.

141. Griffin, C.M., et al., *The relationship between lesion and normal appearing brain tissue abnormalities in early relapsing remitting multiple sclerosis.* J Neurol, 2002. **249**(2): p. 193-9.

142. Miller, D.H., A.J. Thompson, and M. Filippi, *Magnetic resonance studies of abnormalities in the normal appearing white matter and grey matter in multiple sclerosis.* J Neurol, 2003. **250**(12): p. 1407-19.

143. Ceccarelli, A., et al., *Normal-appearing white and grey matter damage in MS. A volumetric and diffusion tensor MRI study at 3.0 Tesla.* J Neurol, 2007. **254**(4): p. 513-8.

144. Davies, G.R., et al., *Normal-appearing grey and white matter T1 abnormality in early relapsing-remitting multiple sclerosis: a longitudinal study.* Mult Scler, 2007. **13**(2): p. 169-77.

145. Dehmeshki, J., et al., *The normal appearing grey matter in primary progressive multiple sclerosis: a magnetisation transfer imaging study.* J Neurol, 2003. **250**(1): p. 67-74.

146. Kapeller, P., et al., *Preliminary evidence for neuronal damage in cortical grey matter and normal appearing white matter in short duration relapsing-remitting multiple sclerosis: a quantitative MR spectroscopic imaging study.* J Neurol, 2001. **248**(2): p. 131-8.

147. Ge, Y., et al., *Magnetization transfer ratio histogram analysis of normal-appearing gray matter and normal-appearing white matter in multiple sclerosis.* J Comput Assist Tomogr, 2002. **26**(1): p. 62-8.

148. Coupe, P., P. Yger, and C. Barillot, *Fast non local means denoising for 3D MR images.* Med Image Comput Comput Assist Interv, 2006. **9**(Pt 2): p. 33-40.

149. Fathi, A. and A.R. Naghsh-Nilchi, *Efficient image denoising method based on a new adaptive wavelet packet thresholding function.* IEEE Trans Image Process, 2012. **21**(9): p. 3981-90.

150. Manjon, J.V., et al., *New methods for MRI denoising based on sparseness and self-similarity.* Med Image Anal, 2012. **16**(1): p. 18-27.

151. Sled, J.G., A.P. Zijdenbos, and A.C. Evans, *A nonparametric method for automatic correction of intensity nonuniformity in MRI data.* IEEE Trans Med Imaging, 1998. **17**(1): p. 87-97.

152. Tustison, N.J., et al., *N4ITK: improved N3 bias correction.* IEEE Trans Med Imaging, 2010. **29**(6): p. 1310-20.

153. Galdames, F.J., F. Jaillet, and C.A. Perez, *An accurate skull stripping method based on simplex meshes and histogram analysis for magnetic resonance images.* J Neurosci Methods, 2012. **206**(2): p. 103-19.

154. Sadananthan, S.A., et al., *Skull stripping using graph cuts.* Neuroimage, 2010. **49**(1): p. 225-39.

155. Segonne, F., et al., *A hybrid approach to the skull stripping problem in MRI.* Neuroimage, 2004. **22**(3): p. 1060-75.

156. Smith, S.M., *Fast robust automated brain extraction.* Hum Brain Mapp, 2002. **17**(3): p. 143-55.

157. Eskildsen, S.F., et al., *BEaST: brain extraction based on nonlocal segmentation technique.* Neuroimage, 2012. **59**(3): p. 2362-73.

158. Arnold, J.B., et al., *Qualitative and quantitative evaluation of six algorithms for correcting intensity nonuniformity effects.* Neuroimage, 2001. **13**(5): p. 931-43.

159. Lin, M., et al., *A new bias field correction method combining N3 and FCM for improved segmentation of breast density on MRI.* Med Phys, 2011. **38**(1): p. 5-14.

160. Zheng, W., M.W. Chee, and V. Zagorodnov, *Improvement of brain segmentation accuracy by optimizing non-uniformity correction using N3.* Neuroimage, 2009. **48**(1): p. 73-83.

161. Belaroussi, B., et al., *Intensity non-uniformity correction in MRI: existing methods and their validation.* Med Image Anal, 2006. **10**(2): p. 234-46.

162. Maintz, J.B. and M.A. Viergever, *A survey of medical image registration.* Med Image Anal, 1998. **2**(1): p. 1-36.

163. van den Elsen, P.A., E.J.D. Pol, and M.A. Viergever, *Medical image matching-a review with classification.* Engineering in Medicine and Biology Magazine, IEEE, 1993. **12**(1): p. 26-39.

164. Oliveira, F.P. and J.M. Tavares, *Medical image registration: a review.* Comput Methods Biomech Biomed Engin, 2012.

165. Glocker, B., et al., *Deformable medical image registration: setting the state of the art with discrete methods.* Annu Rev Biomed Eng, 2011. **13**: p. 219-44.

166. Hill, D.L.G. and P. Batchelor, *Registration Methodology*, in *Medical Image Registration*2001, CRC Press. p. 39-70.

167. Hill, D.L.G. and D.J. Hawkes, *Across-Modality Registration Using Intensity-Based Cost Functions*, in *Handbook of medial imaging*, N.B. Isaac, Editor 2000, Academic Press, Inc. p. 537-554.

168. Woods, R.P., *Within-modality registration using intensity-based cost functions*, in *Handbook of medical imaging*, N.B. Isaac, Editor 2000, Academic Press, Inc. p. 529-536.

169. Woods, R.P., S.R. Cherry, and J.C. Mazziotta, *Rapid automated algorithm for aligning and reslicing PET images.* J Comput Assist Tomogr, 1992. **16**(4): p. 620-33.

170. Shannon, C.E., *A mathematical theory of communication.* Bell System Technical Journal, 1948. **27**: p. 379-423.

171. Mortazavi, D., A.Z. Kouzani, and H. Soltanian-Zadeh, *Segmentation of multiple sclerosis lesions in MR images: a review.* Neuroradiology, 2012. **54**(4): p. 299-320.

172. Wicks, D.A., et al., *Volume measurement of multiple sclerosis lesions with magnetic resonance images. A preliminary study.* Neuroradiology, 1992. **34**(6): p. 475-9.

173. Molyneux, P.D., et al., *Quantitative techniques for lesion load measurement in multiple sclerosis: an assessment of the global threshold technique after non uniformity and histogram matching corrections.* Eur J Neurol, 1998. **5**(1): p. 55-60.

174. Wang, L., et al., *Correction for variations in MRI scanner sensitivity in brain studies with histogram matching.* Magn Reson Med, 1998. **39**(2): p. 322-7.

175. Goldberg-Zimring, D., et al., *Automated detection and characterization of multiple sclerosis lesions in brain MR images.* Magn Reson Imaging, 1998. **16**(3): p. 311-8.

176. Hojjatoleslami, S., F. Kruggel, and D. von Cramon, *Segmentation of White Matter Lesions from Volumetric MR ImagesMedical Image Computing and Computer-Assisted Intervention – MICCAI'99*, C. Taylor and A. Colchester, Editors. 1999, Springer Berlin / Heidelberg. p. 52-61.

177. Pachai, C., et al., *A pyramidal approach for automatic segmentation of multiple sclerosis lesions in brain MRI.* Comput Med Imaging Graph, 1998. **22**(5): p. 399-408.

178. Anbeek, P., et al., *Automatic segmentation of different-sized white matter lesions by voxel probability estimation.* Med Image Anal, 2004. **8**(3): p. 205-15.

179. Anbeek, P., et al., *Probabilistic segmentation of white matter lesions in MR imaging.* Neuroimage, 2004. **21**(3): p. 1037-44.

180. Mohamed, F.B., et al., *Increased differentiation of intracranial white matter lesions by multispectral 3D-tissue segmentation: preliminary results.* Magn Reson Imaging, 2001. **19**(2): p. 207-18.

181. Lao, Z., et al., *Computer-assisted segmentation of white matter lesions in 3D MR images using support vector machine.* Acad Radiol, 2008. **15**(3): p. 300-13.

182. Francis, S.J., *Automatic Lesion Identification in MRI of Multiple Sclerosis Patients*, in *Department of Neurology and Neurosurgery*2004, McGill University: Montreal.

183. Khayati, R., et al., *Fully automatic segmentation of multiple sclerosis lesions in brain MR FLAIR images using adaptive mixtures method and Markov random field model.* Comput Biol Med, 2008. **38**(3): p. 379-90.

184. Harmouche, R., et al. *Bayesian MS Lesion Classification Modeling Regional and Local Spatial Infromation.* in *18th International Conference on Pattern Recognition (ICPR 2006)*. 2006. Hong Kong, China: IEEE Computer Society.

185. Garcia-Lorenzo, D., et al., *Multiple sclerosis lesion segmentation using an automatic multimodal graph cuts.* Med Image Comput Comput Assist Interv, 2009. **12**(Pt 2): p. 584-91.

186. Sajja, B.R., et al., *Unified approach for multiple sclerosis lesion segmentation on brain MRI.* Ann Biomed Eng, 2006. **34**(1): p. 142-51.

187. Van Leemput, K., et al., *Automated segmentation of multiple sclerosis lesions by model outlier detection.* IEEE Trans Med Imaging, 2001. **20**(8): p. 677-88.

188. Zijdenbos, A., R. Forghani, and A. Evans, *Automatic quantification of MS lesions in 3D MRI brain data sets: Validation of INSECT*, in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*1998. p. 439-448.

189. Udupa, J.K., et al., *Multiple sclerosis lesion quantification using fuzzy-connectedness principles.* IEEE Trans Med Imaging, 1997. **16**(5): p. 598-609.

190. Shiee, N., et al., *A topology-preserving approach to the segmentation of brain images with multiple sclerosis lesions.* Neuroimage, 2010. **49**(2): p. 1524-35.

191. Calabrese, M., et al., *Detection of cortical inflammatory lesions by double inversion recovery magnetic resonance imaging in patients with multiple sclerosis.* Arch Neurol, 2007. **64**(10): p. 1416-22.

192. Cuadra, M.B., et al., *Atlas-based segmentation of pathological MR brain images using a model of lesion growth.* IEEE Trans Med Imaging, 2004. **23**(10): p. 1301-14.

193. Thirion, J.-P. and G. Calmon, *Deformation Analysis to Detect and Quantify Active Lesions in Three-Dimensional Medical Image Sequences.* IEEE Trans Med Imaging, 1999. **18**(5): p. 429-441.

194. Barkhof, F., et al., *MRI monitoring of immunomodulation in relapse-onset multiple sclerosis trials.* Nat Rev Neurol, 2012. **8**(1): p. 13-21.

195. Petkau, J., et al., *Magnetic resonance imaging as a surrogate outcome for multiple sclerosis relapses.* Mult Scler, 2008. **14**(6): p. 770-8.

196. Korn, E.L., P.S. Albert, and L.M. McShane, *Assessing surrogates as trial endpoints using mixed models.* Stat Med, 2005. **24**(2): p. 163-82.

197. Sormani, M.P., et al., *Magnetic resonance imaging as surrogate for clinical endpoints in multiple sclerosis: data on novel oral drugs.* Mult Scler, 2011. **17**(5): p. 630-3.

198. Sormani, M.P., et al., *Magnetic resonance active lesions as individual-level surrogate for relapses in multiple sclerosis.* Mult Scler, 2011. **17**(5): p. 541-9.

199. Geurts, J.J., et al., *Consensus recommendations for MS cortical lesion scoring using double inversion recovery MRI.* Neurology, 2011. **76**(5): p. 418-24.

200. Jenkinson, M., et al., *Fsl.* Neuroimage, 2012. **62**(2): p. 782-90.

201. Smith, S.M., et al., *Advances in functional and structural MR image analysis and implementation as FSL.* Neuroimage, 2004. **23 Suppl 1**: p. S208-19.

202. Smith, S.M., et al., *Accurate, robust, and automated longitudinal and cross-sectional brain change analysis.* Neuroimage, 2002. **17**(1): p. 479-89.

203. Zhang, Y., M. Brady, and S. Smith, *Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm.* IEEE Trans Med Imaging, 2001. **20**(1): p. 45-57.

204. Friston, K.J., et al., *The relationship between global and local changes in PET scans.* J Cereb Blood Flow Metab, 1990. **10**(4): p. 458-66.

205. Friston, K.J., et al., *Comparing functional (PET) images: the assessment of significant change.* J Cereb Blood Flow Metab, 1991. **11**(4): p. 690-9.

206. Ashburner, J. and K.J. Friston, *Voxel-based morphometry--the methods.* Neuroimage, 2000. **11**(6 Pt 1): p. 805-21.

207. Ashburner, J. and K.J. Friston, *Unified segmentation.* Neuroimage, 2005. **26**(3): p. 839-51.

208. Fischl, B., et al., *Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain.* Neuron, 2002. **33**(3): p. 341-55.

209. Kim, J.S., et al., *Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification.* Neuroimage, 2005. **27**(1): p. 210-21.

210. MacDonald, D., et al., *Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI.* Neuroimage, 2000. **12**(3): p. 340-56.

211. Han, X., et al., *Topology correction in brain cortex segmentation using a multiscale, graph-based algorithm.* IEEE Trans Med Imaging, 2002. **21**(2): p. 109-21.

212. Jones, S.E., B.R. Buchbinder, and I. Aharon, *Three-dimensional mapping of cortical thickness using Laplace's equation.* Hum Brain Mapp, 2000. **11**(1): p. 12-32.

213. Eskildsen, S.F. and L.R. Ostergaard, *Active surface approach for extraction of the human cerebral cortex from MRI.* Med Image Comput Comput Assist Interv, 2006. **9**(Pt 2): p. 823-30.

214. Eskildsen, S.F. and L.R. Ostergaard, *Quantitative comparison of two cortical surface extraction methods using MRI phantoms.* Med Image Comput Comput Assist Interv, 2007. **10**(Pt 1): p. 409-16.

215. Eskildsen, S.F., M. Uldahl, and L.R. Ostergaard, *Extraction of the cerebral cortical boundaries from MRI for measurement of cortical thickness*, in *SPIE*, M.J. Fitzpatrick and J.M. Reinhardt, Editors. 2005.

216. Xu, C., et al., *Reconstruction of the human cerebral cortex from magnetic resonance images.* IEEE Trans Med Imaging, 1999. **18**(6): p. 467-80.

217. Dale, A.M., B. Fischl, and M.I. Sereno, *Cortical surface-based analysis. I. Segmentation and surface reconstruction.* Neuroimage, 1999. **9**(2): p. 179-94.

218. Han, X., et al., *Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer.* Neuroimage, 2006. **32**(1): p. 180-94.

219. Reuter, M., et al., *Within-subject template estimation for unbiased longitudinal image analysis.* Neuroimage, 2012. **61**(4): p. 1402-18.

220. Kuperberg, G.R., et al., *Regionally localized thinning of the cerebral cortex in schizophrenia.* Arch Gen Psychiatry, 2003. **60**(9): p. 878-88.

221. Salat, D.H., et al., *Thinning of the cerebral cortex in aging.* Cereb Cortex, 2004. **14**(7): p. 721-30.

222. Rosas, H.D., et al., *Regional and progressive thinning of the cortical ribbon in Huntington's disease.* Neurology, 2002. **58**(5): p. 695-701.

223. Fischl, B., M.I. Sereno, and A.M. Dale, *Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system.* Neuroimage, 1999. **9**(2): p. 195-207.

224. Fischl, B., et al., *High-resolution intersubject averaging and a coordinate system for the cortical surface.* Hum Brain Mapp, 1999. **8**(4): p. 272-84.

225. Fischl, B., et al., *Automatically parcellating the human cerebral cortex.* Cereb Cortex, 2004. **14**(1): p. 11-22.

226. Desikan, R.S., et al., *An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest.* Neuroimage, 2006. **31**(3): p. 968-80.

227. Fischl, B. *Surface-based analysis: Intersubject registration and smoothing*. [cited 2012 September 6]; Available from: http://surfer.nmr.mgh.harvard.edu/fswiki/FsTutorial.

228. Lerch, J.P. and A.C. Evans, *Cortical thickness analysis examined through power analysis and a population simulation.* Neuroimage, 2005. **24**(1): p. 163-73.

229. Pratt, W.K., *Digital image processing (2nd ed.)*1991: John Wiley \\& Sons, Inc. 698.

230. De Stefano, N., M. Battaglini, and S.M. Smith, *Measuring brain atrophy in multiple sclerosis.* J Neuroimaging, 2007. **17 Suppl 1**: p. 10S-15S.

231. Fox, N.C., et al., *Progressive cerebral atrophy in MS: a serial study using registered, volumetric MRI.* Neurology, 2000. **54**(4): p. 807-12.

232. Smith, S.M., et al., *Normalized accurate measurement of longitudinal brain change.* J Comput Assist Tomogr, 2001. **25**(3): p. 466-75.

233. Jenkinson, M. and S. Smith, *A global optimisation method for robust affine registration of brain images.* Med Image Anal, 2001. **5**(2): p. 143-56.

234. Worsley, K.J. and J.E. Taylor, *Detecting fMRI activation allowing for unknown latency of the hemodynamic response.* Neuroimage, 2006. **29**(2): p. 649-54.

235. Ceccarelli, A., R. Bakshi, and M. Neema, *MRI in multiple sclerosis: a review of the current literature.* Curr Opin Neurol, 2012. **25**(4): p. 402-9.

236. Filippi, M. and M.A. Rocca, *MR imaging of multiple sclerosis.* Radiology, 2011. **259**(3): p. 659-81.

237. Hulst, H.E. and J.J. Geurts, *Gray matter imaging in multiple sclerosis: what have we learned?* BMC Neurol, 2011. **11**: p. 153.

238. Miller, D.H., et al., *Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance.* Brain, 2002. **125**(Pt 8): p. 1676-95.

239. Giorgio, A., et al., *Brain atrophy assessment in multiple sclerosis: importance and limitations.* Neuroimaging Clin N Am, 2008. **18**(4): p. 675-86, xi.

240. Lanz, M., H.K. Hahn, and H. Hildebrandt, *Brain atrophy and cognitive impairment in multiple sclerosis: a review.* J Neurol, 2007. **254 Suppl 2**: p. II43-8.

241. Chard, D.T., et al., *Brain atrophy in clinically early relapsing-remitting multiple sclerosis.* Brain, 2002. **125**(Pt 2): p. 327-37.

242. De Stefano, N., et al., *Evidence of early cortical atrophy in MS: relevance to white matter changes and disability.* Neurology, 2003. **60**(7): p. 1157-62.

243. Prinster, A., et al., *A voxel-based morphometry study of disease severity correlates in relapsing-- remitting multiple sclerosis.* Mult Scler, 2010. **16**(1): p. 45-54.

244. Sanfilipo, M.P., et al., *The relationship between whole brain volume and disability in multiple sclerosis: a comparison of normalized gray vs. white matter with misclassification correction.* Neuroimage, 2005. **26**(4): p. 1068-77.

245. Tedeschi, G., et al., *Brain atrophy and lesion load in a large population of patients with multiple sclerosis.* Neurology, 2005. **65**(2): p. 280-5.

246. Amato, M.P., et al., *Association of neocortical volume changes with cognitive deterioration in relapsing-remitting multiple sclerosis.* Arch Neurol, 2007. **64**(8): p. 1157-61.

247. Benedict, R.H., et al., *Neocortical atrophy, third ventricular width, and cognitive dysfunction in multiple sclerosis.* Arch Neurol, 2006. **63**(9): p. 1301-6.

248. Valsasina, P., et al., *Evidence for progressive gray matter loss in patients with relapsing-remitting MS.* Neurology, 2005. **65**(7): p. 1126-8.

249. Pirko, I., et al., *Gray matter involvement in multiple sclerosis.* Neurology, 2007. **68**(9): p. 634-42.

250. Fisniku, L.K., et al., *Gray matter atrophy is related to long-term disability in multiple sclerosis.* Ann Neurol, 2008. **64**(3): p. 247-54.

251. Calabrese, M., et al., *Widespread cortical thinning characterizes patients with MS with mild cognitive impairment.* Neurology, 2010. **74**(4): p. 321-8.

252. Sailer, M., et al., *Focal thinning of the cerebral cortex in multiple sclerosis.* Brain, 2003. **126**(Pt 8): p. 1734-44.

253. Rudick, R.A., et al., *Gray matter atrophy correlates with MS disability progression measured with MSFC but not EDSS.* J Neurol Sci, 2008.

254. Tekok-Kilic, A., et al., *Independent contributions of cortical gray matter atrophy and ventricle enlargement for predicting neuropsychological impairment in multiple sclerosis.* Neuroimage, 2007. **36**(4): p. 1294-300.

255. Fisher, E., et al., *Gray matter atrophy in multiple sclerosis: a longitudinal study.* Ann Neurol, 2008. **64**(3): p. 255-65.

256. Tedeschi, G., et al., *Brain atrophy evolution and lesion load accrual in multiple sclerosis: a 2-year follow-up study.* Mult Scler, 2009. **15**(2): p. 204-11.

257. Riccitelli, G., et al., *Cognitive impairment in multiple sclerosis is associated to different patterns of gray matter atrophy according to clinical phenotype.* Hum Brain Mapp, 2011. **32**(10): p. 1535-43.

258. Riccitelli, G., et al., *Voxelwise assessment of the regional distribution of damage in the brains of patients with multiple sclerosis and fatigue.* AJNR Am J Neuroradiol, 2011. **32**(5): p. 874-9.

259. Sastre-Garriga, J., et al., *Brain volumetry counterparts of cognitive impairment in patients with multiple sclerosis.* J Neurol Sci, 2009. **282**(1-2): p. 120-4.

260. Battaglini, M., et al., *Voxel-wise assessment of progression of regional brain atrophy in relapsing-remitting multiple sclerosis.* J Neurol Sci, 2009. **282**(1-2): p. 55-60.

261. Mesaros, S., et al., *Evidence of thalamic gray matter loss in pediatric multiple sclerosis.* Neurology, 2008. **70**(13 Pt 2): p. 1107-12.

262. Ceccarelli, A., et al., *A voxel-based morphometry study of grey matter loss in MS patients with different clinical phenotypes.* Neuroimage, 2008. **42**(1): p. 315-22.

263. Charcot, J.-M., *Lectures on the DIseases of the Nervous System delivered at La Salpetriere*, 1877, New Sydenham Society: London. p. 194-195.

264. LaRocca, N.G., *Cognitive Impairment and Mood Disturbances*, in *Primer on Multiple Sclerosis*, B.S. Giesser, Editor 2011, Oxford University Press: New York. p. 241-262.

265. Cottrell, S.S. and S.A. Wilson, *The affective symptomatology of disseminated sclerosis.* J Neurol Psychopathol, 1926. **7**(25): p. 1-30.

266. Kurtzke, J.F., *Neurologic impairment in multiple sclerosis and the disability status scale.* Acta Neurol Scand, 1970. **46**(4): p. 493-512.

267. Canter, A.H., *Direct and Indirect Measures of Psychological Deficit in Multiple Sclerosis: Part I.* The Journal of General Psychology, 1951. **44**(1): p. 3-25.

268. Peyser, J.M., et al., *Guidelines for neuropsychological research in multiple sclerosis.* Arch Neurol, 1990. **47**(1): p. 94-7.

269. Rao, S.M., et al., *Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction.* Neurology, 1991. **41**(5): p. 685-91.

270. McIntosh-Michaelis, S.A., et al., *The prevalence of cognitive impairment in a community survey of multiple sclerosis.* Br J Clin Psychol, 1991. **30 (Pt 4)**: p. 333-48.

271. van den Burg, W., et al., *Cognitive impairment in patients with multiple sclerosis and mild physical disability.* Arch Neurol, 1987. **44**(5): p. 494-501.

272. Feinstein, A., *The Clinical Neuropsychiatry of Multiple Sclerosis*. Second Edition ed2007, New York: Cambridge University Press. 265.

273. Paul, R.H., et al., *Impairments of attention in individuals with multiple sclerosis.* Mult Scler, 1998. **4**(5): p. 433-9.

274. Prakash, R.S., et al., *Cortical recruitment during selective attention in multiple sclerosis: an fMRI investigation of individual differences.* Neuropsychologia, 2008. **46**(12): p. 2888-95.

275. Balsimelli, S., et al., *Attention impairment associated with relapsing-remitting multiple sclerosis patients with mild incapacity.* Arq Neuropsiquiatr, 2007. **65**(2A): p. 262-7.

276. DeLuca, J., et al., *Is speed of processing or working memory the primary information processing deficit in multiple sclerosis?* J Clin Exp Neuropsychol, 2004. **26**(4): p. 550-62.

277. Macniven, J.A., et al., *Stroop performance in multiple sclerosis: information processing, selective attention, or executive functioning?* J Int Neuropsychol Soc, 2008. **14**(5): p. 805-14.

278. Archibald, C.J. and J.D. Fisk, *Information processing efficiency in patients with multiple sclerosis.* J Clin Exp Neuropsychol, 2000. **22**(5): p. 686-701.

279. Minden, S.L., et al., *Memory impairment in multiple sclerosis.* J Clin Exp Neuropsychol, 1990. **12**(4): p. 566-86.

280. DeLuca, J., et al., *Acquisition and storage deficits in multiple sclerosis.* J Clin Exp Neuropsychol, 1998. **20**(3): p. 376-90.

281. Rao, S.M., G.J. Leo, and P. St Aubin-Faubert, *On the nature of memory disturbance in multiple sclerosis.* J Clin Exp Neuropsychol, 1989. **11**(5): p. 699-712.

282. Beatty, W.W., et al., *Implicit learning in patients with chronic progressive multiple sclerosis.* International Journal of Clinical Neuropsychology, 1990. **12**(3-4): p. 166-172.

283. Amato, M.P., V. Zipoli, and E. Portaccio, *Cognitive changes in multiple sclerosis.* Expert Rev Neurother, 2008. **8**(10): p. 1585-96.

284. Kalmar, J.H., et al., *The relationship between cognitive deficits and everyday functional activities in multiple sclerosis.* Neuropsychology, 2008. **22**(4): p. 442-9.

285. Simioni, S., et al., *Progressive decline of decision-making performances during multiple sclerosis.* J Int Neuropsychol Soc, 2009. **15**(2): p. 291-5.

286. Rao, S.M., et al., *Cognitive dysfunction in multiple sclerosis. II. Impact on employment and social functioning.* Neurology, 1991. **41**(5): p. 692-6.

287. Wishart, H. and D. Sharpe, *Neuropsychological aspects of multiple sclerosis: a quantitative review.* J Clin Exp Neuropsychol, 1997. **19**(6): p. 810-24.

288. Carone, D.A., et al., *Interpreting patient/informant discrepancies of reported cognitive symptoms in MS.* J Int Neuropsychol Soc, 2005. **11**(5): p. 574-83.

289. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.* J Psychiatr Res, 1975. **12**(3): p. 189-98.

290. DePaulo, J.R., Jr. and M.F. Folstein, *Psychiatric disturbances in neurological patients: detection, recognition, and hospital course.* Ann Neurol, 1978. **4**(3): p. 225-8.

291. Huber, S.J., et al., *Magnetic resonance imaging correlates of dementia in multiple sclerosis.* Arch Neurol, 1987. **44**(7): p. 732-6.

292. Franklin, G.M., et al., *Correlation of neuropsychological and MRI findings in chronic/progressive multiple sclerosis.* Neurology, 1988. **38**(12): p. 1826-9.

293. Beatty, W.W. and D.E. Goodkin, *Screening for cognitive impairment in multiple sclerosis. An evaluation of the Mini-Mental State Examination.* Arch Neurol, 1990. **47**(3): p. 297-301.

294. Smith, A., *Symbol digit modalities test*, 1982, Western Psychological Services: Los Angeles, USA.

295. Rao, S.M., *A manual for the brief repeatable battery of neuropsychological tests in multiple sclerosis*, 1990, National Multiple Sclerosis Society: New York.

296. Beatty, W.W., et al., *Identifying multiple sclerosis patients with mild or global cognitive impairment using the Screening Examination for Cognitive Impairment (SEFCI).* Neurology, 1995. **45**(4): p. 718-23.

297. Aupperle, R.L., et al., *Three screening batteries to detect cognitive impairment in multiple sclerosis.* Mult Scler, 2002. **8**(5): p. 382-9.

298. Benedict, R.H., et al., *Minimal neuropsychological assessment of MS patients: a consensus approach.* Clin Neuropsychol, 2002. **16**(3): p. 381-97.

299. Strober, L., et al., *Sensitivity of conventional memory tests in multiple sclerosis: comparing the Rao Brief Repeatable Neuropsychological Battery and the Minimal Assessment of Cognitive Function in MS.* Mult Scler, 2009. **15**(9): p. 1077-84.

300. Rao, S.M., et al., *Correlation of magnetic resonance imaging with neuropsychological testing in multiple sclerosis.* Neurology, 1989. **39**(2 Pt 1): p. 161-6.

301. Benedict, R.H. and R. Zivadinov, *Risk factors for and management of cognitive dysfunction in multiple sclerosis.* Nat Rev Neurol, 2011. **7**(6): p. 332-42.

302. Rovaris, M., G. Comi, and M. Filippi, *MRI markers of destructive pathology in multiple sclerosis-related cognitive dysfunction.* J Neurol Sci, 2006. **245**(1-2): p. 111-6.

303. Zivadinov, R., et al., *A longitudinal study of brain atrophy and cognitive disturbances in the early phase of relapsing-remitting multiple sclerosis.* J Neurol Neurosurg Psychiatry, 2001. **70**(6): p. 773-80.

304. Simon, J.H., et al., *A longitudinal study of brain atrophy in relapsing multiple sclerosis. The Multiple Sclerosis Collaborative Research Group (MSCRG).* Neurology, 1999. **53**(1): p. 139-48.

305. Kalkers, N.F., et al., *Optimizing the association between disability and biological markers in MS.* Neurology, 2001. **57**(7): p. 1253-8.

306. Batista, S., et al., *Basal ganglia, thalamus and neocortical atrophy predicting slowed cognitive processing in multiple sclerosis.* J Neurol, 2012. **259**(1): p. 139-46.

307. Jasperse, B., et al., *Regional brain atrophy development is related to specific aspects of clinical dysfunction in multiple sclerosis.* Neuroimage, 2007. **38**(3): p. 529-37.

308. Houtchens, M.K., et al., *Thalamic atrophy and cognition in multiple sclerosis.* Neurology, 2007. **69**(12): p. 1213-23.

309. Rao, S.M., et al., *Cerebral disconnection in multiple sclerosis. Relationship to atrophy of the corpus callosum.* Arch Neurol, 1989. **46**(8): p. 918-20.

310. Pelletier, J., et al., *A longitudinal study of callosal atrophy and interhemispheric dysfunction in relapsing-remitting multiple sclerosis.* Arch Neurol, 2001. **58**(1): p. 105-11.

311. Amato, M.P., et al., *Neocortical volume decrease in relapsing-remitting MS patients with mild cognitive impairment.* Neurology, 2004. **63**(1): p. 89-93.

312. Sicotte, N.L., et al., *Regional hippocampal atrophy in multiple sclerosis.* Brain, 2008. **131**(Pt 4): p. 1134-41.

313. Benedict, R.H., et al., *Memory impairment in multiple sclerosis: correlation with deep grey matter and mesial temporal atrophy.* J Neurol Neurosurg Psychiatry, 2009. **80**(2): p. 201-6.

314. Achiron, A., et al., *Superior temporal gyrus thickness correlates with cognitive performance in multiple sclerosis.* Brain Struct Funct, 2012.

315. Comi, G., et al., *Clinical, neurophysiological, and magnetic resonance imaging correlations in multiple sclerosis.* J Neurol Neurosurg Psychiatry, 1998. **64 Suppl 1**: p. S21-5.

316. van Buchem, M.A., et al., *Correlation of volumetric magnetization transfer imaging with clinical data in MS.* Neurology, 1998. **50**(6): p. 1609-17.

317. Amato, M.P., et al., *Cognitive assessment and quantitative magnetic resonance metrics can help to identify benign multiple sclerosis.* Neurology, 2008. **71**(9): p. 632-8.

318. Rovaris, M., et al., *Cortical/subcortical disease burden and cognitive impairment in patients with multiple sclerosis.* AJNR Am J Neuroradiol, 2000. **21**(2): p. 402-8.

319. Filippi, M., et al., *Changes in the normal appearing brain tissue and cognitive impairment in multiple sclerosis.* J Neurol Neurosurg Psychiatry, 2000. **68**(2): p. 157-61.

320. Ranjeva, J.P., et al., *Local tissue damage assessed with statistical mapping analysis of brain magnetization transfer ratio: relationship with functional status of patients in the earliest stage of multiple sclerosis.* AJNR Am J Neuroradiol, 2005. **26**(1): p. 119-27.

321. Khaleeli, Z., et al., *Localized grey matter damage in early primary progressive multiple sclerosis contributes to disability.* Neuroimage, 2007. **37**(1): p. 253-61.

322. Audoin, B., et al., *Magnetic resonance study of the influence of tissue damage and cortical reorganization on PASAT performance at the earliest stage of multiple sclerosis.* Hum Brain Mapp, 2005. **24**(3): p. 216-28.

323. Christodoulou, C., et al., *Cognitive performance and MR markers of cerebral injury in cognitively impaired MS patients.* Neurology, 2003. **60**(11): p. 1793-8.

324. Gadea, M., et al., *Spectroscopic axonal damage of the right locus coeruleus relates to selective attention impairment in early stage relapsing-remitting multiple sclerosis.* Brain, 2004. **127**(Pt 1): p. 89-98.

325. Staffen, W., et al., *Magnetic resonance spectroscopy of memory and frontal brain region in early multiple sclerosis.* J Neuropsychiatry Clin Neurosci, 2005. **17**(3): p. 357-63.

326. Pan, J.W., et al., *Cognitive dysfunction lateralizes with NAA in multiple sclerosis.* Appl Neuropsychol, 2001. **8**(3): p. 155-60.

327. Roosendaal, S.D., et al., *Accumulation of cortical lesions in MS: relation with cognitive impairment.* Mult Scler, 2009. **15**(6): p. 708-14.

328. Calabrese, M., et al., *Cortical lesions and atrophy associated with cognitive impairment in relapsing-remitting multiple sclerosis.* Arch Neurol, 2009. **66**(9): p. 1144-50.

329. Derakhshan, M., et al., *An Evaluation of Automatic Techniques for the Quantification of Grey Matter Atrophy in Multiple Sclerosis*, in *Medical Image Computing and Computer-Assisted Interventation - MICCAI'09*2009, Springer: London.

330. Stadelmann, C., et al., *Cortical pathology in multiple sclerosis.* Curr Opin Neurol, 2008. **21**(3): p. 229-34.

331. Bo, L., et al., *Lack of correlation between cortical demyelination and white matter pathologic changes in multiple sclerosis.* Arch Neurol, 2007. **64**(1): p. 76-80.

332. Caramanos, Z., et al., *(1)H-MRSI evidence for cortical gray matter pathology that is independent of cerebral white matter lesion load in patients with secondary progressive multiple sclerosis.* J Neurol Sci, 2009. **282**(1-2): p. 72-9.

333. Patenaude, B., *Bayesian Statistical Models of Shape and Appearance for Subcortical Brain Segmentation*, 2007, University of Oxford.

334. Ghassemi, R., et al., *Lesion distribution in children with clinically isolated syndromes.* Ann Neurol, 2008. **63**(3): p. 401-5.

335. Levesque, I., et al., *The role of edema and demyelination in chronic T1 black holes: a quantitative magnetization transfer study.* J Magn Reson Imaging, 2005. **21**(2): p. 103-10.

336. Sled, J.G., et al., *Regional variations in normal brain shown by quantitative magnetization transfer imaging.* Magn Reson Med, 2004. **51**(2): p. 299-303.

337. Tartaglia, M.C., et al., *Choline is increased in pre-lesional normal appearing white matter in multiple sclerosis.* J Neurol, 2002. **249**(10): p. 1382-90.

338. Nyul, L.G., J.K. Udupa, and X. Zhang, *New variants of a method of MRI scale standardization.* IEEE Trans Med Imaging, 2000. **19**(2): p. 143-50.

339. Warfield, S.K., K.H. Zou, and W.M. Wells, *Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation.* IEEE Trans Med Imaging, 2004. **23**(7): p. 903-21.

340. Zijdenbos, A.P., et al., *Morphometric analysis of white matter lesions in MR images: method and validation.* IEEE Trans Med Imaging, 1994. **13**(4): p. 716-24.

341. Lee, H. and I. Prohovnik, *Cross-validation of brain segmentation by SPM5 and SIENAX.* Psychiatry Res, 2008. **164**(2): p. 172-7.

342. Audoin, B., et al., *Localization of grey matter atrophy in early RRMS : A longitudinal study.* J Neurol, 2006. **253**(11): p. 1495-501.

343. Audoin, B., et al., *Voxel-based analysis of MTR images: a method to locate gray matter abnormalities in patients at the earliest stage of multiple sclerosis.* J Magn Reson Imaging, 2004. **20**(5): p. 765-71.

344. Sastre-Garriga, J., et al., *Brain volumetry counterparts of cognitive impairment in patients with multiple sclerosis.* J Neurol Sci, 2009.

345. Antulov, R., et al., *Gender-related differences in MS: a study of conventional and nonconventional MRI measures.* Mult Scler, 2009. **15**(3): p. 345-54.

346. Khaleeli, Z., et al., *Magnetization transfer ratio in gray matter: a potential surrogate marker for progression in early primary progressive multiple sclerosis.* Arch Neurol, 2008. **65**(11): p. 1454-9.

347. Portaccio, E., et al., *Neocortical volume decrease in relapsing-remitting multiple sclerosis with mild cognitive impairment.* J Neurol Sci, 2006. **245**(1-2): p. 195-9.

348. Prinster, A., et al., *Grey matter loss in relapsing-remitting multiple sclerosis: a voxel-based morphometry study.* Neuroimage, 2006. **29**(3): p. 859-67.

349. Chard, D.T., et al., *The reproducibility and sensitivity of brain tissue volume measurements derived from an SPM-based segmentation methodology.* J Magn Reson Imaging, 2002. **15**(3): p. 259-67.

350. Chard, D., et al., *[P692] Reducing the effects of white matter lesions on measures of brain grey and white matter volumes*, in *European Commitee for Treatment and Research in Multiple Sclerosis (ECTRIMS)*2009: Dusseldorf, Germany. p. P692.

351. Battaglini, M., et al., *[P705] Influence of white matter lesions in the measurement of MRI-derived brain columes. Implications for clinical multiple sclerosis studies*, in *European Commitee for Treatment and Research in Multiple Sclerosis (ECTRIMS)*2009: Dusseldorf, Germany. p. P705.

352. Anderson, V.M., N.C. Fox, and D.H. Miller, *Magnetic resonance imaging measures of brain atrophy in multiple sclerosis.* J Magn Reson Imaging, 2006. **23**(5): p. 605-18.

353. Bermel, R.A. and R. Bakshi, *The measurement and clinical relevance of brain atrophy in multiple sclerosis.* Lancet Neurol, 2006. **5**(2): p. 158-70.

354. Derakhshan, M., et al. *Regions of reduced cortical magnetization transfer ratio detected in MS patients using surface-based techniques*. in *International Society for Magnetic Resonance in Medicine - ISMRM*. 2009. Honolulu, USA.

355. Derakhshan, M., et al. *Surface-based techniques reveal regions of reduced cortical magnetization transfer ratio in patients with MS*. in *International Society for Magnetic Resonance in Medicine - ISMRM*. 2010. Stockholm, Sweeden.

356. Kollia, K., et al., *First clinical study on ultra-high-field MR imaging in patients with multiple sclerosis: comparison of 1.5T and 7T.* AJNR Am J Neuroradiol, 2009. **30**(4): p. 699-702.

357. Fisniku, L.K., et al., *Magnetization transfer ratio abnormalities reflect clinically relevant grey matter damage in multiple sclerosis.* Mult Scler, 2009. **15**(6): p. 668-77.

358. Jure, L., et al., *Individual voxel-based analysis of brain magnetization transfer maps shows great variability of gray matter injury in the first stage of multiple sclerosis.* J Magn Reson Imaging, 2010. **32**(2): p. 424-8.

359. Derakhshan, M., et al., *Evaluation of automated techniques for the quantification of grey matter atrophy in patients with multiple sclerosis.* Neuroimage, 2010. **52**(4): p. 1261-7.

360. Benedetti, B., et al., *Influence of aging on brain gray and white matter changes assessed by conventional, MT, and DT MRI.* Neurology, 2006. **66**(4): p. 535-9.

361. Ge, Y., et al., *Age-related total gray matter and white matter changes in normal adult brain. Part II: quantitative magnetization transfer ratio histogram analysis.* AJNR Am J Neuroradiol, 2002. **23**(8): p. 1334-41.

362. Lee, K.Y., et al., *Age-related changes in conventional and magnetization transfer MR imaging in elderly people: comparison with neurocognitive performance.* Korean J Radiol, 2004. **5**(2): p. 96-101.

363. Collins, D.L., et al., *Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space.* J Comput Assist Tomogr, 1994. **18**(2): p. 192-205.

364. Ihaka, R. and R. Gentleman, *R: A Language for Data Analysis and Graphics.* Journal of Computational and Graphical Statistics, 1996. **5**(3): p. 299-314.

365. Glasser, M.F. and D.C. Van Essen, *Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI.* J Neurosci, 2011. **31**(32): p. 11597-616.

366. Lassmann, H. and C.F. Lucchinetti, *Cortical demyelination in CNS inflammatory demyelinating diseases.* Neurology, 2008. **70**(5): p. 332-3.

367. Wonderlick, J.S., et al., *Reliability of MRI-derived cortical and subcortical morphometric measures: effects of pulse sequence, voxel geometry, and parallel imaging.* Neuroimage, 2009. **44**(4): p. 1324-33.

368. Kutzelnigg, A. and H. Lassmann, *Cortical demyelination in multiple sclerosis: a substrate for cognitive deficits?* J Neurol Sci, 2006. **245**(1-2): p. 123-6.

369. Derakhshan, M., et al. *Global and Regional Cortical Thickness Correlates of Cognitive Function in Cognitively Impaired Patients with Multiple Sclerosis*. in *American Academy of Neurology*. 2011. Honolulu, USA.

370. Benedict, R.H., et al., *Validity of the minimal assessment of cognitive function in multiple sclerosis (MACFIMS).* J Int Neuropsychol Soc, 2006. **12**(4): p. 549-58.

371. Bobholz, J.A. and S.M. Rao, *Cognitive dysfunction in multiple sclerosis: a review of recent developments.* Curr Opin Neurol, 2003. **16**(3): p. 283-8.

372. Sepulcre, J., et al., *Cognitive impairment in patients with multiple sclerosis using the Brief Repeatable Battery-Neuropsychology test.* Mult Scler, 2006. **12**(2): p. 187-95.

373. Bodling, A.M., D.R. Denney, and S.G. Lynch, *Individual variability in speed of information processing: an index of cognitive impairment in multiple sclerosis.* Neuropsychology, 2012. **26**(3): p. 357-67.

374. Covey, T.J., et al., *Information processing speed, neural efficiency, and working memory performance in multiple sclerosis: differential relationships with structural magnetic resonance imaging.* J Clin Exp Neuropsychol, 2011. **33**(10): p. 1129-45.

375. Rao, S.M., P. St Aubin-Faubert, and G.J. Leo, *Information processing speed in patients with multiple sclerosis.* J Clin Exp Neuropsychol, 1989. **11**(4): p. 471-7.

376. Forn, C., et al., *Information-processing speed is the primary deficit underlying the poor performance of multiple sclerosis patients in the Paced Auditory Serial Addition Test (PASAT).* J Clin Exp Neuropsychol, 2008. **30**(7): p. 789-96.

377. Wechsler, D., *The measurement of adult intelligence*. 3rd ed1944, Baltimore, Md.: Williams & Wilkins. vii, 258 p.

378. Gronwall, D.M., *Paced auditory serial-addition task: a measure of recovery from concussion.* Percept Mot Skills, 1977. **44**(2): p. 367-73.

379. Glanz, B.I., et al., *Cognitive dysfunction in patients with clinically isolated syndromes or newly diagnosed multiple sclerosis.* Mult Scler, 2007. **13**(8): p. 1004-10.

380. Rosti, E., et al., *PASAT in detecting cognitive impairment in relapsing-remitting MS.* Appl Neuropsychol, 2007. **14**(2): p. 101-12.

381. Feuillet, L., et al., *Early cognitive impairment in patients with clinically isolated syndrome suggestive of multiple sclerosis.* Mult Scler, 2007. **13**(1): p. 124-7.

382. Shallice, T., *Specific impairments of planning.* Philos Trans R Soc Lond B Biol Sci, 1982. **298**(1089): p. 199-209.

383. Culbertson, W.C. and E.A. Zillmer, *Tower of London - Drexel University (TOL DX)*, 2001, Multi-Health Systems: Chicago, IL.

384. Sanchez-Benavides, G., et al., *Problem-solving abilities and frontal lobe cortical thickness in healthy aging and mild cognitive impairment.* J Int Neuropsychol Soc, 2010. **16**(5): p. 836-45.

385. Berg, W.K., et al., *Deconstructing the tower: parameters and predictors of problem difficulty on the Tower of London task.* Brain Cogn, 2010. **72**(3): p. 472-82.

386. Dagher, A., et al., *Mapping the network for planning: a correlational PET activation study with the Tower of London task.* Brain, 1999. **122 (Pt 10)**: p. 1973-87.

387. Greenwood, K.E., et al., *Tower of London versus Real Life Analogue Planning in Schizophrenia with Disorganization and Psychomotor Poverty Symptoms.* J Int Neuropsychol Soc, 2011: p. 1-11.

388. MacAllister, W.S., et al., *Assessment of executive functioning in childhood epilepsy: the Tower of London and BRIEF.* Child Neuropsychol, 2012. **18**(4): p. 404-15.

389. Rasser, P.E., et al., *Functional MRI BOLD response to Tower of London performance of first-episode schizophrenia patients using cortical pattern matching.* Neuroimage, 2005. **26**(3): p. 941-51.

390. Zook, N.A., et al., *Working memory, inhibition, and fluid intelligence as predictors of performance on Tower of Hanoi and London tasks.* Brain Cogn, 2004. **56**(3): p. 286-92.

391. Chiaravalloti, N.D. and J. DeLuca, *Cognitive impairment in multiple sclerosis.* Lancet Neurol, 2008. **7**(12): p. 1139-51.

392. Wechsler, D., *Wechsler memory scale*1945, San Antonio, TX: The Psychological Corporation.

393. Woodcock, R., K.S. McGrew, and N. Mather, *Woodcock-Johnson III normative Update (NU) Tests of Cognitive Abilities.*, 2001, Riverside Publishing: Rolling Measows, IL.

394. Heaton, R.K., *Wisconsin card sorting test manual*1993, Odessa, Florida: Psychological Asseessment Resources.

395. Kaplan, E., H. Goodglass, and S. Weintraub, *Boston naming test*1983, Philadelphia: Lea & Febiger.

396. Delis, D.C., E. Kaplan, and J.H. Kramer, *Delis-Kaplan executive function system (D-KEFS)*, 2001, The Psychological Corporation: San Antonio, TX.

397. Leonard, G., et al. *Sequencing of complex bimanual movements: Sex differences as related to complexity*. in *Society for Neuroscience*. 2010. San Diego, USA.

398. Thurstone, L.L., *A factorial study of perception*1944, Chicago, IL, US: University of Chicago Press. vi, 148.

399. Leonard, G., B. Milner, and L. Jones, *Performance on unimanual and bimanual tapping tasks by patients with lesions of the frontal or temporal lobe.* Neuropsychologia, 1988. **26**(1): p. 79-91.

400. Derakhshan, M., et al., *Surface-based analysis reveals regions of cortical magnetization transfer ratio in patients with multiple sclerosis: a proposed method for imaging subpial demyelination.* Human Brain Mapping, 2012.

401. Collins, D.L. and J.C. Pruessner, *Towards accurate, automatic segmentation of the hippocampus and amygdala from MRI by augmenting ANIMAL with a template library and label fusion.* Neuroimage, 2010. **52**(4): p. 1355-66.

402. Akaike, H., *A new look at the statistical model identification.* Automatic Control, IEEE Transactions on, 1974. **19**(6): p. 716-723.

403. Burnham, K.P. and D.R. Anderson, *Model selection and multimodel inference : a practical information-theoretic approach*. 2nd ed2002, New York: Springer. xxvi, 488 p.

404. Mazerolle, M.J., *Mouvements et reprodcution des amphibiens en tourbieres perturbees. Appendix 1: Making sense out of Akaike's Information Criterion (AIC)*, in *Faculte de foresterie et de geomatique*2004, University Laval: Laval, QC.

405. Wagenmakers, E.J. and S. Farrell, *AIC model selection using Akaike weights.* Psychon Bull Rev, 2004. **11**(1): p. 192-6.

406. Bever, C., et al., *The brief repeatable battery of neuropsychological tests for multiple sclerosis: a preliminary serial study.* Multiple Sclerosis, 1995. **1**(3): p. 165-169.

407. Forn, C., et al., *Anatomical and functional differences between the Paced Auditory Serial Addition Test and the Symbol Digit Modalities Test.* J Clin Exp Neuropsychol, 2011. **33**(1): p. 42-50.

408. Bermel, R.A., et al., *Bicaudate ratio as a magnetic resonance imaging marker of brain atrophy in multiple sclerosis.* Arch Neurol, 2002. **59**(2): p. 275-80.

409. Cohen, J., *Statistical power analysis for the behavioral sciences*. 2nd ed1988, Hillsdale, N.J.: L. Erlbaum Associates. xxi, 567 p.

410. Nakamura, K. and E. Fisher, *Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients.* Neuroimage, 2009. **44**(3): p. 769-76.

411. Derakhshan, M., et al. *Combining SIENA and SIENAx for improved quantification of grey and white matter atrophy*. in *International Society for Magnetic Resonance in Medicine (ISMRM)*. 2012. Melbourne, Australia.

412. Derakhshan, M., et al. *Hippocampal magentization transfer ratio and not hippocampal atrophy best explains memory dysfunction in patients with multiple sclerosis*. in *International Society for Magnetic Resonance in Medicine (ISMRM)*. 2013. Salt Lake City, USA.