DEVELOPMENT AND APPLICATION OF NOVEL ALGORITHMS FOR QUANTITATIVE ANALYSIS OF MAGNETIC RESONANCE IMAGING IN MULTIPLE SCLEROSIS

Michael G. DWYER III

PhD by Published Work

School of Engineering, Design, and Technology

UNIVERSITY OF BRADFORD

2013

Abstract

DEVELOPMENT AND APPLICATION OF NOVEL ALGORITHMS FOR QUANTITATIVE ANALYSIS OF MAGNETIC RESONANCE IMAGING IN MULTIPLE SCLEROSIS

Michael G. DWYER III

Key Words:

This document is a critical synopsis of prior work by Michael Dwyer, submitted in support of a PhD by published work. The selected work is focused on the application of quantitative magnetic resonance imaging (MRI) analysis techniques to the study of multiple sclerosis (MS).

MS is a debilitating disease with a multi-factorial pathology, progression, and clinical presentation. Its most salient feature is focal inflammatory lesions, but it also includes significant parenchymal atrophy and microstructural damage. As a powerful tool for *in vivo* investigation of tissue properties, MRI can provide important clinical and scientific information regarding these various aspects of the disease, but precise, accurate quantitative analysis techniques are needed to detect subtle changes and to cope with the vast amount of data produced in an MRI session.

To address this, eight new techniques were developed by Michael Dwyer and his co-workers to better elucidate focal, atrophic, and occult/"invisible" pathology. These included: a method to better evaluate errors in lesion identification; a method to quantify differences in lesion distribution between scanner strengths; a method to measure optic nerve atrophy; a more precise method to quantify tissue-specific atrophy; a method sensitive to dynamic myelin changes; and a method to quantify iron in specific brain structures.

Taken together, these new techniques are complementary and improve the ability of clinicians and researchers to reliably assess various key elements of MS pathology *in vivo*.

Acknowledgements

No scientific work occurs in a vacuum, and so it is often difficult to decide who deserves mention and in what order. This is not the case for me. I owe most of what I am academically and professionally today to Professor Robert Zivadinov. He has helped and guided me in innumerable ways beyond the possibility to list here, but the most important was his faith in me. It has been in striving to match his expectations and trust in me that I have become the scientist and person that I am today. I am honoured and proud to call him a mentor, colleague, and most of all a friend.

There are many others who have provided support, guidance, and encouragement. In particular, I would like to thank Professor Clive Beggs, who has served as my mentor through this PhD program. His generosity of time and advice and talents for organization and clear thought have turned what could have been a stressful process into an enjoyable one.

I would also like to thank my wife Sarah and my son Henry for their endless patience through years of long workdays and stolen weekends dedicated to scientific endeavours rather than the family picnics they deserve.

Finally, I would like to thank my parents Karen and Michael for their constant love and support, and for sacrificing their early savings to buy a young child a career-inspiring computer before such things were popular.

Contents

Chapter 1 – Introduction

1.1 Purpose

This document is a synopsis of the published work of Michael Dwyer (hereafter referred to as the author), which is herewith submitted for a PhD by published work.

1.2 Selected publications

The body of published work submitted for this PhD is itemized in Table 1, which comprises eleven journal papers and one conference proceeding. Copies of these publications are included in Appendix B. They have been selected because they represent both a significant contribution to the scientific knowledge base and either primary or substantial involvement from the author. Additional supporting material, which comprises the eleven journal papers itemized in Table 2, is included in Appendix C.

Table 1. Publications submitted for inclusion in the PhD

Ref. No.	Publication	Status[*]
1	**Semi-automatic brain region extraction (SABRE) reveals superior cortical and deep gray matter atrophy in MS.** Carone DA, Benedict RH, Dwyer MG, Cookfair DL, Srinivasaraghavan B, Tjoa CW, Zivadinov R. Neuroimage. 2006 Jan 15;29(2):505-14. Epub 2005 Oct 5**.**	RP***

Ref. No.	Publication	Status[*]
2	**Effect of MRI coregistration on serial short-term brain volume changes in multiple sclerosis.** Fritz DA, Dwyer MG, Bagnato F, Watts KL, Bratina A, Zorzon M, Durastanti V, Locatelli L, Millefiorini E, Zivadinov R. Neurol Res. 2006 Apr;28(3):275-9.	RP***
3	**Neocortical atrophy, third ventricular width, and cognitive dysfunction in multiple sclerosis.** Benedict RH, Bruce JM, Dwyer MG, Abdelrahman N, Hussein S, Weinstock-Guttman B, Garg N, Munschauer F, Zivadinov R. *Arch Neurol. 2006 Sep;63(9):1301-6.*	RP***
4	**Quantitative diffusion weighted imaging measures in patients with multiple sclerosis.** Tavazzi E, Dwyer MG, Weinstock-Guttman B, Lema J, Bastianello S, Bergamaschi R, Cosi V, Benedict RH, Munschauer FE 3rd, Zivadinov R. *Neuroimage. 2007 Jul 1;36(3):746-54. Epub 2007 Apr 10.*	RP***
5	**Application of hidden Markov random field approach for quantification of perfusion/diffusion mismatch in acute ischemic stroke.** Dwyer MG, Bergsland N, Saluste E, Sharma J, Jaisani Z, Durfee J, Abdelrahman N, Minagar A, Hoque R, Munschauer FE 3rd, Zivadinov R. *Neurol Res. 2008 Oct;30(8):827-34. doi: 10.1179/174313208X340987.*	RP**
6	**A sensitive, noise-resistant method for identifying focal demyelination and remyelination in patients with multiple sclerosis via voxel-wise changes in magnetization transfer ratio.** Dwyer MG, Bergsland N, Hussein S, Durfee JE, Wack DS, Zivadinov R. *J Neurol Sci. 2009 Jul 15;282(1-2):86-95. doi: 10.1016/j.jns.2009.03.016. Epub 2009 Apr 22.*	RP**
7	**Relationship of optic nerve and brain conventional and non-conventional MRI measures and retinal nerve fiber layer thickness, as assessed by OCT and GDx: a pilot study.** Frohman EM, Dwyer MG, Frohman T, Cox JL, Salter A, Greenberg BM, Hussein S, Conger A, Calabresi P, Balcer LJ, Zivadinov R. *J Neurol Sci. 2009 Jul 15;282(1-2):96-105. doi: 10.1016/j.jns.2009.04.010. Epub 2009 May 12.*	RP***
8	**Signal abnormalities on 1.5 and 3 Tesla brain MRI in multiple sclerosis patients and healthy controls. A morphological and spatial quantitative comparison study.** Di Perri C, Dwyer MG, Wack DS, Cox JL, Hashmi K, Saluste E, Hussein S, Schirda C, Stosic M, Durfee J, Poloni GU, Nayyar N, Bergamaschi R, Zivadinov R. *Neuroimage. 2009 Oct 1;47(4):1352-62. doi: 10.1016/j.neuroimage.2009.04.019. Epub 2009 Apr 14.*	RP***
9	**Abnormal subcortical deep-gray matter susceptibility-weighted imaging filtered phase measurements in patients with multiple sclerosis: a case-control study.** Zivadinov R, Heininen-Brown M, Schirda CV, Poloni GU, Bergsland N, Magnano CR, Durfee J, Kennedy C, Carl E, Hagemeier J, Benedict RH, Weinstock-Guttman B, Dwyer MG. *Neuroimage. 2012 Jan 2;59(1):331-9. doi: 10.1016/j.neuroimage.2011.07.045. Epub 2011 Jul 27.*	RP***
10	**Improved assessment of multiple sclerosis lesion segmentation agreement via detection and outline error estimates.** Wack DS, Dwyer MG, Bergsland N, Di Perri C, Ranza L, Hussein S, Ramasamy D, Poloni G, Zivadinov R. *BMC Med Imaging. 2012 Jul 19;12:17. doi: 10.1186/1471-2342-12-17.*	RP***
11	**Improved longitudinal gray matter atrophy assessment via a combination of SIENA and a 4-dimensional hidden Markov random field model.** Dwyer MG, Bergsland N, Zivadinov R. *28th European Committee for Treatment and Research in Multiple Sclerosis, Lyon, France, October 10-13, 2012:P840.*	P**

Table 2. Supporting documents

Publication	Status[*]
Detection of cortical lesions is dependent on choice of slice thickness in patients with multiple sclerosis. Dolezal O, Dwyer MG, Horakova D, Havrdova E, Minagar A, Balachandran S, Bergsland N, Seidl Z, Vaneckova M, Fritz D, Krasensky J, Zivadinov R. *Int Rev Neurobiol. 2007;79:475-89.*	RP***
Diffusion-weighted imaging predicts cognitive impairment in multiple sclerosis. Benedict RH, Bruce J, Dwyer MG, Weinstock-Guttman B, Tjoa C, Tavazzi E, Munschauer FE, Zivadinov R. *Mult Scler. 2007 Jul;13(6):722-30. Epub 2007 Mar 15.*	RP***
Comparison of three different methods for measurement of cervical cord atrophy in multiple sclerosis. Zivadinov R, Banas AC, Yella V, Abdelrahman N, Weinstock-Guttman B, Dwyer MG. *AJNR Am J Neuroradiol. 2008 Feb;29(2):319-25. Epub 2007 Nov 1.*	RP***
The place of conventional MRI and newly emerging MRI techniques in monitoring different aspects of treatment outcome. Zivadinov R, Stosic M, Cox JL, Ramasamy DP, Dwyer MG. *J Neurol. 2008 Mar;255 Suppl 1:61-74. doi: 10.1007/s00415-008-1009-1. Review.*	RP***
Use of perfusion- and diffusion-weighted imaging in differential diagnosis of acute and chronic ischemic stroke and multiple sclerosis. Zivadinov R, Bergsland N, Stosic M, Sharma J, Nussenbaum F, Durfee J, Hani N, Abdelrahman N, Jaisani Z, Minagar A, Hoque R, Munschauer FE 3rd, Dwyer MG. *Neurol Res. 2008 Oct;30(8):816-26. doi: 10.1179/174313208X341003.*	RP***
Gray matter atrophy and disability progression in patients with early relapsing-remitting multiple sclerosis: a 5-year longitudinal study. Horakova D, Dwyer MG, Havrdova E, Cox JL, Dolezal O, Bergsland N, Rimes B, Seidl Z, Vaneckova M, Zivadinov R. *J Neurol Sci. 2009 Jul 15;282(1-2):112-9. doi: 10.1016/j.jns.2008.12.005. Epub 2009 Jan 24.*	RP***
Voxel-wise magnetization transfer imaging study of effects of natalizumab and IFNβ-1a in multiple sclerosis. Zivadinov R, Dwyer MG, Hussein S, Carl E, Kennedy C, Andrews M, Hojnacki D, Heininen-Brown M, Willis L, Cherneva M, Bergsland N, Weinstock-Guttman B. *Mult Scler. 2012 Aug;18(8):1125-34. doi: 10.1177/1352458511433304. Epub 2011 Dec 22.*	RP***
Iron deposition on SWI-filtered phase in the subcortical deep gray matter of patients with clinically isolated syndrome may precede structure-specific atrophy. Hagemeier J, Weinstock-Guttman B, Bergsland N, Heininen-Brown M, Carl E, Kennedy C, Magnano C, Hojnacki D, Dwyer MG, Zivadinov R. *AJNR Am J Neuroradiol. 2012 Sep;33(8):1596-601. doi: 10.3174/ajnr.A3030. Epub 2012 Mar 29.*	RP***
Subcortical and cortical gray matter atrophy in a large sample of patients with clinically isolated syndrome and early relapsing-remitting multiple sclerosis. Bergsland N, Horakova D, Dwyer MG, Dolezal O, Seidl ZK, Vaneckova M, Krasensky J, Havrdova E, Zivadinov R. *AJNR Am J Neuroradiol. 2012 Sep;33(8):1573-8. doi: 10.3174/ajnr.A3086. Epub 2012 Apr 12.*	RP***
Bimonthly Evolution of Cortical Atrophy in Early Relapsing-Remitting Multiple Sclerosis over 2 Years: A Longitudinal Study. Zivadinov R, Tekwe C, Bergsland N, Dolezal O, Havrdova E, Krasensky J, Dwyer MG, Seidl Z, Ramasamy DP, Vaneckova M, Horakova D. *Mult Scler Int. 2013;2013:231345. doi: 10.1155/2013/231345. Epub 2013 Jan 10.*	RP***

Publication	Status[*]
Gray matter SWI-filtered phase and atrophy are linked to disability in MS. Hagemeier J, Weinstock-Guttman B, Heininen-Brown M, Poloni GU, Bergsland N, Schirda C, Magnano CR, Kennedy C, Carl E, Dwyer MG, Minagar A, Zivadinov R. *Front Biosci (Elite Ed). 2013 Jan 1;5:525-32.*	RP***
Improved longitudinal gray and white matter atrophy assessment via application of a 4-dimensional hidden Markov random field model. Dwyer MG, Bergsland NP, Zivadinov R. [Unpublished manuscript]	N-P**

[*]Status: P = available in the public domain; R = refereed journal paper; N-P = not yet available; * = sole author; ** = principal author; *** = joint author

1.3 Statement of work

This document summarizes, synthesizes, and critically evaluates published work of the author. This work involves the development and application of novel computational algorithms for quantitatively analysing magnetic resonance imaging (MRI) to better understand and characterize the pathology of multiple sclerosis (MS), its progression, and the response of patients to treatment.

For the past decade, the author has served as senior computer scientist and technical director at the Buffalo Neuroimaging Analysis Center (BNAC, University at Buffalo, NY, USA), a multi-disciplinary institute specializing in quantitative MRI analysis. His role in the BNAC is one of independent conceptual leadership with regard to the application of computer science to MRI analysis techniques.

In collaboration with Prof. Robert Zivadinov, whose expertise is in clinical neurology, the author has developed, adapted, and applied a range of MRI image analysis algorithms, designed to accurately quantify *in vivo* changes in the brain of MS patients, thus advancing both the work of BNAC and the discipline as a whole. Throughout his tenure, he has also served as the key technical planner in the design of analysis protocols for over 70 individual studies.

The purpose of this submission is not to exhaustively catalogue this work, but rather to provide evidence supporting conferral of a degree of PhD by published

work. Therefore, the specific subset of papers selected is based on the cohesiveness of the work reported and the fact that it demonstrates substantial independent leadership and involvement by the author. While most of these papers relate to MS in one way or another, a paper on the assessment of stroke is also included, because the methodologies developed in that study laid important groundwork for subsequent investigations of MS patients.

1.4 Justification for the work

Because it is non-invasive, non-ionizing, and can distinguish with great accuracy pathological changes in soft tissue, MRI has become widely used in the diagnosis, management, and study of neurological conditions such as MS. Although it has already facilitated major clinical and scientific advances, it is nonetheless a relatively new technology, with the first commercial whole-body scanners only appearing in 1980 (Ai et al. 2012). While these early machines were primitive, over the subsequent decades rapid technological improvements have been made which allow MRI to quickly and accurately map many aspects of tissue at resolutions on the order of a cubic millimetre. Consequently, MRI is capable of harvesting a wealth of raw data from patients and delivering new insights into the nature of many neurological diseases.

While advances in MRI physics have been rapid, they have not consistently translated into improved ability of clinicians and researchers to interpret the raw data into meaningful information. In practice, within a single one-hour scanning session, ten or more full conventional and/or non-conventional 3-dimensional tissue maps can be produced, each emphasizing different aspects of the local chemical environment and/or pathophysiology. Indeed, a single MRI scanning session may produce over a gigabyte of data comprising multiple intensity readings in millions of voxels (3D pixels). Without sophisticated quantitative tools to cope with this vast amount of information, many radiologists' only recourse is to make qualitative judgements or simple hand calculations. However, this approach is subjective and prone to operator error and misinterpretation (Drew et al. 2013). Furthermore, the use of qualitative assessment can mean that only a small proportion of the data collected is utilized, with much potentially useful information discarded. Consequently, quantitative methodologies have been developed in recent years that enable

the data to be collected and analysed more reproducibly and in much greater detail.

By conferring the ability to synthesize data, elucidate subtle effects, and refine and standardize outcome measures, quantitative MRI image analysis holds great potential to lead advances in both basic science and clinical treatment. With its roots in signal processing, statistics, and pattern recognition, it has become a field in its own right (Dhawan 2011; Dougherty 2009), and has already seen many early successes. Nevertheless, it is a young field, with great potential for improvement and many unexplored frontiers.

The aetiology and pathophysiology of MS is poorly understood, with the result that the full clinical relevance of MRI data collected may not be known. Although MS is known to be associated with many MRI factors, the clinico-radiological paradox – the fact that MRI findings do not correlate well with clinical outcome – remains troublesomely unresolved (Barkhof 1999; Barkhof 2002). Understanding the precise impact of therapy on brain atrophy and tissue integrity is complicated by factors like pseudoatrophy (Zivadinov, Reder, et al. 2008) and measurement error, and the best way to deal with these remains an open question. Even seemingly straightforward issues like quantifying lesion burden remain fraught with difficulties. How does one distinguish between lesion and more subtle dirty-appearing white matter (WM)? How does one quantify pathology with extremely diffuse edges? Consequently, there is need to develop new, accurate techniques that can reliably quantify tissue integrity, lesion burden, and atrophy in greater detail, so that disease onset and progression can be better understood and therapeutic impact can be better assessed.

In addition to the need for new analysis techniques, it is important to ensure consistency in the fast moving field of MRI. For example, how does one deal with subjective decisions made by different MRI analysts, regarding the choice of regions of interest (ROI) to investigate? Furthermore, how does one ensure that increasing the power of the MRI machine (e.g. from 1.5T to 3T) does not completely invalidate cross-group comparisons? These and many other 'softer' issues have not been sufficiently discussed. However, given that clinical and scientific findings are often based on the outcome of MRI image analysis, it is

important that consistency of analysis is maintained and that robust protocols are developed.

1.5 Aims and objectives

The overall aim of the work was to develop, adapt, or refine quantitative image analysis methodologies to improve both the precise characterization of MS pathology and progression and the assessment of therapeutic interventions.

The specific objectives of the work were to:

(i) Develop a new methodology for accurately quantifying volumetric loss in the optic nerves of MS patients.

(ii) Develop an improved methodology for accurately quantifying grey matter (GM) atrophy in MS patients.

(iii) Develop an improved methodology for assessing the heterogeneity of microstructural tissue damage in MS patients.

(iv) Develop an improved methodology for accurately and reliably quantifying pathological intracranial features with diffuse boundaries on MRI images.

(v) Develop a new methodology for accurately quantifying and localizing ongoing demyelination and remyelination in the whole brain of MS patients.

(vi) Develop a standard methodology for quantifying iron deposition in specific structures in the brains of MS patients.

(vii) Evaluate the extent to which increasing the field strength of the MRI machine influences the outcome of MS lesion analysis.

(viii) Evaluate the extent to which operator variability regarding choice of ROI influences the quality of MRI lesion measurements in MS patients.

Chapter 2 – Background and context

This chapter presents an overview of the background research underpinning the work of the author and his collaborators. This work is interdisciplinary in nature and overlaps between three key areas: MRI, MS, and image processing and analysis. Therefore, to provide context and to familiarize readers with these fields, a brief overview is provided below. For readers completely unfamiliar with these fields or desiring more information, Appendix D reviews each in further depth and provides additional key references.

2.1 MRI

MRI is a young but extremely versatile medical imaging modality. Through a combination of superconducting magnet, gradient-producing coils, and radiofrequency (RF) coils, it is capable of interrogating tissue protons to elucidate many details about their chemical environment (Haacke 1999). The technique is three-dimensional, non-invasive, and non-ionizing. Although more time consuming than computed tomography (CT) it can produce a unique array of contrasts, each of which can provide specific clinical or scientific information not otherwise obtainable (Fig 1). Classically, these contrasts included proton density (PD), T1, and T2, but many more have been discovered as well. Of particular note are diffusion (Le Bihan 2003), which can quantify the molecular motion of water, and magnetization transfer (Henkelman et al. 2001), which is particularly sensitive to macromolecules.
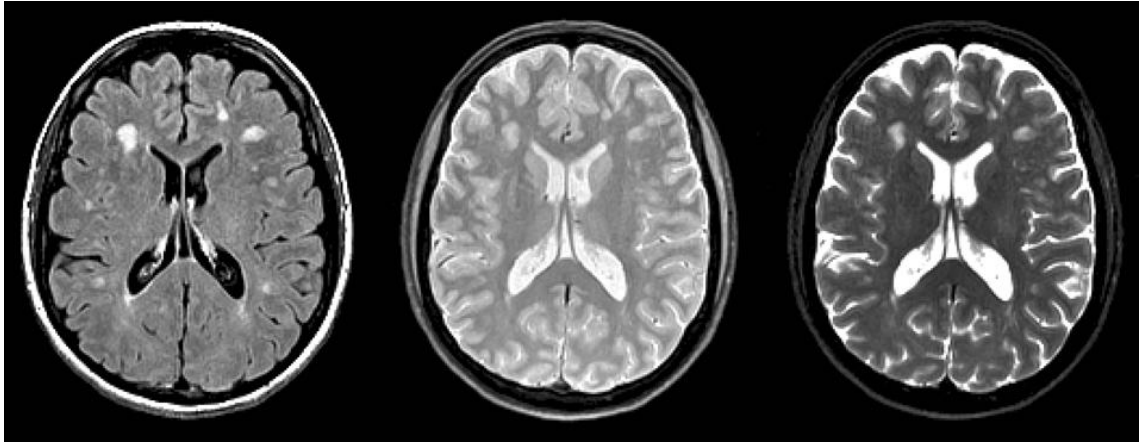
Figure 1: Conventional MRI sequences showing different tissue properties: FLAIR (left), proton density (center), and T2 (right). Images have been co-registered to show exactly the same slice, but each has a different contrast mechanism.

Most neuroimaging MRI scanning sessions produce five to ten different three-dimensional image volumes with resolutions on the order of one to five cubic millimetres per voxel. They are usually interpreted qualitatively by trained radiologists, but they can also be directly digitally transferred for more rigorous quantification.

2.2 Multiple sclerosis

Multiple sclerosis is a chronic, often debilitating disease of the central nervous system, thought to affect more than two million individuals worldwide (Rosati 2001). It is the most common neurologically disabling disease in young adults, and currently has no known cure. It usually begins with a relapsing/remitting phase followed by later conversion to progressive deterioration, but other courses are possible and the effects vary widely between individuals (Poser et al. 1982).

The hallmark pathology of MS is the presence of focal inflammatory demyelinating plaques (Fig 2) that disseminate in both time and space. These lesions can appear in many areas, and are common in periventricular, juxtacortical, infratentorial, and spinal regions (Polman and Reingold 2011). During the acute phase of their formation, there is usually evidence of blood-

brain-barrier breakdown. Over time, many result in permanent gliosis and diminished or absent axonal conduction (Frohman et al. 2006).
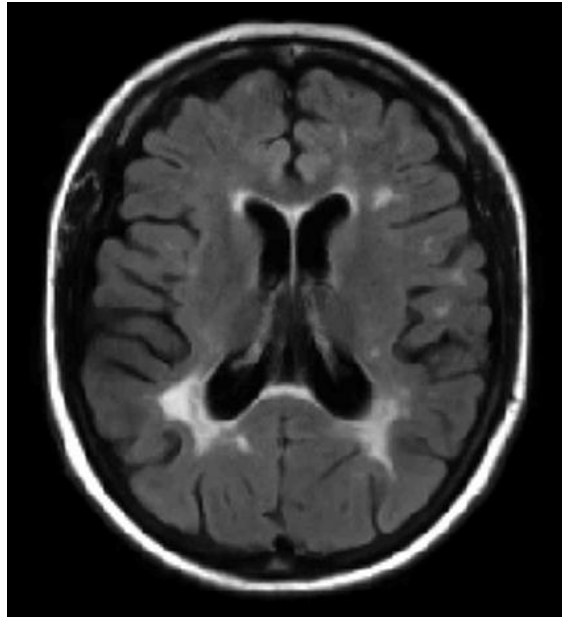
Figure 2: FLAIR image of a representative secondary progressive MS patient showing hallmark focal lesions, predominantly in the periventricular regions. Additionally, enlarged ventricles and thinning cortical gyri indicate brain atrophy.

Since MS was identified by Jean-Martin Charcot in 1866, many disparate factors have been implicated in its aetiology and progression. Environmental studies have revealed associations with latitude, sunlight, infectious diseases (particularly Epstein-Barr), vitamin D intake, and smoking (Ascherio and Munger 2007a; Ascherio and Munger 2007b). A genetic component is also implicated based on higher co-occurrence in twins (Willer et al. 2003), and genome-wide studies have identified specific candidate loci (Hafler et al. 2007; De Jager et al. 2009). Immunology has provided many insights, including the role of CD4+ and CD8+ T cells, and B cells are currently receiving increased attention (Kasper and Shoemaker 2010). Histopathology has also shed much light, demonstrating widespread damage and neurodegeneration beyond simple demyelinating focal lesions (Trapp and Nave 2008) as well as the presence of substantial GM lesion burden (Geurts et al. 2009). However, despite significant progress in all these areas no single factor has been fully explanatory, and a comprehensive disease model remains elusive.

Although a cure has not been found, treatment options have expanded greatly, and have now gone beyond symptom management to actual slowing of disease progression. The most popular therapies are the beta interferons and glatiramer acetate, but others like natalizumab (Polman et al. 2006) and fingolimod (Kappos et al. 2010) have also proven valuable, and even more such as alemtuzamab (Coles et al. 2008) and ocrelizumab (Kappos et al. 2011) are on the horizon. None of these drugs are without side effects, though, and individual responses often vary.

2.3 Medical image processing and analysis

For much of its history, neurological MRI was mainly qualitative in nature, but more recently clinicians, MRI physicists, and researchers have begun to draw on the fields of statistics, image analysis and signal processing to develop new methodologies. However, medical images pose their own problems, and so adaptation of general algorithms developed in other fields to neurological MRI remains a major challenge. Nonetheless, substantial progress has been made.

MRI is capable of producing a wealth of raw data, but the transition from this raw data to meaningful information is not always straightforward. A single MRI scanning session may produce multiple intensity readings for over nine million voxels (Fig 3). Consequently, a qualitative approach to MRI analysis is likely to result not only in diagnostic inconsistency, but also in much valuable data being discarded. Furthermore, even if it were possible to fully evaluate all of this data by eye, there would remain the difficult question of which pieces of information are clinically relevant and how they interrelate. Quantitative MRI image analysis can address both of these issues. By bringing modern medical image analysis techniques to bear, it is possible to process the massive raw datasets provided by MRI quickly and reliably. Because it can distil raw data into specific metrics, quantitative MRI can also help address questions of relative importance and interrelation by facilitating the use of statistical modelling and/or data mining techniques to correlate MRI findings with other meaningful clinical outcomes like physical disability and cognition. Such applications range from the relatively intuitive – e.g. the realization that lesion counts may not be as important or predictive as volumetric measurement in MS (Fisniku, Brex, et al. 2008) – to the

completely new and exciting, as in structural and functional connectomics (Bullmore and Bassett 2011).
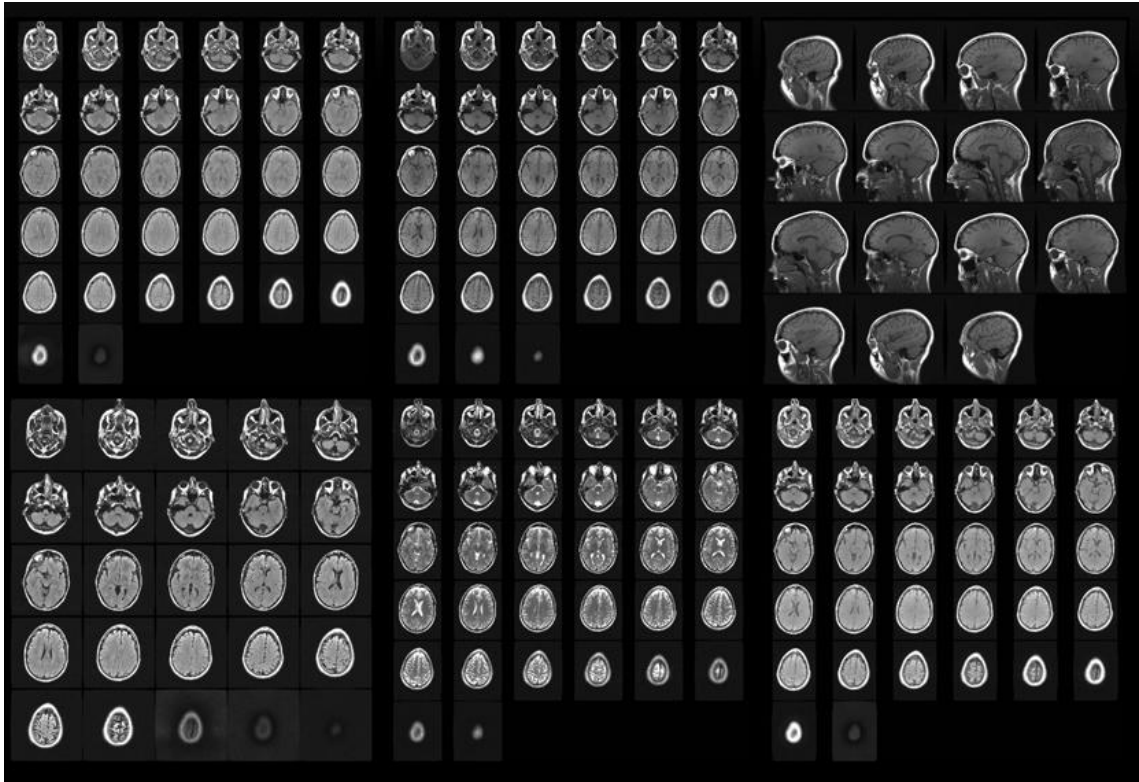
Figure 3: Visual overview of typical MRI data produced by a single, clinical scanning session. (N.B. Research protocols can comprise much more data.)

To a reasonable first approximation, medical image analysis can be split up into pre-processing and processing stages. Pre-processing is concerned with such tasks as noise attenuation, artefact correction, and post-acquisition alignment. Removal of noise can be accomplished in many ways, from simple smoothing (Gonzalez and Woods 2008) to more complex methods like anisotropic diffusion (Black and Sapiro 1998). Artefact correction includes such techniques as bias field removal (Sled et al. 1998) and distortion correction (Andersson et al. 2003). Post-acquisition alignment is particularly important, as it allows data from multiple contrast types, time-points, or individuals to be synthesized. Consequently, it is a relatively mature field with robust techniques for both co-registration (linear, within-individual alignment) (Jenkinson and Smith 2001) and normalization (non-linear warping for inter-individual alignment) (Klein et al. 2009). Furthermore, high-resolution, modern MRI atlases have been developed

that allow for the alignment of scans into a common space that facilitates collaboration between researchers and uniform understanding of scientific findings (Mazziotta et al. 2001). In general, pre-processing techniques take images as input and produce modified or refined images as output.

Processing techniques are more challenging to summarize, as they are highly varied and application-dependent. They can be qualitative, also producing images as output, or quantitative, producing well-defined metrics as output. In either case, they can provide an important alternative or complement to qualitative radiological reads. General techniques include contouring, segmentation, and anatomical parcellation. Contouring, either automated or semi-automated, can be used to identify focal lesions, and has been substantially refined via the use of edge-finding techniques and iso-contours (Grimaud et al. 1996). Segmentation is usually fully automated, and involves the classification of individual voxels into tissue categories (e.g. GM, WM, CSF), sometimes including pathology (Van Leemput et al. 2001; Zhang et al. 2001; Ashburner and Friston 2005). Anatomical parcellation uses varying models and techniques to reliably extract specific structures such as deep GM nuclei or specific cortical gyri (Patenaude et al. 2011; Fischl 2012).

In addition to general techniques, more contrast-specific processing algorithms have been developed, including diffusion-based fibre tracking (Ciccarelli et al. 2008), magnetization transfer ratio-based macromolecular quantification (Filippi et al. 1998), and BOLD-based functional MRI analysis (Jezzard et al. 2002). These specific techniques continue to evolve in tandem with improvements and discoveries in MRI sequence development.

2.4 MRI image analysis in MS – overview of the field

Concurrent with and interwoven into advances in epidemiology, immunology, and histopathology, MRI has already played a dramatic role in MS, where it has both answered and generated many questions. Given the multifactorial nature of MS, it is in many ways an ideal target for quantitative MRI research since it can simultaneously provide information about so many different aspects of CNS tissue (Fig. 4).
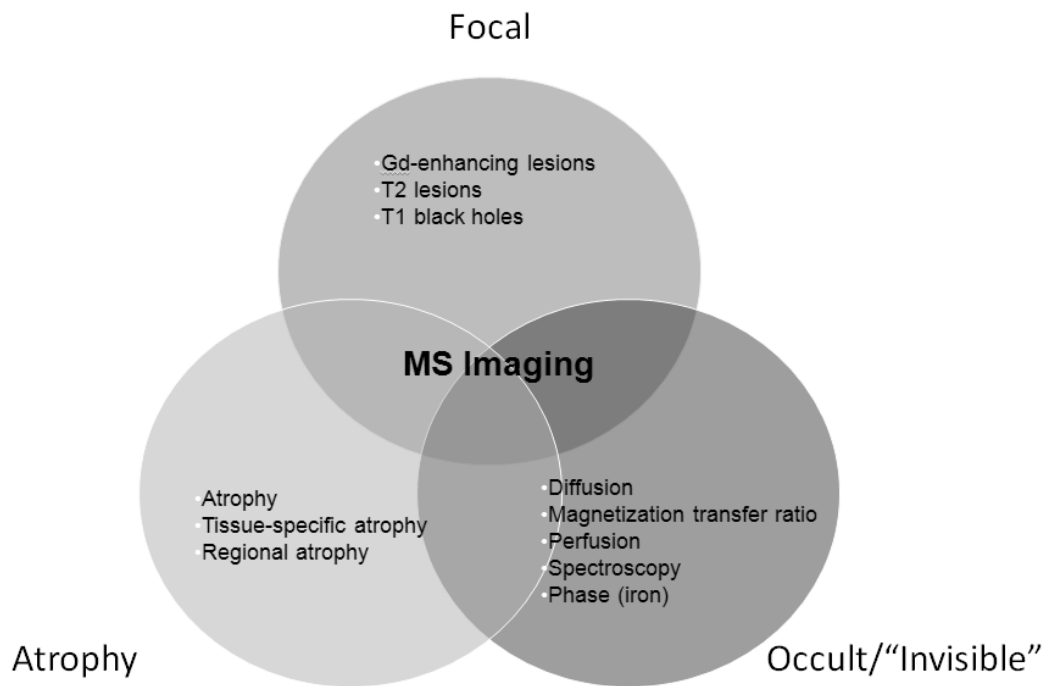
Figure 4: MR imaging in MS has been used to elucidate many different aspects of the disease. Although there are many exceptions, most current techniques can be broadly categorized as measuring focal pathology, tissue atrophy, or "invisible" damage.

Focal pathology (in the form of lesions) was the earliest recognized element of MS, and remains one of the most important clinical outcomes. From a clinical perspective, MRI sensitivity to lesions has substantially improved diagnostic criteria and differential diagnosis (Miller and Weinshenker 2008; Polman and Reingold 2011). In many clinics, it has also become a common standard of care to perform routine MRI to assess neurologically silent progression and assist in evaluating response to treatment. There has even been some evidence that lesions could be used as a primary endpoint in clinical trials (Sormani and Bruzzi 2013), but this has been debated (Rudick and Cutter 2013). From a research perspective, the spatial and temporal resolution achievable with MRI has led to a much better understanding of the formation, progression, and eventual fate of lesions (Ciccarelli et al. 1999). Also of particular interest, through the interplay of histopathology and MRI, chronic T1 hypointensities ("black holes") have been identified as an important subset of lesions that are more indicative of axonal loss and neurodegeneration than standard T2-hyperintense lesions (van Walderveen et al. 1998). Remaining problems in this

area mainly stem from difficulty in standardizing lesions measures across scanners and operators (Grimaud et al. 1996; Zijdenbos et al. 2002), although novel correlative approaches still continue to shed light on the disease process (Gourraud et al. 2013) and the precise relationship between lesions and atrophy is still under investigation (Mühlau et al. 2013).

Another aspect of MS that has been greatly illuminated by MRI is atrophy. Some patients show extensive atrophy late in the disease, and its precise role has been historically not well understood. With quantitative MRI analysis, it has become possible to more precisely characterize tissue atrophy and to reduce it to precise numerical quantities. This has led to the understanding that MS is not simply a focal inflammatory disease, but also a neurodegenerative condition (Chard et al. 2002; Losseff et al. 1996; Miller et al. 2002; Simon et al. 1999). Typically, young healthy adults lose approximately 0.15% of their brain volume per year (Ge et al. 2002), whereas an annual loss of 1.5% or more has been observed in MS cohorts (Ge et al. 2000; Zivadinov et al. 2001). As such, assessment of atrophy is now increasingly becoming an important end-point in MS therapeutic trials (Filippi et al. 2004; Filippi et al. 2001; Rudick et al. 2000). One notable finding has been that pathology in MS is markedly different between GM and WM (Geurts and Barkhof 2008), and this has led to increased interest in measuring their individual atrophy dynamics. Atrophy occurs at broadly different rates between the GM and WM (Ge et al. 2001; Sanfilipo et al. 2006). GM volume appears to have a closer relationship with disability (Fisniku, Chard, et al. 2008) and cognitive impairment (Benedict et al. 2006; Sanfilipo et al. 2006; Riccitelli et al. 2011) than is the case with the WM, although both are clearly important (Bodini et al. 2009; Bodini, Cercignani, Khaleeli, et al. 2013; Papadopoulou et al. 2013), and less GM pathology is seen in benign MS (Calabrese et al. 2013). The ability to reliably and independently detect subtle GM and WM volumetric changes *in vivo* is therefore an issue of critical importance, with some arguing that atrophy should be "upgraded" to a gold standard outcome measure in MS treatment (Rudick et al. 2013). However, despite the acceptance of atrophy as a key component of MS, its precise nature, cause(s), and mechanism(s) are still not fully understood. A number of possibilities have been discussed and explored, (Trapp and Nave 2008;

Zivadinov, Reder, et al. 2008) but further work is still needed to fully understand the phenomenon.

Beyond lesions and atrophy, MRI has also been extensively used to investigate more subtle, "occult" tissue changes in MS. For example, myelin is a critical target for both research and therapeutic intervention, and as MRI techniques have improved, there have been some promising approaches to achieve more myelin-sensitive imaging. Although no MRI marker has been found to have high specificity, one technique, MTR (Berry et al. 1999), achieves good sensitivity by taking advantage of the fact that water bound to large macromolecules, particularly myelin, responds with resonant behaviour to a wider range of radio frequencies than free water. By comparing images with and without the additional application of an off-resonant pulse, the local concentration of macromolecules can be indirectly inferred. MTR has been used to successfully detect myelin changes in lesions (van Waesberghe et al. 1998; Brown et al. 2012). MTR changes have also been observed in normal-appearing tissue (Filippi, Campi, et al. 1995; Bodini, Cercignani, Toosy, et al. 2013), but it is not clear whether these represent changes in myelin content or other processes such as axonal loss, since primary demyelination is not known to occur outside lesions. Similarly, diffusion imaging has also been used to evaluate the microstructure of tissue and to detect *in vivo* myelin changes (Ciccarelli et al. 2003), and diffusion tractography has shown how connectivity changes can affect cognition (Bozzali et al. 2013). Beyond this, MS spectroscopy has also been used to investigate changes in specific metabolites in the CNS (Srinivasan et al. 2005; De Stefano et al. 2007; Ciccarelli et al. 2007), although its resolution remains relatively crude. More recently, interest in iron dynamics in MS have also been revived due to the emergence of new phase-sensitive MRI techniques (Deistung et al. 2013; Zheng et al. 2013).

Although significant advances in MS image analysis have been made on all these fronts in recent years, there are still significant 'holes' in the knowledge base. For example, precision in lesion measurements remains challenging, the early detection of atrophy is extremely difficult, and dynamic myelin changes are difficult to localize and quantify *in vivo*. From a broader perspective, and perhaps most importantly, the clinico-radiological paradox remains unresolved.

Consequently, there is need to develop new methodologies with which to better understand and interpret MRI data regarding MS.

2.5 MRI image analysis in MS – statistical modelling

Given these challenges, the author initiated a programme of work geared toward addressing them (see Chapter 3 for details). Although individual problems require individual solutions, in broad terms the author applied statistical modelling techniques to address many of the objectives outlined in section 1.5. Statistical modelling provides a rigorous and useful framework, and allowed the author to draw on much existing work. In particular, the concepts of regularization, robust statistics, and multiple comparison correction were utilized. In order to aid the reader, these concepts are briefly introduced here. As this is a summary and synthesis document, formal mathematical rigor will be avoided in the interest of providing a more intuitive, high-level overview.

2.5.1 Models and regularization

MRI researchers and clinicians are consistently faced with a trade-off between scanning time, noise, and spatial resolution. Since time is limited, these last two factors are almost always compromised to some degree, creating significant challenges for quantitative image analysis. In the absence of any information about the underlying system, these challenges would be overwhelming. Fortunately however, this is not the case, because nearly all real-world MRI images have some underlying structure to them that can be exploited. For example, neurological MRI images usually have a great deal of spatial consistency – although there are many transitions between tissue types, they are almost always at least contiguous. Perhaps surprisingly, even this simple observation can dramatically improve many quantification approaches. Of course, exploiting such information is not completely free, and one must be careful to balance *a priori* notions against actual observations.

This balance is usually referred to as the bias/variance trade-off, and is a relatively mature field in mathematics and optimization theory (Hastie et al. 2011). The abstract intuition described above is formalized in a number of more mathematically precise and rigorous ways in different applications, but most fall under the categories of regularization or model fitting. In the context of the

author's work, one of the most important and illustrative of these approaches is the hidden Markov random field (HMRF) model (Winkler 2003).

At the core of this model is the Markov property, which roughly asserts that the probability of each state transition for a given system only depends on the current state, and not on any prior states. The Markov property is a significant over-simplification of the real world, but is often surprisingly tenable in many practical applications. For example, disparate areas like gambling, stock markets, and speech can all be modelled as Markovian processes. In the current context, the Markov property is also important from a spatial perspective rather than a temporal one, and the conventional Markov chain is generalized to the concept of a 2- or 3-dimensional Markov random field. In essence, taking the Markov assumption allows tissues on MRI to be modelled in such a way that only the direct neighbours matter in classifying or segmenting a given voxel. In turn, this allows for a direct calculation of how likely a given configuration is *a priori*. Finally, by combining this *a priori* spatial configuration information with observed voxel intensities, a maximum *a posteriori* (Bayesian) estimate can be created. The Markov random field approach was introduced for MRI tissue segmentation in (Held et al. 1997) and substantially refined by (Zhang et al. 2001), who added the "hidden" aspect to explicitly model the noise in image acquisition.

Using HMRFs or other similar models, the ultimate balance between bias and variance can usually be reduced to a single, clear parameter choice. In theory, Bayesian approaches can be used to calculate the optimal value for this parameter to minimize the total actual expected error. In practice, though, this is difficult to determine analytically and so is often determined using empirical approaches like simulation, bagging, cross-validation, or L-curve modelling.

2.5.2 Non-parametric statistics

The study of non-parametric statistics is a very broad field, of which two particular elements are primarily relevant to the author's body of work: robust statistics and permutation testing.

Robust statistics (Huber and Ronchetti 2009) provide a precise and rigorous means for estimation of distribution parameters in the face of substantial noise

or even minor bias. These are generally split into estimators of location and estimators of scale. Estimators of location are analogous to the conventional use of the mean, and can be as simple as trimmed mean or median and as complex as Huber or Hempel ψ-function based M-estimators. Estimators of scale are analogous to standard deviation, and can similarly range from familiar approaches like interquartile range and median absolute deviation to more unique approaches like Q_n and S_n estimators. The Q_n estimator is of particular interest, because it is efficient, does not depend on a location estimate, and is computationally feasible to calculate. It is essentially the median value of all the pairwise differences between all items in the set being studied.

Permutation testing solves a different but related problem: how to assign formal significance values when the underlying distribution is not known. Conventional techniques like Student's t-test assume a known family of parameters underlying the null hypothesis, such that the probability of observing a given value by chance can be analytically determined. In contrast, permutation testing makes almost no assumptions about the nature of the underlying data, and determines the likelihood of observing a given value empirically from the data itself. For example, the equivalent of a t-test is performed by continually re-assigning group labels, and determining in what percentage of random labellings the observed difference is more than in the true labelling. Permutation testing is not a new concept (Pitman 1937), but is very computationally intensive. The rise of cheap, powerful computers have therefore contributed to a marked rise in their use and applicability. Even today, though, permutation tests still usually only use a percentage of the total number possible combinations via Monte Carlo sampling.

2.5.3 Multiple comparison correction

In theory, performing statistical tests per voxel can provide unparalleled regional information that would otherwise be "lost in the average". However, it also severely exacerbates the problem of multiple comparisons – there are millions of voxels in many MRI images. Conventional Bonferroni correction is also useless, resulting in highly over-conservative conclusions. The key to solving this problem is that meaningful results generally do not occur in single isolated voxels, but rather in clusters. This observation can be exploited to create more

appropriate and suitably powerful statistical tests. One of the early (and still widely used) approaches in this area is Gaussian random field (GRF) theory (Worsley et al. 1996). Like nearly all parametric statistical approaches, GRF theory makes certain assumptions about the data in exchange for power. However, these assumptions may not be valid for a given experiment, and are sometimes difficult to defend in novel research areas where data behaviour is not yet well understood. In particular, application of GRF theory generally requires a strong *a priori* hypothesis about the degree of tissue change that will be observed, in order to choose an appropriate cluster-forming threshold. Permutation testing is a more recent, non-parametric alternative to this that can provide more appropriate results while making minimal assumptions. With this approach, group labels are randomly permuted and values are calculated in order to build up an empirical null distribution. It has been used with great success to make statistical inferences in voxel-wise statistical mapping problems (Nichols and Holmes 2002).

Chapter 3 – The author's contribution to the knowledge base

This chapter presents an overview of the author's original work (as set out in his published papers) and his contribution to the field of MS research.

3.1 Narrative framework

In order to contextualize the author's work and to aid the reader in understanding the author's contribution to the knowledge base, a narrative framework (shown in Table 3) has been constructed, which describes in a cohesive manner the body of the work presented. For ease of reference, this framework groups the author's various papers into distinct sections, with each section dealing with related 'contributions', which are mapped against the specific objectives identified in section 1.5.

The overall goal of the author's work has been to advance the use of quantitative analysis of MRI images in order to better understand MS. Over the years in which the author has been working in this field, pursuit of this goal has led him to explore many avenues of research, with much of the work developing organically, rather than fitting any pre-planned scheme. However with retrospective reflection, it is clear that the body of work broadly fits into the three distinct, but related sub-areas (tissue atrophy; occult pathology; and focal pathology), outlined in Table 3. Taken together, these sub-areas highlight different aspects of MS pathology, and serve as a contextual framework for the

author's work. Prior to 2007, the author also published a number of preliminary studies associated with the quantification of tissue atrophy, which laid the foundation for later more sophisticated work. Because of the formative nature of this work, these preliminary studies are reported in a separate section in Table 3.

Table 3: Narrative framework of the author's published work

Section	Description	Objectives met	Relevant papers
3.2 Preliminary work	This section presents an overview of the author's early work, which formed the basis of the later studies relating to tissue atrophy.	N/A	1, 2, & 3
3.3 Tissue atrophy	Neurodegeneration, characterized by tissue atrophy, is a major component of MS disease progression, and is predictive of future disease course, disability, and cognitive outcome. As such, it is an important target of research and therapeutic development. Therefore, methods for its accurate quantification and tissue-specific assessment are highly important. This section presents an overview of the author's work relating to tissue atrophy. In particular, a new methodology for accurate optic nerve atrophy quantification is presented, together with a dramatically improved methodology (algorithm) for accurately quantifying GM atrophy in MS patients.	(i) & (ii)	7 & 11
3.4 Occult pathology	The term 'occult pathology' refers to pathologies that are beyond the range of ordinary knowledge and understanding. Numerous investigations have demonstrated that the so-called 'normal-appearing' tissue in MS is often damaged in subtle ways that are not immediately apparent on conventional MRI. These include diffuse white matter axonal injury, abnormal iron accrual, and cortical demyelination. This section presents an overview of the author's work relating to some of the occult pathologies associated with MS. In particular, an improved methodology is presented for quantifying pathological intracranial features with diffuse boundaries on MRI images, along with an improved methodology for assessing the heterogeneity of microstructural tissue damage in MS patients. Also presented is a new methodology for accurately quantifying ongoing demyelination and remyelination in the whole brain of MS patients. Finally, a standard methodology for quantifying iron deposition in specific structures in the brains of MS patients is	(iii), (iv), (v) & (vi)	4, 5, 6 & 9

	described.		
3.5 Focal pathology (error minimization)	Although focal lesions are one of the best understood aspects of MS, their precise quantification remains difficult to standardize and reproduce. Errors can occur due to subjective decisions regarding the choice of regions of interest (ROI) to investigate. Also, the field strength of the MRI machine influences the outcome of lesion analysis. This section presents an overview of the author's work relating to error minimization and standardization in the assessment of focal lesions in MS. In particular, operator variability regarding choice of ROI is evaluated, as is the impact of increasing the field strength of the MRI machine on the outcome of lesion analysis.	(vii) & (viii)	8 & 10

3.2 Preliminary work

Qualitative and quantitative MRI have provided much insight into atrophy in MS, but many key questions still remain. For example, precisely how and where does atrophy occur in MS? What portion of atrophy is related to demyelination compared to actual loss of cell bodies? How does MS atrophy relate to clinical and cognitive outcomes? What is the best way to measure atrophy? Can atrophy be accurately detected over smaller time periods and/or in smaller groups of subjects in order to reduce the risk entailed in experimental clinical trials?

In response to the above questions, the author undertook the work reported here. Its beginnings can be traced to 2005, with some crude but important investigations into the basic nature of atrophy in MS and the factors affecting its measurement. In (Carone et al. 2006), a technique developed by (Dade et al. 2004) was adapted to study region-specific atrophy in MS based on parcellation via familiar Talairach landmarks. This work confirmed that GM atrophy was present in MS, and that it was not necessarily uniform. However, the regions studied were relatively coarse and did not directly follow meaningful anatomical boundaries. In addition, by modern standards the GM/WM segmentation used was relatively imprecise. At the same time, the author and his co-workers evaluated the impact of technical aspects like coregistration on a variety of brain

atrophy measurement techniques, and also underscored the significant difficulty of measuring brain changes in the short term (Fritz et al. 2006).

Building on this early work, it was hypothesized that a simple and highly reproducible proxy for GM atrophy might be a valuable tool. Third ventricular width was proposed, based on its location between the left and right thalami, and the author and his co-workers were able to demonstrate that this measure explained significant variation in a number of important neuropsychological tests (Benedict et al. 2006). However, it was recognized that although third ventricular width may be important, it was not sufficient due to the heterogeneity of atrophy in MS. The author and his co-workers therefore undertook additional work to develop reliable techniques for measuring other complementary aspects of atrophy, including that of the spinal cord (Zivadinov, Banas, et al. 2008). This preliminary work made it clear that studying small, well-defined structures could have substantial advantages, and additionally underscored the importance of precision in measurement. It also reiterated the need to better understand the actual nature of atrophy.

3.3 Tissue atrophy

3.3.1 Optic nerve and RNFL analysis

Optic neuritis is one of the most common initial symptoms of MS (Sørensen et al. 1999), with nearly two thirds of patients suffering at least one acute episode (McDonald and Barnes 1992). Therefore, a natural extension of the early preliminary studies was to focus on the optic nerve and retina. The optic nerve and retina are important targets for better understanding the precise nature of tissue atrophy. Unlike other peripheral nerves, the optic nerve is actually a direct outgrowth of the diencephalon, deriving from embryonic retinal ganglion cells. As such, it is myelinated by oligodendrocytes rather than Schwann cells. Furthermore, the retinal nerve fibre layer (RNFL) is formed from the axons of the optic nerve, and is normally un-myelinated (Nolte 2008). Therefore, this two-part system provides a unique opportunity to look at the same axons in two different places and states – one where they are myelinated and one where they are not – with very few additional confounds. Evaluating atrophy in both these structures at the same time thus has the potential to provide information about the specific atrophy related to axons as compared to demyelination.

Optical coherence tomography (OCT) already provides a fast, reliable, and non-invasive way to quantify RNFL (Huang et al. 1991). However, there were not previously any reliable methods for optic nerve atrophy quantification beyond simple ROI-based techniques. Such techniques, while easily applied, are highly subject to error due to motion, slice angle, and scanner artefacts. To address these concerns, and to ensure that differences seen were due to real anatomical change and not just measurement error, a new approach was required.

The author's solution to this problem was achieved via the creation of a shape-based model fitting technique (Frohman et al. 2009). In this type of approach, a parameterized mathematical model is used to adjust an idealized shape such that it best matches acquired data. In most cases, the motivation for this is twofold. First, it dramatically reduces the number of parameters involved. In theory, fitting something like the optic nerve is highly multidimensional, with one parameter per voxel (i.e., partial volume of the structure in the voxel). By changing to a parametric shape representation, the challenge can be greatly reduced from potentially thousands of parameters to just a few. Second, by imposing an *a priori* morphology to the structure to be quantified, shape fitting serves as a form of regularization. Like other Bayesian approaches, this regularization can greatly assist with precision in the face of noise and artefacts.

In this case, a circularly extruded cubic spline tube model was used, resulting in a total of only 9 parameters as a function of slice location (four for x, four for y, and one for diameter). At each point, a circular cross section was extruded with orientation determined by the spline's tangent vector and radius set by the equivalent model parameter. Given this model, fitting was performed in stages. First, cubic splines were fit to operator-identified points at the centre of the optic nerve on each slice. This fitting served both to characterize the path of the nerve and to smooth out any small errors in the manual point placement. In addition, deviation metrics were used to correct for movement between slices. After this, fitting of the diameter was performed by a brute force search through potential diameter values, by identifying the diameter value minimizing a gradient-matching cost function.

Overall, the approach was very successful, and atrophy measurement using this technique proved to be extremely reproducible. Specifically, it demonstrated a mean absolute scan-rescan error of only 0.06 mm (less than 5% of the nerve diameter). This precision is more impressive when considering the fact that it is only a little more than one tenth of the voxel width (0.49 mm).

3.3.2 Precise longitudinal GM atrophy quantification

Although atrophy has been suggested as a gold standard treatment outcome, others dispute this stance (Rudick et al. 2013). It is notable, though, that many of the counter-arguments involve technical difficulties associated with assessing tissue atrophy, rather than on atrophy per se. For example, pseudoatrophy (Zivadinov, Reder, et al. 2008) is an important phenomenon that results in paradoxically larger brain shrinkage with the initiation of therapy. Additionally, it is argued that atrophy is an end-stage effect that may not be apparent for years. Although these are important caveats, they are not insurmountable. Mounting evidence demonstrates that pseudoatrophy has less impact on GM than WM, so studying GM in particular may alleviate that problem. In sufficiently powered studies, volume changes are seen even in early CIS patients (Bergsland et al. 2012; Henry et al. 2008) so it seems that the inability to see short-term atrophy is likely a result of imprecise measurement rather than an actual lack of tissue changes. In this case, it would be extremely important to improve the precision of existing GM atrophy measurement techniques.

To address this issue, the author developed an improved algorithm for quantifying GM atrophy (Dwyer et al. 2012). Rather than work from the ground up, the widely accepted FAST tool (Zhang et al. 2001) and elements of SIENAX and SIENA (Smith et al. 2002) were taken as a starting point. Briefly, FAST works by applying a hidden Markov random field (HMRF) (Winkler 2003) expectation maximization (EM) model to classify tissue as GM, WM, or CSF and to estimate partial volumes within voxels. FAST's model is three-dimensional, though – concerned only with the spatial neighbours of individual voxels. In the current work, this model was extended to be four-dimensional in order to account for temporal neighbours as well. Although dealt with more formally in the referenced paper, the intuitive motivation for this is relatively straightforward. In a real set of MRIs, many voxels have intensities midway

between the average intensities of tissue classes. In some cases, it's contextually clear what these voxels should be called (e.g. a single voxel completely surrounded by other voxels that are clearly WM), and in these cases FAST will classify them appropriately. In other cases (e.g. along borders), the correct classification is not clear at all, even to an expert human reader. In this case, the choice between tissues will be seemingly random, although FAST will attempt to be spatially consistent about that random classification. In this second case, a key difference between an intelligent human reader and FAST is that when reviewing a follow-up image, the human reader will be conservative and not consider the tissue to be changing unless it is clearly different in intensity. FAST will simply classify it as whichever intensity it is closest to, even if 'closest' means it only changes in intensity by a fraction of a percent for voxels that were directly between classes.

Initial testing with this improved model was very promising, and it was subjected to more rigorous validation techniques including scan-rescan error assessment, tests of agreement with controlled simulation, and ability to dichotomize clinically meaningful subject groups. In all cases, the model demonstrated significant improvement compared to standard SIENAX, including lower errors (absolute deviation), better correlation with whole-brain (non-tissue specific) measures, and significantly improved effect sizes (up to a 68% increase).

3.4 Occult pathology

As discussed previously, post-mortem and histopathological investigations have revealed abnormalities in MS tissue *outside* of traditional focal lesions (Allen et al. 2001). However, such investigations are usually only performed in terminal cases and so cannot shed much light on the more dynamic aspects of any occult pathologies that may be present. Because of this, it is unclear whether such diffuse tissue damage is a primary or secondary effect, and what role such damage plays in disease progression and outcome.

3.4.1 Quantification of pathological features with diffuse boundaries

Although the work described in this section was undertaken in the context of ischemic stroke, the algorithms developed by the author laid important groundwork for later work in MS patients. In acute stroke, a primary option for

treatment is thrombolytic therapy, but this therapy carries a high risk of intracranial bleeding. Therefore, before proceeding it is extremely important to know whether any tissue is actually still viable but at risk. Commonly, this has been evaluated on MRI via so-called "perfusion-diffusion" mismatch, in which a diffusion abnormality is taken to represent unsalvageable tissue and a potential surrounding area of depressed but non-zero perfusion is taken to represent tissue that is salvageable but will become infarcted without intervention. The salvageable tissue is called the "ischemic penumbra". Because time is essential in stroke, an automatic approach for independently assessing volumes of infarct core and ischemic penumbra would be of great value, but available approaches based on simple standard deviation thresholding were highly susceptible to noise (Røhl et al. 2001; Takasawa et al. 2008). Also, unlike the infarcted core, the edges of the penumbra are often extremely diffuse; without a clear border, small variations in threshold can dramatically change volume. As such, assessment of the size of penumbra presents clinicians with a considerable challenge.

In order to address this problem, the author developed a more noise-resistant and statistically rigorous approach. As discussed previously, hidden Markov random field (HMRF) modelling is a powerful tool providing just these capabilities in situations where observations are spatially correlated (i.e., where neighbouring voxels aren't truly independent). However, an HRMF model generally assumes corrupted viewing of categorically different underlying classes, and in this case the penumbra class was not well defined. To address this problem, a unique combination of M-estimators (Yuan and Bentler 1998), abnormality indexing, and HMRF was employed by the author (Dwyer et al. 2008). An M-estimator was used to robustly calculate the intensity distribution characteristics of the infarct core and then a Mahalanobis distance metric was used to categorize the abnormality level of tissue. Finally, an HMRF model was used on this derived abnormality field to classify tissue. Intuitively, infarct core and contralateral normal tissue were made to "compete" for ownership of the intermediate data, with tissue ultimately more similar to infarct being retained as penumbra.

Although this work appears at first sight to be unrelated to MS, it provided important experience and insight into the difficult problem of precisely quantifying tissue pathology that is both subtle and without sharply demarcated borders. More specifically, it also confirmed the applicability of M-estimators in assessing empirical image intensity distribution parameters, which was critical in solving the Monte Carlo MTR change problem discussed below.

3.4.2 Microstructural tissue damage

An early portion of the author's work aimed to employ diffusion MRI to better characterize microstructural tissue damage in normal appearing brain tissue. As mentioned briefly above, diffusion MRI uses specially tuned, temporally separated gradients of opposite polarity to specifically impact mobile nuclei (Le Bihan 2003). The effects of these gradients vary with the extent of molecular diffusion, and are processed to produce a quantitative map of the degree of diffusivity. In a healthy human brain, diffusivity is highly restricted by hydrophobic myelin sheaths, and so is considerably lower in WM than in GM. Loss of myelin (or cellular structure in general) reduces the barriers to molecular motion, and so increases diffusivity locally.

In this work, diffusion metrics were evaluated by the author and his co-workers in the whole brain in addition to focal MS lesions (Tavazzi et al. 2007). Of particular note, the author introduced a novel metric: entropy. Whereas the conventional metric, mean apparent diffusivity (ADC), provides an estimate of the overall water diffusion in the tissue, entropy provides an estimate of the macroscopic variation in diffusivity levels. Although standard deviation is somewhat analogous, entropy is arguably more appropriate in this case. Although standard deviation and entropy are both minimized by all observed diffusivities being identical, they are maximized differently. Standard deviation is maximized by a bimodal distribution with peaks at the highest and lowest diffusion values. On the other hand, entropy is maximized by a completely uniform distribution where all possible diffusion values are equally observed in different tissue areas. Organized structure may well result in clusters of different diffusivity values, but disease is more likely to spread the distribution out. Confirming this theoretical concern, entropy was empirically demonstrated to be the best diffusion-based predictor of clinical outcome in the studied dataset.

Later work using the same technique demonstrated that entropy was also the best diffusion predictor of cognitive ability in MS (Benedict et al. 2007) and was also successfully used to show the microstructural effects of glatiramer acetate therapy (Zivadinov et al. 2011).

3.4.3 *In vivo* mapping and quantification of dynamic myelin changes

Remyelination is recognized as a competing process to destructive demyelination in MS (Bunge et al. 1961; Prineas and Connell 1979). So-called "shadow plaques" containing thin fibres have been shown to be partially remyelinated with a thinner sheath, and remyelinated axons have been found to have unique morphology (Prineas et al. 1993). However, nearly all of the information gleaned regarding this phenomenon is from cross-sectional data via post-mortem analysis or from animal models not perfectly equivalent to MS. Consequently, the timing and interplay of demyelinating and remyelinating processes *in vivo* in real MS patients was poorly understood, with the impact of therapy not yet precisely evaluated. Accurate *in vivo* detection of the active processes involved in demyelination and remyelination is difficult to achieve from a technical standpoint, primarily due to a lack of MRI specificity and very noisy images. Preliminary MTR-based work by (Chen et al. 2007) provided a threshold-based technique for detecting these changes, but it was restricted to areas within lesions. Because more subtle demyelination and remyelination may also occur outside of overt T2 lesions (e.g. in cortical lesions or potentially in pre- or peri-lesional WM), the author developed a novel, statistically rigorous method for quantifying them in the whole brain.

The work drew on pre-existing techniques for cluster-based voxel-wise inference, but faced unique challenges – whereas VBM and fMRI methods rely on consistency in effect between subjects, myelination changes can occur in unique locations in different patients. This necessitates evaluation of voxel-wise changes on a per-subject basis, and makes standard permutation testing impossible. It is feasible in this case to fall back to a GRF theoretic approach. However, given the potential for large areas of subtly altered myelin content, it is difficult to confidently choose a sufficiently meaningful cluster-forming threshold.

During this time another group published a method called TFCE, for "threshold-free cluster enhancement" (Smith and Nichols 2009). TFCE provides a means

to elegantly sidestep the cluster-forming threshold drawback of GRF theory by nonlinearly modifying statistic images based on their local cluster support at multiple scales. Unfortunately, though, TFCE had no known closed form probability distribution function, and permutation testing could not be applied to a single subject. To resolve this, the author created a new technique to make inferences using a novel Monte Carlo simulation coupled with an M-estimator for variance estimation. This resulted in the ability to reliably estimate the null distribution for TFCE values in a single subject, and to thereby assign meaningful p-values on a multiple-comparison corrected, per-voxel basis. The final technique was validated in three separate ways, using scan-rescan, simulation, and application to real clinical data.

3.4.4 Iron assessment

Another potentially intriguing front in the exploration of not-conventionally-visible MS pathology and progression is iron detection and quantification. Iron toxicity has been implicated in other diseases including Parkinson's (Dexter et al. 1991) and Alzheimer's (Connor and Menzies 1992), and some early histopathological work indicated abnormal iron deposition in MS (Craelius et al. 1982; LeVine 1997). Until recently, though, iron deposition was very difficult to image *in vivo*. Initial work with T2 shortening was promising, but was also limited by the relative non-specificity of T2 contrast (Bakshi et al. 2000). In the interim, new MRI techniques based on phase data have been proposed. Traditionally discarded in favour of magnitude data after Fourier reconstruction, phase data is exquisitely sensitive to magnetic field fluctuations imposed by the high susceptibility of iron. Unfortunately, it is plagued by serious and substantial artefacts including aliasing, background field contributions, and non-locality. However, modern processing approaches have been proposed that are capable of alleviating the majority of these artefacts and recovering much of the information available. These include the susceptibility-weighted imaging (SWI) technique (Haacke et al. 2004), and more recently QSM (Langkammer et al. 2012).

With these techniques, it has become possible to investigate iron deposition in MS with much more sensitivity than T2 shortening could provide (Haacke et al. 2009). However, two concerns encouraged a careful, regional approach rather

than whole brain measures. First, even in healthy people the time-course of iron deposition is substantially different in different regions of the brain (Hallgren and Sourander 1958). Second, myelin changes can cause non-iron phase changes in WM (Yablonskiy et al. 2012). Therefore, the author initiated a programme of work to develop a GM specific means to quantify phase on a structure-by-structure basis.

To accomplish this, the author employed a novel combination of FSL's FIRST subcortical parcellation tool (Patenaude et al. 2011) and SWI (Zivadinov, Heininen-Brown, et al. 2012). FIRST uses shape models as described previously to reliably identify left and right thalamus, caudate, putamen, pallidum, hippocampus, nucleus accumbens, and amygdala on high-resolution T1-weighted images. By implementing a system for automatically running FIRST and then realigning and warping the FIRST-identified regions into SWI image space, it was possible to extract region-specific phase data for each region. In addition, normative data was collected from healthy controls and used to implement region-specific phase thresholds indicative of abnormal iron content. Using this, a novel measure of mean phase of abnormal phase voxels (MP-APV) was introduced.

3.5 Focal pathology

Although focal pathology is the most classically well-understood aspect of MS, it remains challenging to make full use of the amount of data that even conventional MRI produces. In particular, reproducibility has been relatively low despite improvements to MRI. Whereas atrophy reliability is within the 1% range, even optimistic reports of inter-rater lesion agreement have indicated variances of near 7% (Filippi, Horsfield, et al. 1995), and others have reported up to 20% (Jackson et al. 1993). In terms of actual voxel-level agreement, kappa values below 0.7 have been demonstrated between centres (Zijdenbos et al. 2002). Beyond this, variation between sequence parameters and a trend toward higher field scanners makes inter-site comparison difficult at best. Consequently, there is considerable scope for errors to occur, which might compromise otherwise sound technical advances. In attempt to better understand the sources of potential errors, the author and his co-workers undertook the following studies involving MS patients.

3.5.1 Operator impact on lesion measures

Although automatic lesion detection has been, and continues to be an area of extensive active research, it has still not been widely adopted as a de facto method due to performance and precision issues (García-Lorenzo et al. 2013). As such, semi-automated operator-guided techniques have remained the gold standard – a situation that is likely to continue for the near future. Given that such semi-automatic classification still has a relatively high rate of error, improved precision is an important goal. To facilitate such improved precision, it is important to have meaningful measurements indicating the source(s) of imprecision. Otherwise, attempts to make improvements may easily be misdirected. Realizing this, the author and his co-workers initiated a programme of work aimed at providing an easily standardized and widely applicable means to separate out the key factors that contribute to error in lesion measurements.

An appropriate measure or set of measures should have certain desirable characteristics, including ability to separate operator ability from lesion burden, ability to direct improvements toward specific areas, and resistance to "averaging out" of measurement noise. The commonly used similarity index and variants like Kappa or concordance provide the last characteristic, but not the first two. Therefore, the author and his co-workers developed a new set of measures termed detection error and outline error (Wack et al. 2012). A key aspect of these measures is that they operate in a tiered lesion-wise/voxel-wise manner rather than a purely voxel-wise manner, in contrast to many other methods. Detection error reflects the lesion-wise probability of complete disagreement in marking a lesion, whereas outline error represents the degree of volume disagreement for lesions that were mutually agreed to exist.

These measures were applied by the author to a real dataset calculated multiple times by different operators, and were found to be considerably more independent of overall scan lesion volume than similarity index. They also provided better feedback to individual operators to allow them to improve by focusing either on recognizing lesions or more accurately delineating them, as necessary.

3.5.2 Scanner impact on lesion measures

Another major class of errors in lesion measures is scanner induced, either via hardware differences, changes in acquisition parameters, or both. From a hardware perspective, a major contributor to lesion salience is field strength. In particular, many research groups and clinical centres have moved from 1.5 tesla to 3 tesla scanners. Although the change in overall lesion volume was previously investigated (Sicotte et al. 2003), it was not clear if the impact of improved field strength was spatially homogenous. To address this question, the author and his colleagues adapted a recently introduced lesion probability mapping technique (Enzinger et al. 2006) to compare lesion maps calculated from paired images acquired on 1.5 and 3 tesla scanners (Di Perri et al. 2009). Specifically, the technique was modified to use pairwise statistics rather than group comparisons, and TFCE was substituted for the previous multi-resolution smoothing approach. The results agreed with previous work in showing a significantly higher lesion load at 3 tesla, but also revealed that lesions were more commonly differentially detected in specific regions – particularly the occipital horns of the lateral ventricles.

Chapter 4 – Discussion and conclusions

4.1 Evaluation, limitations, and improvements

The subset of the author's work described above produced a number of new algorithms and techniques that have already furthered the study of MS. In addition, many have continuing value and are still being used in ongoing studies.

Two of the author's algorithms dealt with methods for better understanding the nature of tissue atrophy in MS, and included a new methodology for accurately quantifying volumetric loss in the optic nerves (Frohman et al. 2009) and an improved methodology for accurately quantifying GM atrophy (Dwyer et al. 2012). This work was important in a number of ways. From a technical perspective, the optic nerve approach was an early and novel use of shape models in MRI of MS. Such shape models have continued to prove extremely useful in providing accurate and reliable semi-automated and automated analysis, and have notably been used in FMRIB's freely available and widely used FIRST software (Patenaude et al. 2011). From a scientific and clinical perspective, the optic nerve approach provided detailed data about the interplay between OCT, MRI, and visual outcome measures, and was important in elucidating the relative behaviour of optic nerve and RNFL. In particular, the low correlation of two highly precise measures – optic nerve diameter and RNFL thickness – was an important clue that volumetric changes are not entirely or necessarily even largely due to axonal loss. Along with the concurrent work of others, this helped to motivate the importance of studying WM and GM

35

separately. The improved method for quantifying GM atrophy is more practical, but equally important since it is capable of reducing the number of subjects required to demonstrate atrophy-reducing efficacy in clinical trials. This both helps to protect more patients from the risk inherent in trials, and also allows more funds to be diverted to continued exploration of other therapeutic avenues. In fact, the approach has already been used in a large study and demonstrated improved ability to discern treatment effects (Zivadinov, Bergsland, et al. 2013). Also, the technique can potentially detect changes earlier, which may help to address some of the concerns with atrophy as a gold standard discussed above.

As with any work, both of these techniques could have been improved, and more future work is required to build on the foundations laid. In particular, it would be beneficial to study the relationship between RNFL thickness and optic nerve diameter with more statistical rigor, and in particular to evaluate more complex models to better understand all sources of variance. Also, although initial investigations indicated that the optic nerve diameter is relatively constant over the area studied, the bias/variance trade-off of adding a linearly varying diameter parameter to the model was not systematically evaluated. In a similar vein, the GM quantification approach uses an empirical regularization parameter that could benefit from more formal study – e.g. with L-curves or related methods. Additionally, the emergence of very precise cross-sectional techniques (Dahnke et al. 2013), more anatomically targeted techniques (Vrenken et al. 2013), and availability of higher-quality MRI images may eventually obviate the need for such regularization.

Four more methods were developed by the author to better elucidate occult pathology: an improved methodology for accurately and reliably quantifying pathological intracranial features with diffuse boundaries (Dwyer et al. 2008), an improved methodology for assessing the heterogeneity of microstructural tissue damage (Tavazzi et al. 2007), a new methodology for accurately quantifying and localizing ongoing demyelination and remyelination in the whole brain (Dwyer et al. 2009), and a standard methodology for quantifying iron deposition in specific brain structures (Zivadinov, Heininen-Brown, et al. 2012). The first technique laid important groundwork for the others in addition to its direct benefit in acute

ischemic stroke. The method for measuring microstructural tissue damage was important at the time it was done, but has since been largely overtaken by diffusion tensor imaging (DTI) and tractography (Ciccarelli et al. 2008), which is capable of answering much more detailed questions and it is sensitive not only to the magnitude of diffusion but also to its directionality. However, the author's work did introduce a novel use of entropy that might still be worthwhile to apply to DTI-specific metrics.

The voxel-wise myelination detection technique developed by the author was more directly applicable, and led to the ability to investigate demyelination and remyelination not just in the natural course of the disease, but also on the effect of therapies on MTR. In particular, it was applied to a cohort of patients treated with either interferon beta 1-a or natalizumab (Zivadinov, Dwyer, et al. 2012), the results of which strongly suggested that natalizumab may significantly promote remyelination compared to interferon beta 1-a. Work with this method is still ongoing, and it is hoped that it can continue to be improved in a number of ways. First, given its basis on MTR it is sensitive but not entirely specific to myelin, but may also include some sensitivity to other macromolecules and therefore result in some false positives. Although other investigators have provided convincing evidence that this concern is not overly large (Giacomini et al. 2009), it would clearly be better to more completely eliminate it. Since this work was completed, other investigators have continued to expand the capabilities of MRI to do just this by detecting and quantifying myelin more specifically. In particular, novel sequences have been proposed that address these specificity concerns with MTR while still retaining high sensitivity, including qMTI (Janve et al. 2013), which may be clinically feasible (Dortch et al. 2011; Stikov et al. 2011), mcDESPOT (Spader et al. 2013), and robust myelin water fraction (Kwon et al. 2013). Since the majority of work on the VW-MTR technique described is related to statistical modelling, it is largely pulse-sequence-agnostic. It would therefore almost certainly be instructive to apply similar methodology to these improved myelin-specific images. However, it is important to note that none of these new methods are themselves perfectly specific to myelin, and the development of such an MRI marker is still an open research goal.

The iron assessment approach was also directly applicable. Although intuitively somewhat convoluted, the MP-APV measure was empirically shown to be well correlated with other non-iron metrics, including disability. In another study, it was shown that MP-APV may indicate iron deposition preceding atrophy in clinically isolated syndrome (CIS) patients (Hagemeier et al. 2012). This finding has also been replicated by another study (Al-Radaideh et al. 2013). Although cross-sectional results must be interpreted with caution, this observation is particularly intriguing, because it supports the idea that iron may have a causative or modulatory role in MS rather than simply being the end-product of myelin debris. Clearly, longitudinal studies are needed to substantiate this suggestion, but this work shows that they are certainly warranted. Further improvements to the technique will also help provide a more complete picture. In particular, phase imaging, though very sensitive to iron, is not perfectly specific and is also plagued by many artefacts. Modern approaches like quantitative susceptibility mapping (Langkammer et al. 2012) may provide more accurate measurements, and are currently being investigated by the author and his colleagues.

Finally, two methods were introduced by the author to study influences on calculating more conventional focal pathology, including a method for decomposing the sources of inter-rater disagreement in semi-automated lesion assessment (Wack et al. 2012) and a technique for comparing lesion visibility at different scanner strengths (Di Perri et al. 2009). Although it seems likely that in the longer-term future automated identification techniques will largely replace manual assessment of lesions, the decomposition work remains relevant. The slow progress of automated techniques demonstrates that an overly simplistic solution is unlikely, and the ability to better identify specific sources of error in a rigorous and quantitative manner can also help direct development of future classification algorithms just as it can aid in the training of human operators.

The method for comparing lesions at different scanner strengths clearly demonstrated the incompatibility of direct comparison between measures and, like (Sicotte et al. 2003), indicated the importance of using field strength as a statistical covariate or otherwise controlling for it in multi-centre trials. Furthermore, it lent weight to the idea that hyperintensities observed at the

posterior horns may be a normal variant, although the matter is not settled. Additionally, the pair-wise spatial approach might also be used for other applications, including measuring coil- and software-related changes.

4.2 Overall impact

The author's work as described above has had a significant impact on the work of the BNAC and that of other researchers, both in terms of direct contributions to the knowledge base and also in the creation of new techniques for future research.

Overall, the author's work has contributed to the understanding of atrophy and occult pathology associated with MS, and has highlighted some consistency issues associated with quantifying focal pathology. Within the BNAC, the techniques developed have been used in numerous studies in different disease subgroups (Benedict et al. 2007; Horakova et al. 2009; Zivadinov, Heininen-Brown, et al. 2012; Zivadinov, Dwyer, et al. 2012; Hagemeier et al. 2012; Bergsland et al. 2012; Zivadinov, Tekwe, et al. 2013; Hagemeier et al. 2013). Outside of the BNAC, they have complemented the ongoing work of others in the field in investigating the precise MRI-observable impact and course of MS, and the author's relevant publications have garnered over 250 citations at the time of this writing. In particular, it is now clear that atrophy is a complex phenomenon differentially affecting tissue classes, and that subtle atrophy occurs very early in the disease (Bergsland et al. 2012; Horakova et al. 2009). Furthermore, it is clear that atrophy occurs at different rates and with different impact in different areas and specific structures, and that these changes have an important impact on cognition (Bergsland et al. 2012; Benedict et al. 2006). With regard to occult pathology, diffusion entropy has been linked specifically to cognitive deficits (Benedict et al. 2007). Additionally, the VW-MTR technique has provided a better understanding of the locations and heterogeneity of subtle demyelination, and has also shown clear cortical changes (Dwyer et al. 2009; Zivadinov, Dwyer, et al. 2012). The author's phase imaging work has also led to particularly interesting findings, including the observations that iron likely relates to disability and that its deposition may precede atrophy (Hagemeier et al. 2012). Taken together, all of these findings seem to agree with the observations of others that the neurodegenerative component is not a removed, secondary

effect, but rather potentially a primary process (Stys et al. 2012). Finally, the author's work with focal pathology, although perhaps less scientifically exciting than the other work, has provided important information about precisely where 1.5 and 3 tesla scanners differ in lesion sensitivity, and has also indicated how to better train operators for agreement.

As a whole and alongside the complementary work of many other researchers, these various contributions provide guidance for future investigations in addition to their direct impact. With respect to atrophy, the author's work underscores the need for better understanding of the disparate mechanisms associated with this phenomenon, in order to separately quantify the contribution of fluid-based, axonal, and myelin-based tissue losses towards overall brain atrophy. For occult pathology, the work motivates further steps to elucidate the causes of non-focal pathology, and perhaps suggests the need for more appropriate animal models that include iron abnormalities and diffuse demyelination. Finally, for more conventional focal pathology, the work indicates that more needs to be done in terms of standardization, and future investigations might better investigate specific scan parameter-based impacts and the precise causes of operator disagreement in outlining lesions.

Outside of contributions to the knowledge base and indications of future directions, the work has also more directly resulted in the production of a number of useful algorithms that are both inherently valuable and can also be relatively easily adapted or applied to new problems. These algorithms include a semi-automated parametric shape model, a 4-D HMRF model for longitudinal tissue segmentation, a system for statistically robust quantification of subtle tissue changes (demyelination and remyelination), a technique for reliable quantification of structure-specific phase (representative of iron), and a decomposition method for quantifying specific error contributions to lesion identification. The techniques for voxel-wise MTR analysis and for improved GM classification will undoubtedly prove useful to monitor short and long term brain changes in a continuing variety of disease subclasses and therapeutic investigations. Equally importantly, the advances in understanding of lesion measurements can provide for more precise lesion burden measures in near-term clinical trials.

Finally, it seems even more clear that, as suggested previously (Zivadinov, Stosic, et al. 2008), a true understanding of MS is unlikely to come from any single modality or technique, but rather from the fusion of many into a single composite picture. This type of fusion is likely to require substantial statistical modelling work on large, standardized datasets containing information about many facets of the disease. Accurate, reliable quantitative techniques including those presented by the author are therefore vital precursors to the goal of completely overcoming the clinico-radiological paradox.

4.3 Conclusions

MS is a serious, debilitating disease that currently has no cure and whose precise etiology and evolution are still not fully understood. Despite this, it is clear that rapid progress is being made on a number of fronts, including a better understanding of the nature of the disease itself and a much more effective armamentarium of therapeutic options. Therefore, there is cause for some optimism that the future will continue to hold meaningful improvements for patients with MS.

The author's work described in this document has served a critical part in furthering this goal, both by directly answering important questions and by providing tools to answer future questions. Like the underlying disease, the related activities were multi-factorial, and dealt individually with atrophy, occult pathology, and focal pathology.

With reference to tissue atrophy, the author has demonstrated that it is possible to use quantitative techniques to precisely measure volumetric loss both in specific structures and in individual tissue compartments. In particular, optic nerve atrophy can be measured with a precision approaching one-tenth of a voxel, and overall GM atrophy can be assessed reliably with up to 50% fewer subjects than previously required.

With reference to occult pathologies, the author has demonstrated that higher-order statistical measures like entropy provide valuable additional information, and also that these pathologies can be meaningfully quantified even when their boundaries are relatively diffuse. Furthermore, the author has shown that the extent of demyelination and remyelination can be measured *in vivo*, and has

also provided a reliable means for reproducibly quantifying phase (iron) in the basal ganglia.

Finally, with reference to focal pathologies, the author has demonstrated that there are multiple, separable sources of error, and has provided techniques that can be employed to disentangle their relative impacts.

The work described here has taken over a decade of concerted effort, and has proceeded as a combination of top-down planning and bottom-up reaction to empirical findings. In total, it has resulted in the creation of eight new techniques and many related scientific findings. Taken together, these both directly pushed forward the current knowledge of MS and provided re-usable tools to answer future questions.

Reference list

Note: publications by the author are listed in bold

Ai, T., Morelli, J., Hu, X. and Hao, D. 2012. A Historical Overview of Magnetic Resonance Imaging, Focusing on Technological Innovations. *Investigative radiology* 47(12), pp. 725–41.

Allen, I. V, McQuaid, S., Mirakhur, M. and Nevin, G. 2001. Pathological abnormalities in the normal-appearing white matter in multiple sclerosis. *Neurological sciences* 22(2), pp. 141–4.

Al-Radaideh, A.M., Wharton, S.J., Lim, S., Tench, C.R., Morgan, P.S., Bowtell, R.W., Constantinescu, C.S. and Gowland, P. A. 2013. Increased iron accumulation occurs in the earliest stages of demyelinating disease: an ultra-high field susceptibility mapping study in Clinically Isolated Syndrome. *Multiple sclerosis (Houndmills, Basingstoke, England)* 19(7), pp. 896–903.

Andersson, J.L.R., Skare, S. and Ashburner, J. 2003. How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging. *NeuroImage* 20(2), pp. 870–88.

Ascherio, A. and Munger, K.L. 2007a. Environmental risk factors for multiple sclerosis. Part I: the role of infection. *Annals of neurology* 61(4), pp. 288–99.

Ascherio, A. and Munger, K.L. 2007b. Environmental risk factors for multiple sclerosis. Part II: Noninfectious factors. *Annals of neurology* 61(6), pp. 504–13.

Ashburner, J. and Friston, K. 2005. Unified segmentation. *Neuroimage* 26(3), pp. 839–851.

Bakshi, R., Shaikh, Z. and Janardhan, V. 2000. MRI T2 shortening ('black T2') in multiple sclerosis: frequency, location, and clinical correlation. *Neuroreport* 11(1), pp. 15–21.

Barkhof, F. 1999. MRI in multiple sclerosis: correlation with expanded disability status scale (EDSS). *Multiple sclerosis* 50(5), pp. 1282–1288.

Barkhof, F. 2002. The clinico-radiological paradox in multiple sclerosis revisited. *Current opinion in neurology* 15(3), pp. 239–245.

Benedict, R.H.B., Bruce, J., Dwyer, M.G., Weinstock-Guttman, B., Tjoa, C., Tavazzi, E., Munschauer, F.E. and Zivadinov, R. 2007. Diffusion-weighted imaging predicts cognitive impairment in multiple sclerosis. *Multiple sclerosis* 13(6), pp. 722–30.

Benedict, R.H.B.R., Bruce, J.M.J., Dwyer, M.G., Abdelrahman, N., Hussein, S., Weinstock-Guttman, B., Garg, N., Munschauer, F. and Zivadinov, R. 2006. Neocortical atrophy, third ventricular width, and cognitive dysfunction in multiple sclerosis. *Archives of neurology* 63(9), pp. 1301–6.

Bergsland, N., Horakova, D., Dwyer, M.G., Dolezal, O., Seidl, Z.K., Vaneckova, M., Krasensky, J., Havrdova, E. and Zivadinov, R. 2012. Subcortical and Cortical Gray Matter Atrophy in a Large Sample of Patients with Clinically Isolated Syndrome and Early Relapsing-Remitting Multiple Sclerosis. *American Journal of Neuroradiology\ 33(8), pp. 1573–1578.*

Berry, I., Barker, G.J., Barkhof, F., Campi, A., Dousset, V., Franconi, J.M., Gass, A., Schreiber, W., Miller, D.H. and Tofts, P.S. 1999. A multicenter measurement of magnetization transfer ratio in normal white matter. *Journal of magnetic resonance imaging: JMRI* 9(3), pp. 441–6.

Le Bihan, D. 2003. Looking into the functional architecture of the brain with diffusion MRI. *Nature reviews. Neuroscience* 4(6), pp. 469–80.

Black, M. and Sapiro, G. 1998. Robust anisotropic diffusion. *Image Processing, IEEE Transactions on* 7(3), pp. 421–432.

Bodini, B., Cercignani, M., Khaleeli, Z., Miller, D.H., Ron, M., Penny, S., Thompson, A.J. and Ciccarelli, O. 2013. Corpus callosum damage predicts disability progression and cognitive dysfunction in primary-progressive MS after five years. *Human brain mapping* 34(5), pp. 1163–72.

Bodini, B., Cercignani, M., Toosy, A., De Stefano, N., Miller, D.H., Thompson, A.J. and Ciccarelli, O. 2013. A novel approach with "skeletonised MTR" measures tract-specific microstructural changes in early primary-progressive MS. *Human brain mapping* 000(July 2012), pp. 1–11.

Bodini, B., Khaleeli, Z., Cercignani, M., Miller, D.H., Thompson, A.J. and Ciccarelli, O. 2009. Exploring the relationship between white matter and gray matter damage in early primary progressive multiple sclerosis: an in vivo study with TBSS and VBM. *Human brain mapping* 30(9), pp. 2852–61.

Bozzali, M., Spanò, B., Parker, G.J.M., Giulietti, G., Castelli, M., Basile, B., Rossi, S., Serra, L., Magnani, G., Nocentini, U., Caltagirone, C., Centonze, D. and Cercignani, M. 2013. Anatomical brain connectivity can assess cognitive dysfunction in multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England)* 19(9), pp. 1161–8.

Brown, R. A., Narayanan, S. and Arnold, D.L. 2012. Segmentation of magnetization transfer ratio lesions for longitudinal analysis of demyelination and remyelination in multiple sclerosis. *NeuroImage* 66C, pp. 103–109.

Bullmore, E.T. and Bassett, D.S. 2011. Brain graphs: graphical models of the human brain connectome. *Annual review of clinical psychology* 7, pp. 113–40.

Bunge, M.B., Bunge, R.P. and Ris, H. 1961. Ultrastructural study of remyelination in an experimental lesion in adult cat spinal cord. *The Journal of biophysical and biochemical cytology* 10, pp. 67–94.

Calabrese, M., Favaretto, A., Poretto, V., Romualdi, C., Rinaldi, F., Mattisi, I., Morra, A., Perini, P. and Gallo, P. 2013. Low degree of cortical pathology is associated with benign course of multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England)* 19(7), pp. 904–11.

Carone, D.A., Benedict, R.H.B., Dwyer, M.G., Cookfair, D.L., Srinivasaraghavan, B., Tjoa, C.W. and Zivadinov., R. 2006. Semi-automatic brain region extraction (SABRE) reveals superior cortical and deep gray matter atrophy in MS. *Neuroimage* 29(2), pp. 505–514.

Chard, D.T., Griffin, C.M., Parker, G.J.M., Kapoor, R., Thompson, A.J. and Miller., D.H. 2002. Brain atrophy in clinically early relapsing–remitting multiple sclerosis. *Brain* 125(2), pp. 327–337.

Chen, J.T., Kuhlmann, T., Jansen, G.H., Collins, D.L., Atkins, H.L., Freedman, M.S., O'Connor, P.W. and Arnold, D.L. 2007. Voxel-based analysis of the evolution of magnetization transfer ratio to quantify remyelination and

demyelination with histopathological validation in a multiple sclerosis lesion. *Neuroimage* 36(4), pp. 1152–8.

Ciccarelli, O., Catani, M., Johansen-Berg, H., Clark, C. and Thompson, A. 2008. Diffusion-based tractography in neurological disorders: concepts, applications, and future developments. *Lancet neurology* 7(8), pp. 715–27.

Ciccarelli, O., Giugni, E., Paolillo, A., Mainero, C., Gasperini, C., Bastianello, S. and Pozzilli, C. 1999. Magnetic resonance outcome of new enhancing lesions in patients with relapsing-remitting multiple sclerosis. *European journal of neurology* 6(4), pp. 455–9.

Ciccarelli, O., Werring, D.J., Barker, G.J., Griffin, C.M., Wheeler-Kingshott, C. a M., Miller, D.H. and Thompson, A.J. 2003. A study of the mechanisms of normal-appearing white matter damage in multiple sclerosis using diffusion tensor imaging--evidence of Wallerian degeneration. *Journal of neurology* 250(3), pp. 287–92.

Ciccarelli, O., Wheeler-Kingshott, C.A., McLean, M.A., Cercignani, M., Wimpey, K., Miller, D.H. and Thompson, A.J. 2007. Spinal cord spectroscopy and diffusion-based tractography to assess acute disability in multiple sclerosis. *Brain* 130(Pt 8), pp. 2220–31.

Coles, A.J., Compston, D.A., Selmaj, K., Lake, S.L., Moran, S., Margolin, D.H., Norris, K. and Tandon, P.K. 2008. Alemtuzumab vs. interferon beta-1a in early multiple sclerosis. *N Engl J Med* 359(17), pp. 1786–1801.

Connor, J. and Menzies, S. 1992. A histochemical study of iron, transferrin, and ferritin in Alzheimer's diseased brains. *Journal of neuroscience research* 31(1), pp. 75–83.

Craelius, W., Migdal, M.W., Luessenhop, C.P., Sugar, A. and Mihalakis., I. 1982. Iron deposits surrounding multiple sclerosis plaques. *Archives of pathology & laboratory medicine* 106(8), pp. 397–399.

Dade, L., Gao, F., Kovacevic, N. and Roy, P. 2004. Semiautomatic brain region extraction: a method of parcellating brain regions from structural magnetic resonance images. *Neuroimage* 22(4), pp. 1492–1502.

Dahnke, R., Yotter, R.A. and Gaser, C. 2013. Cortical thickness and central surface estimation. *NeuroImage* 65, pp. 336–48.

Deistung, A., Schäfer, A., Schweser, F., Biedermann, U., Turner, R. and Reichenbach, J.R. 2013. Toward in vivo histology: a comparison of quantitative susceptibility mapping (QSM) with magnitude-, phase-, and R2*-imaging at ultra-high magnetic field strength. *NeuroImage* 65, pp. 299–314.

Dexter, D., Carayon, A., Javoy-Agid, F. and Agid, Y. 1991. Alterations in the levels of iron, ferritin and other trace metals in Parkinson's disease and other neurodegenerative diseases affecting the basal ganglia. *Brain* 114(4), pp. 1953–1975.

Dhawan, A.P. 2011. *Medical Image Analysis*. John Wiley & Sons.

Dortch, R.D., Li, K., Gochberg, D.F., Welch, E.B., Dula, A.N., Tamhane, A.A., Gore, J.C. and Smith, S.A. 2011. Quantitative magnetization transfer imaging in human brain at 3 T via selective inversion recovery. *Magnetic resonance in medicine* 66(5), pp. 1346–52.

Dougherty, G. 2009. *Digital Image Processing for Medical Applications*. Cambridge University Press.

Drew, T., Vo, M.L. and Wolfe, J.M. 2013. The Invisible Gorilla Strikes Again: Sustained Inattentional Blindness in Expert Observers. *Psychological Science*, p. EPUB Ahead of print (DOI: 10.1177/0956797613479386.

Dwyer, M., Bergsland, N., Hussein, S., Durfee, J., Wack, D. and Zivadinov, R. 2009. A sensitive, noise-resistant method for identifying focal demyelination and remyelination in patients with multiple sclerosis via voxel-wise changes in magnetization transfer ratio. *Journal of the neurological sciences* 282(1-2), pp. 86–95.

Dwyer, M., Bergsland, N., Saluste, E., Sharma, J., Jaisani, Z., Durfee, J., Abdelrahman, N., Minagar, A., Hoque, R., Munschauer, F.E. and Zivadinov, R. 2008. Application of hidden Markov random field approach for quantification of perfusion/diffusion mismatch in acute ischemic stroke. *Neurological research* 30(8), pp. 827–34.

Dwyer, M.G., Bergsland, N. and Zivadinov, R. 2012. Improved longitudinal gray matter atrophy assessment via a combination of SIENA and a 4-dimensional hidden Markov random field model. In: *28th European Committee for Treatment and Research in Multiple Sclerosis*. Lyon, France.

Enzinger, C., Smith, S. and Fazekas, F. 2006. Lesion probability maps of white matter hyperintensities in elderly individuals. *Journal of neurology* 253(8), pp. 1064–1070.

Filippi, M., Campi, A., Dousset, V., Baratti, C., Martinelli, V., Canal, N., Scotti, G. and Comi, G. 1995. A magnetization transfer imaging study of normal-appearing white matter in multiple sclerosis. *Neurology* 45(3 Pt 1), pp. 478–82.

Filippi, M., Horsfield, M.A., Bressi, S., Martinelli, V., Baratti, C., Reganati, P., Campi, A., Miller, D.H. and Comi, G. 1995. Intra- and inter-observer agreement of brain MRI lesion volume measurements in multiple sclerosis. *Brain* 118(6), pp. 1593–1600.

Filippi, M., Rocca, M.A., Martino, G., Horsfield, M.A. and Comi., G. 1998. Magnetization transfer changes in the normal appering white matter precede the appearance of enhancing lesions in patients with multiple sclerosis. *Annals of neurology* 43(6), pp. 809–814.

Filippi, M., Rovaris, M., Inglese, M., Barkhof, F., De Stefano, N., Smith, S. and Comi, G. 2004. Interferon beta-1a for brain tissue loss in patients at presentation with syndromes suggestive of multiple sclerosis: a randomised, double-blind, placebo-controlled trial. *The Lancet* 364(9444), pp. 1489–1496.

Filippi, M., Rovaris, M., Rocca, M.A., Sormani, M.P., Wolinsky, J.S. and Comi, G. 2001. Glatiramer acetate reduces the proportion of new MS lesions evolving into " black holes ." *Neurology* 57(4), pp. 731–733.

Fischl, B. 2012. FreeSurfer. *Neuroimage* 62(2), pp. 774–781.

Fisniku, L.K., Brex, P.A., Altmann, D.R., Miszkiel, K.A., Benton, C.E., Lanyon, R., Thompson, A.J. and Miller, D.H. 2008. Disability and T2 MRI lesions: a 20-year follow-up of patients with relapse onset of multiple sclerosis. *Brain* 131(3), pp. 808–817.

Fisniku, L.K., Chard, D.T., Jackson, J.S., Anderson, V.M., Altmann, D.R., Miszkiel, K.A., Thompson, A.J. and Miller, D.H. 2008. Gray matter atrophy is related to long term disability in multiple sclerosis. *Annals of neurology* 64(3), pp. 247–254.

Fritz, D.A., Dwyer, M.G., Bagnato, F., Watts, K.L., Bratina, A., Zorzon, M., Durastanti, V., Locatelli, L., Millefiorini, E. and Zivadinov, R. 2006. Effect of MRI coregistration on serial short-term brain volume changes in multiple sclerosis. *Neurological research* **28(3), pp. 275–279.**

Frohman, E., Racke, M. and Raine, C. 2006. Multiple sclerosis—the plaque and its pathogenesis. *New England Journal of Medicine* 354(9), pp. 942–955.

Frohman, E.M., Dwyer, M.G., Frohman, T., Cox, J.L., Salter, A., Greenberg, B.M., Hussein, S., Conger, A., Calabresi, P., Balcer, L.J. and Zivadinov, R. 2009. Relationship of optic nerve and brain conventional and non-conventional MRI measures and retinal nerve fiber layer thickness, as assessed by OCT and GDx: a pilot study. *Journal of the neurological sciences* **282(1-2), pp. 96–105.**

García-Lorenzo, D., Francis, S., Narayanan, S., Arnold, D.L. and Collins, D.L. 2013. Review of automatic segmentation methods of multiple sclerosis white matter lesions on conventional magnetic resonance imaging. *Medical image analysis* 17(1), pp. 1–18.

Ge, Y., Grossman, R.I., Babb, J.S., Rabin, M.L., Mannon, L.J. and Kolson, D.L. 2002. Age-Related Total Gray Matter and White Matter Changes in Normal Adult Brain. Part I: Volumetric MR Imaging Analysis. *AJNR* 23(8), pp. 1327–1333.

Ge, Y., Grossman, R.I., Udupa, J.K., Babb, J.S., Nyul, L.G. and Kolson, D.L. 2001. Brain Atrophy in Relapsing-Remitting Multiple Sclerosis: Fractional Volumetric Analysis of Gray Matter and White Matter. *Radiology* 220(3), pp. 606–610.

Ge, Y., Grossman, R.I., Udupa, J.K., Wei, L., Mannon, L.J., Polansky, M. and Kolson, D.L. 2000. Brain Atrophy in Relapsing-Remitting Multiple Sclerosis and Secondary Progressive Multiple Sclerosis: Longitudinal Quantitative Analysis. *Radiology* 214(3), pp. 665–670.

Geurts, J.J.G. and Barkhof, F. 2008. Grey matter pathology in multiple sclerosis. *Lancet neurology* 7(9), pp. 841–51.

Geurts, J.J.G., Stys, P.K., Minagar, A., Amor, S. and Zivadinov, R. 2009. Gray matter pathology in (chronic) MS: modern views on an early observation. *Journal of the neurological sciences* 282(1-2), pp. 12–20.

Giacomini, P.S., Levesque, I.R., Ribeiro, L., Narayanan, S., Francis, S.J., Pike, G.B. and Arnold, D.L. 2009. Measuring demyelination and remyelination in acute multiple sclerosis lesion voxels. *Archives of neurology* 66(3), pp. 375–81.

Gonzalez, R.C. and Woods, R.E. 2008. *Digital Image Processing.* Prentice Hall.

Gourraud, P., Sdika, M., Khankhanian, P., Henry, R.G., Beheshtian, A., Matthews, P.M., Hauser, S.L., Oksenberg, J.R., Pelletier, D. and Baranzini, S.E. 2013. A genome-wide association study of brain lesion distribution in multiple sclerosis. *Brain : a journal of neurology* 136(Pt 4), pp. 1012–24.

Grimaud, J., Lai, M., Thorpe, J., Adeleine, P., Wang, L., Barker, G.J., Plummer, D.L., Tofts, P.S., McDonald, W.I. and Miller., D.H. 1996. Quantification of MRI

lesion load in multiple sclerosis: a comparison of three computer-assisted techniques. *Magnetic resonance imaging* 14(5), pp. 495–505.

Haacke, E.M. 1999. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley.

Haacke, E.M., Makki, M., Ge, Y., Maheshwari, M., Sehgal, V., Hu, J., Selvan, M., Wu, Z., Latif, Z., Xuan, Y., Khan, O., Garbern, J. and Grossman, R.I. 2009. Characterizing iron deposition in multiple sclerosis lesions using susceptibility weighted imaging. *Journal of magnetic resonance imaging : JMRI* 29(3), pp. 537–44.

Haacke, E.M., Xu, Y., Cheng, Y. and Reichenbach, J.R. 2004. Susceptibility weighted imaging (SWI). *Magnetic resonance in medicine* 52(3), pp. 612–8.

Hafler, D.A., Compston, A., Sawcer, S., Lander, E.S., Daly, M.J., De Jager, P.L. and De Bakker, P.I. 2007. Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study. *The New England journal of medicine* 357(9), p. 851.

Hagemeier, J., Weinstock-Guttman, B., Bergsland, N., Heininen-Brown, M., Carl, E., Kennedy, C., Magnano, C., Hojnacki, D., Dwyer, M.G. and Zivadinov, R. 2012. Iron deposition on SWI-filtered phase in the subcortical deep gray matter of patients with clinically isolated syndrome may precede structure-specific atrophy. *AJNR. American journal of neuroradiology* 33(8), pp. 1596–601.

Hagemeier, J., Weinstock-Guttman, B., Heininen-Brown, M., Poloni, G.U., Bergsland, N., Schirda, C., Magnano, C.R., Kennedy, C., Carl, E., Dwyer, M.G., Minagar, A. and Zivadinov, R. 2013. Gray matter SWI-filtered phase and atrophy are linked to disability in MS. *Frontiers in bioscience (Elite edition)* 5, pp. 525–32.

Hallgren, B. and Sourander, P. 1958. The effect of age on the non-haemin iron in the human brain. *J Neurochem* 3, pp. 41–51.

Hastie, T., Tibshirani, R. and Friedman, J. 2011. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition (Springer Series in Statistics)*. Springer.

Held, K., Rota Kops, E., Krause, B.J., Wells, W.M., Kikinis, R. and Müller-Gärtner, H.W. 1997. Markov random field segmentation of brain MR images. *IEEE transactions on medical imaging* 16(6), pp. 878–86.

Henkelman, R., Stanisz, G. and Graham, S. 2001. Magnetization transfer in MRI: a review. *NMR in Biomedicine* 14(2), pp. 57–64.

Henry, R.G., Shieh, M., Okuda, D.T., Evangelista, A., Gorno-Tempini, M.L. and Pelletier, D. 2008. Regional grey matter atrophy in clinically isolated syndromes at presentation. *Journal of Neurology, Neurosurgery & Psychiatry* 79(11), pp. 1236–1244.

Horakova, D., Dwyer, M.G., Havrdova, E., Cox, J.L., Dolezal, O., Bergsland, N., Rimes, B., Seidl, Z., Vaneckova, M. and Zivadinov, R. 2009. Gray matter atrophy and disability progression in patients with early relapsing-remitting multiple sclerosis: a 5-year longitudinal study. *Journal of the neurological sciences* 282(1-2), pp. 112–9.

Huang, D., Swanson, E.A., Lin, C.P., Schuman, J.S., Stinson, W.G., Chang, W., Hee, M.R., Flotte, T., Gregory, K. and Puliafito, C.A. 1991. Optical coherence tomography. *Science* 254(5035), pp. 1178–1181.

Huber, P.J. and Ronchetti, E.M. 2009. *Robust Statistics (Wiley Series in Probability and Statistics)*. Wiley.

Jackson, E.F., Narayana, P.A., Wolinsky, J.S. and Doyle, T.J. 1993. Accuracy and reproducibility in volumetric analysis of multiple sclerosis lesions. *Journal of computer assisted tomography* 17(2), pp. 200–5.

De Jager, P.L., Jia, X., Wang, J., de Bakker, P.I.W., Ottoboni, L., Aggarwal, N.T., Piccio, L., Raychaudhuri, S., Tran, D., Aubin, C., Briskin, R., Romano, S., Baranzini, S.E., McCauley, J.L., Pericak-Vance, M.A., Haines, J.L., Gibson, R.A., Naeglin, Y., Uitdehaag, B., Matthews, P.M., Kappos, L., Polman, C., McArdle, W.L., Strachan, D.P., Evans, D., Cross, A.H., Daly, M.J., Compston, A., Sawcer, S.J., Weiner, H.L., Hauser, S.L., Hafler, D.A. and Oksenberg, J.R. 2009. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nature genetics* 41(7), pp. 776–82.

Janve, V. a, Zu, Z., Yao, S., Li, K., Zhang, F.L., Wilson, K.J., Ou, X., Does, M.D., Subramaniam, S. and Gochberg, D.F. 2013. The radial diffusivity and magnetization transfer pool size ratio are sensitive markers for demyelination in a rat model of type III multiple sclerosis (MS) lesions. *NeuroImage* 74, pp. 298–305.

Jenkinson, M. and Smith, S. 2001. A global optimisation method for robust affine registration of brain images. *Medical image analysis* 5(2), pp. 143–156.

Jezzard, P., Matthews, P. and Smith, S. eds. 2002. *Functional MRI: An Introduction to Methods*. Oxford University Press, USA.

Kappos, L., Li, D., Calabresi, P.A., O'Connor, P., Bar-Or, A., Barkhof, F. and Yin, M. 2011. Ocrelizumab in relapsing-remitting multiple sclerosis: a phase 2, randomised, placebo-controlled, multicentre trial. *The Lancet* 378(9805), pp. 1779–1787.

Kappos, L., Radue, E., O'Connor, P., Polman, C., Hohlfeld, R., Calabresi, P. and Selmaj, K. 2010. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *New England Journal of Medicine* 362(5), pp. 387–401.

Kasper, L.H. and Shoemaker, J. 2010. Multiple sclerosis immunology: The healthy immune system vs the MS immune system. *Neurology* 74(1 Supplement 1), pp. S2–S8.

Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J. and Parsey, R. V 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46(3), pp. 786–802.

Kwon, O.I., Woo, E.J., Du, Y.P. and Hwang, D. 2013. A tissue-relaxation-dependent neighboring method for robust mapping of the myelin water fraction. *NeuroImage* 74, pp. 12–21.

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

Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A. and Suetens, P. 2001. Automated segmentation of multiple sclerosis lesions by model outlier detection. *Medical Imaging, IEEE Transactions on* 20(8), pp. 677–688.

LeVine, S. 1997. Iron deposits in multiple sclerosis and Alzheimer's disease brains. *Brain research* 760(1), pp. 298–303.

Losseff, N.A., Wang, L., Lai, H.M., Yoo, D.S., Gawne-Cain, M.L., McDonald, W.I., Miller, D.H. and Thompson, A.J. 1996. Progressive cerebral atrophy in multiple sclerosis A serial MRI study. *Brain* 119(6), pp. 2009–2019.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K. and Woods, R. 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356(1412), pp. 1293–1322.

McDonald, W.I. and Barnes, D. 1992. The ocular manifestations of multiple sclerosis. 1. Abnormalities of the afferent visual system. *Journal of neurology, neurosurgery, and psychiatry* 55(9), pp. 747–52.

Miller, D. and Weinshenker, B. 2008. Differential diagnosis of suspected multiple sclerosis: a consensus approach. *Multiple sclerosis* 14(9) pp. 1157-1174.

Miller, D.H., Barkhof, F., Frank, J.A., Parker, G.J. and Thompson, A.J. 2002. Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. *Brain* 125(8), pp. 1676–1695.

Mühlau, M., Buck, D., Förschler, A., Boucard, C.C., Arsic, M., Schmidt, P., Gaser, C., Berthele, A., Hoshi, M., Jochim, A., Kronsbein, H., Zimmer, C., Hemmer, B. and Ilg, R. 2013. White-matter lesions drive deep gray-matter atrophy in early multiple sclerosis: support from structural MRI. *Multiple sclerosis* 19(11), pp. 1485–92.

Nichols, T.E. and Holmes, A. 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Human brain mapping* 15(1), pp. 1–25.

Nolte, J. 2008. *The Human Brain: An Introduction to Its Functional Anatomy*. Mosby.

Papadopoulou, A., Müller-Lenke, N., Naegelin, Y., Kalt, G., Bendfeldt, K., Kuster, P., Stoecklin, M., Gass, A., Sprenger, T., Radue, E.W., Kappos, L. and Penner, I. 2013. Contribution of cortical and white matter lesions to cognitive impairment in multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England)* 19(10), pp. 1290–6.

Patenaude, B., Smith, S.M., Kennedy, D.N.D. and Jenkinson, M. 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 56(3), pp. 907–22.

Di Perri, C., Dwyer, M.G., Wack, D.S., Cox, J.L., Hashmi, K., Saluste, E., Hussein, S., Schirda, C., Stosic, M., Durfee, J., Poloni, G.U., Nayyar, N., Bergamaschi, R. and Zivadinov, R. 2009. Signal abnormalities on 1.5 and 3 Tesla brain MRI in multiple sclerosis patients and healthy controls. A morphological and spatial quantitative comparison study. *NeuroImage* 47(4), pp. 1352–62.

Pitman, E. 1937. Significance tests which may be applied to samples from any populations. *Supplement to the Journal of the Royal Statistical Society* 4(1), pp. 119–130.

Polman, C. and Reingold, S. 2011. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Annals of Neurology*.

Polman, C.H., O'Connor, P.W., Havrdova, E., Hutchinson, M., Kappos, L., Miller, D.H. and Phillips, J.T. 2006. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *New England Journal of Medicine* 354(9), pp. 899–910.

Poser, S., Raun, N. and Poser, W. 1982. Age at onset, initial symptomatology and the course of multiple sclerosis. *Acta Neurologica Scandinavica* 66(3), pp. 355–362.

Prineas, J.W., Barnard, R.O., Kwon, E.E., Sharer, L.R. and Cho, E.S. 1993. Multiple sclerosis: remyelination of nascent lesions. *Annals of neurology* 33(2), pp. 137–51.

Prineas, J.W. and Connell, F. 1979. Remyelination in multiple sclerosis. *Annals of neurology* 5(1), pp. 22–31.

Riccitelli, G., Rocca, M. A., Pagani, E., Rodegher, M.E., Rossi, P., Falini, A., Comi, G. and Filippi, M. 2011. Cognitive impairment in multiple sclerosis is associated to different patterns of gray matter atrophy according to clinical phenotype. *Human brain mapping* 32(10), pp. 1535–43.

Røhl, L., Østergaard, L., Simonsen, C.Z., Vestergaard-Poulsen, P., Andersen, G., Sakoh, M., Bihan, D. Le and Gyldensted, C. 2001. Viability thresholds of ischemic penumbra of hyperacute stroke defined by perfusion-weighted MRI and apparent diffusion coefficient. *Stroke* 32(5), pp. 1140–1146.

Rosati, G. 2001. The prevalence of multiple sclerosis in the world: an update. *Neurological sciences* 22(2), pp. 117–139.

Rudick, R. A. and Cutter, G. 2013. MRI lesions: a surrogate for relapses in multiple sclerosis? *Lancet neurology* 12(7), pp. 628–30.

Rudick, R.A., Fisher, E., Filippi, M. and Rocca, M. 2013. Preventing brain atrophy should be the gold standard of effective therapy in MS (after the first year of treatment): Yes. *Multiple Sclerosis Journal* 19(8), pp. 1003–1004.

Rudick, R.A., Fisher, E., Lee, J., Duda, J.T. and Simon, J. 2000. Brain atrophy in relapsing multiple sclerosis: relationship to relapses, EDSS, and treatment with interferon β-1a. *Multiple sclerosis* 6(6), pp. 365–372.

Sanfilipo, M.P., Benedict, R.H.B., Weinstock-Guttman, B. and Bakshi, R. 2006. Gray and white matter brain atrophy and neuropsychological impairment in multiple sclerosis. *Neurology* 66(5), pp. 685–92.

Sicotte, N.L., Voskuhl, R.R., Bouvier, S., Klutch, R., Cohen, M.S. and Mazziotta, J.C. 2003. Comparison of multiple sclerosis lesions at 1.5 and 3.0 Tesla. *Investigative radiology* 38(7), pp. 423–7.

Simon, J.H.J., Jacobs, L.D., Campion, M.K., Rudick, R.A., Cookfair, D.L., Herndon, R.M., Richert, J.R., Salazar, A.M., Fischer, J.S., Goodkin, D.E., Simonian, N., Lajaunie, M., Miller, D.E., Wende, K., Martens-Davidson, A., Kinkel, R.P., Munschauer, F.E. and Brownscheidle, C.M. 1999. A longitudinal

study of brain atrophy in relapsing multiple sclerosis. *Neurology* 53(1), pp. 139–139.

Sled, J.G., Zijdenbos, A.P. and Evans, A.C. 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE transactions on medical imaging* 17(1), pp. 87–97.

Smith, S. and Nichols, T. 2009. Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44(1), pp. 83–98.

Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, A. and De Stefano, N. 2002. Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis. *NeuroImage* 17(1), pp. 479–489.

Sørensen, T.L., Frederiksen, J.L., Brønnum-Hansen, H. and Petersen, H.C. 1999. Optic neuritis as onset manifestation of multiple sclerosis: A nationwide, long-term survey. *Neurology* 53(3), pp. 473–473.

Sormani, M.P. and Bruzzi, P. 2013. MRI lesions as a surrogate for relapses in multiple sclerosis: a meta-analysis of randomised trials. *Lancet neurology* 12(7), pp. 669–76.

Spader, H.S., Ellermeier, A., O'Muircheartaigh, J., Dean, D.C., Dirks, H., Boxerman, J.L., Cosgrove, G.R. and Deoni, S.C.L. 2013. Advances in myelin imaging with potential clinical application to pediatric imaging. *Neurosurgical focus* 34(4), p. E9.

Srinivasan, R., Sailasuta, N., Hurd, R., Nelson, S. and Pelletier, D. 2005. Evidence of elevated glutamate in multiple sclerosis using magnetic resonance spectroscopy at 3 T. *Brain : a journal of neurology* 128(Pt 5), pp. 1016–25.

De Stefano, N., Filippi, M., Miller, D., Pouwels, P.J., Rovira, A., Gass, A., Enzinger, C., Matthews, P.M. and Arnold, D.L. 2007. Guidelines for using proton MR spectroscopy in multicenter clinical MS studies. *Neurology* 69(20), pp. 1942–52.

Stikov, N., Perry, L.M., Mezer, A., Rykhlevskaia, E., Wandell, B.A., Pauly, J.M. and Dougherty, R.F. 2011. Bound pool fractions complement diffusion measures to describe white matter micro and macrostructure. *NeuroImage* 54(2), pp. 1112–21.

Stys, P.K., Zamponi, G.W., Minnen, J. van and Geurts, J.J. 2012. Will the real multiple sclerosis please stand up? *Nature Reviews Neuroscience* 13(7), pp. 507–514.

Takasawa, M., Jones, P.S., Guadagno, J. V., Christensen, S., Fryer, T.D., Harding, S. and Gillard, J.H. 2008. How reliable is perfusion MR in acute stroke? Validation and determination of the penumbra threshold against quantitative PET. *Stroke* 39(3), pp. 870–877.

Tavazzi, E., Dwyer, M.G., Weinstock-Guttman, B., Lema, J., Bastianello, S., Bergamaschi, R., Cosi, V., Benedict, R.H., Munschauer, F.E. and Zivadinov, R. 2007. Quantitative diffusion weighted imaging measures in patients with multiple sclerosis. *Neuroimage* 36(3), pp. 746–754.

Trapp, B. and Nave, K. 2008. Multiple sclerosis: an immune or neurodegenerative disorder? *Annu. Rev. Neurosci.* 31, pp. 247–269.

Vrenken, H., Vos, E.K., van der Flier, W.M., Sluimer, I.C., Cover, K.S., Knol, D.L. and Barkhof, F. 2013. Validation of the automated method VIENA: An accurate, precise, and robust measure of ventricular enlargement. *Human brain mapping* (November 2012).

Wack, D.S., Dwyer, M.G., Bergsland, N., Di Perri, C., Ranza, L., Hussein, S., Ramasamy, D., Poloni, G. and Zivadinov, R. 2012. Improved assessment of multiple sclerosis lesion segmentation agreement via detection and outline error estimates. *BMC medical imaging* 12, p. 17.

Van Waesberghe, J.H., van Walderveen, M.A., Castelijns, J.A., Scheltens, P., Lycklama à Nijeholt, G.J., Polman, C.H. and Barkhof, F. 1998. Patterns of lesion development in multiple sclerosis: longitudinal observations with T1-weighted spin-echo and magnetization transfer MR. *American journal of neuroradiology* 19(4), pp. 675–83.

Van Walderveen, M.A.A., Kamphorst, W., Scheltens, P., van Waesberghe, J.H.T.M., Ravid, R., Valk, J., Polman, C.H. and Barkhof, F. 1998. Histopathologic correlate of hypointense lesions on T1-weighted spin-echo MRI in multiple sclerosis. *Neurology* 50(5), pp. 1282–1288.

Willer, C.J., Dyment, D.A., Risch, N.J., Sadovnick, A.D. and Ebers, G.C. 2003. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proceedings of the National Academy of Sciences of the United States of America* 100(22), pp. 12877–82.

Winkler, G. 2003. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction*. Springer.

Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J. and Evans, A.C. 1996. A unified statistical approach for determining significant signals in images of cerebral activation. *Human brain mapping* 4(1), pp. 58–73.

Yablonskiy, D.A., Luo, J., Iyer, A., Sukstanskii, A.L. and Cross, A.H. 2012. On the Nature of Phase Contrast in Multiple Sclerosis Lesions. In: *Proc Intl Soc Mag Reson Med 20 (2012)*. Melbourne, Australia, p. 502.

Yuan, K. and Bentler, P. 1998. Robust mean and covariance structure analysis. *British Journal of Mathematical and Statistical Psychology* 51(1), pp. 63–88.

Zhang, Y., Brady, M. and Smith, S. 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging* 20(1), pp. 45–57.

Zheng, W., Nichol, H., Liu, S., Cheng, Y.N. and Haacke, E.M. 2013. Measuring iron in the brain using quantitative susceptibility mapping and X-ray fluorescence imaging. *NeuroImage* 78, pp. 68–74.

Zijdenbos, A.P., Forghani, R. and Evans, A.C. 2002. Automatic "pipeline" analysis of 3-D MRI data for clinical trials: application to multiple sclerosis. *IEEE transactions on medical imaging* 21(10), pp. 1280–91.

Zivadinov, R., Banas, A.C., Yella, V., Abdelrahman, N., Weinstock-Guttman, B. and Dwyer, M.G. 2008. Comparison of three different methods for measurement of cervical cord atrophy in multiple sclerosis. *American journal of neuroradiology* 29(2), pp. 319–25.

Zivadinov, R., Bergsland, N., Dolezal, O., Hussein, S., Seidl, Z., Dwyer, M.G., Vaneckova, M., Krasensky, J., Potts, J.A., Kalincik, T., Havrdová, E. and Horáková, D. 2013. Evolution of Cortical and Thalamus Atrophy and

Disability Progression in Early Relapsing-Remitting MS during 5 Years. *American journal of neuroradiology*, p. EPUB (DOI: 10.3174/ajnr.A3503).

Zivadinov, R., Dwyer, M.G., Hussein, S., Carl, E., Kennedy, C., Andrews, M., Hojnacki, D., Heininen-Brown, M., Willis, L., Cherneva, M., Bergsland, N. and Weinstock-Guttman, B. 2012. Voxel-wise magnetization transfer imaging study of effects of natalizumab and IFNβ-1a in multiple sclerosis. *Multiple sclerosis* 18(8), pp. 1125–34.

Zivadinov, R., Heininen-Brown, M., Schirda, C. V, Poloni, G.U., Bergsland, N., Magnano, C.R., Durfee, J., Kennedy, C., Carl, E., Hagemeier, J., Benedict, R.H.B., Weinstock-Guttman, B. and Dwyer, M.G. 2012. Abnormal subcortical deep-gray matter susceptibility-weighted imaging filtered phase measurements in patients with multiple sclerosis: a case-control study. *NeuroImage* 59(1), pp. 331–9.

Zivadinov, R., Hussein, S., Stosic, M., Durfee, J., Cox, J.L., Cookfair, D.L., Hashmi, K., Abdelrahman, N., Garg, N., Dwyer, M.G. and Weinstock-Guttman, B. 2011. Glatiramer acetate recovers microscopic tissue damage in patients with multiple sclerosis. A case-control diffusion imaging study. *Pathophysiology* 18(1), pp. 61–8.

Zivadinov, R., Reder, a T., Filippi, M., Minagar, a, Stüve, O., Lassmann, H., Racke, M.K., Dwyer, M.G., Frohman, E.M. and Khan, O. 2008. Mechanisms of action of disease-modifying agents and brain volume changes in multiple sclerosis. *Neurology* 71(2), pp. 136–44.

Zivadinov, R., Sepcic, J., Nasuelli, D., Masi, R. De, Bragadin, L.M., Tommasi, M.A. and Zambito-Marsala, S. 2001. A longitudinal study of brain atrophy and cognitive disturbances in the early phase of relapsing-remitting multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* 70(6), pp. 773–780.

Zivadinov, R., Stosic, M., Cox, J.L., Ramasamy, D.P. and Dwyer, M.G. 2008. The place of conventional MRI and newly emerging MRI techniques in monitoring different aspects of treatment outcome. *Journal of neurology* 255 Suppl , pp. 61–74.

Zivadinov, R., Tekwe, C., Bergsland, N., Dolezal, O., Havrdova, E., Krasensky, J., Dwyer, M.G., Seidl, Z., Ramasamy, D.P., Vaneckova, M. and Horakova, D. 2013. Bimonthly Evolution of Cortical Atrophy in Early Relapsing-Remitting Multiple Sclerosis over 2 Years: A Longitudinal Study. *Multiple sclerosis international* 2013, p. 231345.

Appendix A – Statement of independence

This is an age of "Big Science" in which work is multi-disciplinary and highly collaborative. Nowhere is this truer than in the modern approach to understanding neurological disease, which requires the concerted effort of neurologists, radiologists, pathologists, immunologists, epidemiologists, biostatisticians, MR physicists, and computer scientists. As such, virtually none of the work discussed here was performed in a vacuum, as is reflected by the sometimes extensive author listings.

This disclaimer notwithstanding, all of the primary publications were selected because they involved a substantial, independent, and necessary individual contribution from the author. To clarify this, a more specific summary is given below for each publication. Numbering is consistent with that given previously in Table 1. First, a brief description of the division of labour is given. Then, specific independent contributions that can be *solely* attributed to the author are listed. Note that these are not necessarily the full extent of the author's involvement, but rather represent those areas for which responsibility was not shared and which would not have been possible without him.

A.1 Specific contributions by publication

1. Semi-automatic brain region extraction (SABRE) reveals superior cortical and deep gray matter atrophy in MS. Carone DA, Benedict RH, Dwyer MG, Cookfair DL, Srinivasaraghavan B, Tjoa CW, Zivadinov R. *Neuroimage. 2006 Jan 15;29(2):505-14. Epub 2005 Oct 5.*

Description: This work involved the parcellation of specific brain structures in MS patients based on standard Talairach landmarks. In particular, it required the adaptation of the SABRE algorithm. The author was the sole computer scientist involved in the process, and was independently responsible for algorithm implementation and modification. The other major contribution to the work was by DA Carone, a neuropsychologist who performed all manual landmark identification. RH Benedict and R Zivadinov, a neuropsychologist and a clinician, respectively, provided project guidance.

Independent contribution(s): algorithm implementation and adjustment

2. Effect of MRI coregistration on serial short-term brain volume changes in multiple sclerosis. Fritz DA, Dwyer MG, Bagnato F, Watts KL, Bratina A, Zorzon M, Durastanti V, Locatelli L, Millefiorini E, Zivadinov R. *Neurol Res. 2006 Apr;28(3):275-9.*

Description: In this work, the author functioned as senior computer scientist and both supervised and worked jointly with DA Fritz, a junior computer scientist, in implementing the actual analysis. The author provided direct oversight and guidance, including project planning, and troubleshooting. The manuscript was written by DA Fritz with critical review by the author. R Zivadinov provided high-level project guidance.

Independent contribution(s): None *(nearly all meaningful work was joint, so no contribution was completely independent)*

3. Neocortical atrophy, third ventricular width, and cognitive dysfunction in multiple sclerosis. Benedict RH, Bruce JM, Dwyer MG, Abdelrahman N, Hussein S, Weinstock-Guttman B, Garg N, Munschauer F, Zivadinov R. *Arch Neurol. 2006 Sep;63(9):1301-6.*

Description: This work involved the calculation of specific atrophy measures and their correlation with cognition. The author was the sole computer scientist involved, and was independently responsible for determining and implementing the proper algorithms to make these measurements. Individual case analyses we performed by N Abdelrahman and S Hussein, and statistical analyses were performed by RH Benedict and JM Bruce.

Independent contribution(s): algorithm selection and implementation; writing or relevant methods section(s).

4. Quantitative diffusion weighted imaging measures in patients with multiple sclerosis. Tavazzi E, Dwyer MG, Weinstock-Guttman B, Lema J, Bastianello S, Bergamaschi R, Cosi V, Benedict RH, Munschauer FE 3rd, Zivadinov R. *Neuroimage. 2007 Jul 1;36(3):746-54. Epub 2007 Apr 10.*

Description: In this publication, the author was directly and independently responsible for all technical aspects of the study. This included research, design, and implementation of the analysis pipeline. In addition, the idea to use entropy and its execution within the study were solely the author's. E Tavazzi applied the created pipeline to specific clinical cases. Writing responsibility was shared between the authors.

Independent contribution(s): algorithm implementation and adjustment; proposal to use entropy as a new measure; writing of relevant methods section

5. Application of hidden Markov random field approach for quantification of perfusion/diffusion mismatch in acute ischemic stroke. Dwyer MG, Bergsland N, Saluste E, Sharma J, Jaisani Z, Durfee J, Abdelrahman N, Minagar A, Hoque R, Munschauer FE 3rd, Zivadinov R. *Neurol Res. 2008 Oct;30(8):827-34. doi: 10.1179/174313208X340987.*

Description: The author was individually responsible for the entire planning and organization of this study. In addition, he was the sole designer and implementer of the actual computational algorithm used, and the methodological conceptualization was entirely due to him, as was the actual writing of the paper. Other authors were responsible for assisting with MRI acquisition (E Saluste), pathology delineation (J Sharma, Z Jaisani, J Durfee, N Abdelrahman), or critical review and guidance.

Independent contribution(s): overall study design; algorithm conceptualization and implementation; writing

6. A sensitive, noise-resistant method for identifying focal demyelination and remyelination in patients with multiple sclerosis via voxel-wise changes in magnetization transfer ratio. Dwyer MG, Bergsland N, Hussein S, Durfee JE, Wack DS, Zivadinov R. *J Neurol Sci. 2009 Jul 15;282(1-2):86-95. doi: 10.1016/j.jns.2009.03.016. Epub 2009 Apr 22.*

Description: The author was individually responsible for the entire planning and organization of this study. In addition, he was the sole designer and implementer of the actual computational algorithm used, and the methodological conceptualization was entirely his.

Independent contribution(s): overall study design; algorithm conceptualization and implementation; simulation design and implementation; writing

7. Relationship of optic nerve and brain conventional and non-conventional MRI measures and retinal nerve fiber layer thickness, as assessed by OCT and GDx: a pilot study. Frohman EM, Dwyer MG, Frohman T, Cox JL, Salter A, Greenberg BM, Hussein S, Conger A, Calabresi P, Balcer LJ, Zivadinov R. *J Neurol Sci. 2009 Jul 15;282(1-2):96-105. doi: 10.1016/j.jns.2009.04.010. Epub 2009 May 12.*

Description: In this publication, the author was independently responsible for the optic nerve radius measurement technique conceptualization, execution, and documentation. The ratings are given as "moderate" because other authors were responsible for OCT measurements and lesion analysis, and overall writing responsibility was shared.

Independent contribution(s): optic nerve algorithm conceptualization and implementation; optic nerve atrophy measurement; optic nerve atrophy validation design and implementation; writing optic nerve measurement section of paper

8. Signal abnormalities on 1.5 and 3 Tesla brain MRI in multiple sclerosis patients and healthy controls. A morphological and spatial quantitative comparison study. Di Perri C, Dwyer MG, Wack DS, Cox JL, Hashmi K, Saluste E, Hussein S, Schirda C, Stosic M, Durfee J, Poloni GU, Nayyar N,

Bergamaschi R, Zivadinov R. *Neuroimage. 2009 Oct 1;47(4):1352-62. doi: 10.1016/j.neuroimage.2009.04.019. Epub 2009 Apr 14.*

Description: In this publication, the author was responsible for the conceptualization and implementation of the spatial lesion probability mapping approach to compare scanners, as well as the technical aspects of preparing the analysis in a halfway space. Delineation of lesions and non-spatial statistical analysis were performed by other authors.

Independent contribution(s): probability map portion of study design; implementation of probability mapping and adjustment to paired tests; writing of relevant methods section

9. Abnormal subcortical deep-gray matter susceptibility-weighted imaging filtered phase measurements in patients with multiple sclerosis: a case-control study. Zivadinov R, Heininen-Brown M, Schirda CV, Poloni GU, Bergsland N, Magnano CR, Durfee J, Kennedy C, Carl E, Hagemeier J, Benedict RH, Weinstock-Guttman B, Dwyer MG. *Neuroimage. 2012 Jan 2;59(1):331-9. doi: 10.1016/j.neuroimage.2011.07.045. Epub 2011 Jul 27.*

Description: In this publication, the author was responsible for designing the overall pipeline for data processing, including combining FIRST and unwarping steps. In addition, he performed the analysis for the normative case sets and introduced the MPV-LP measure.

Independent contribution(s): probability map portion of study design; implementation of probability mapping and adjustment to paired tests; writing of relevant methods section

10. Improved assessment of multiple sclerosis lesion segmentation agreement via detection and outline error estimates. Wack DS, Dwyer MG, Bergsland N, Di Perri C, Ranza L, Hussein S, Ramasamy D, Poloni G, Zivadinov R. *BMC Med Imaging. 2012 Jul 19;12:17. doi: 10.1186/1471-2342-12-17.*

Description: The author was responsible for joint conceptualization and oversight of this project, as well as for selection of the dataset used and the overall analysis pipeline including unbiased alignment into a halfway space.

Independent contribution(s): None *(nearly all meaningful work was joint, so no contribution was completely independent)*

11. Improved longitudinal gray matter atrophy assessment via a combination of SIENA and a 4-dimensional hidden Markov random field model. Dwyer MG, Bergsland N, Zivadinov R. *28th European Committee for Treatment and Research in Multiple Sclerosis, Lyon, France, October 10-13, 2012:P840.*

Description: In this work, the author both conceived and implemented the mathematical framework and algorithm for the 4-dimension HMRF model. In addition, he performed all the simulation and scan-rescan work. The addition of other SIENA elements (halfway space and combined brain mask) was conceptualized jointly with N Bergsland. Analysis on clinical datasets was shared between authors.

Independent contribution(s): Study planning; conceptualization and implementation of novel 4D HMRF algorithm; creation of abstract and poster; design and implementation of validation simulation

A.2 Co-author confirmation

To substantiate these claims, a supporting document from Professor Robert Zivadinov, BNAC Director, is provided below. As the senior researcher for the centre, he was usually most responsible for oversight of proper authorship apportionment for these papers. In addition, he was either last or first author for many of them.

Supporting statements from additional authors can be provided upon request.

BNAC

Buffalo Neuroimaging Analysis Center
100 High Street ❖ Buffalo, NY 14203
TEL: 716.859.7040 ❖ FAX: 716.859.7066
www.BNAC.net

Robert Zivadinov, MD PhD FAAN
Buffalo Neuroimaging Analysis Center
Department of Neurology
University at Buffalo
100 High St.
Buffalo, NY 14203

August 27, 2013

School of Engineering, Design, and Technology
University of Bradford
Richmond Road
Bradford, West Yorkshire, BD7 1DP, UK

To Whom It May Concern:

This letter serves as acknowledgement that I have read and agree with "Appendix A – Statement of independence" in Michael Dwyer's submission for consideration of a Ph.D. entitled "Development and application of novel algorithms for quantitative analysis of magnetic resonance imaging in multiple sclerosis".

As a co-author, I confirm that the overall descriptions and specific contributions listed are true to the best of my knowledge. In particular, I can attest to the veracity of the claims for the following specific works listed in Table 1:

Ref. No.	Publication	Status[*]
1	**Semi-automatic brain region extraction (SABRE) reveals superior cortical and deep gray matter atrophy in MS.** Carone DA, Benedict RH, Dwyer MG, Cookfair DL, Srinivasaraghavan B, Tjoa CW, Zivadinov R. Neuroimage. 2006 Jan 15;29(2):505-14. Epub 2005 Oct 5.	RP***
2	**Effect of MRI coregistration on serial short-term brain volume changes in multiple sclerosis.** Fritz DA, Dwyer MG, Bagnato F, Watts KL, Bratina A, Zorzon M, Durastanti V, Locatelli L, Millefiorini E, Zivadinov R. Neurol Res. 2006 Apr;28(3):275-9.	RP***
3	**Neocortical atrophy, third ventricular width, and cognitive dysfunction in multiple sclerosis.** Benedict RH, Bruce JM, Dwyer MG, Abdelrahman N, Hussein S, Weinstock-Guttman B, Garg N, Munschauer F, Zivadinov R. Arch Neurol. 2006 Sep;63(9):1301-6.	RP***
4	**Quantitative diffusion weighted imaging measures in patients with multiple sclerosis.** Tavazzi E, Dwyer MG, Weinstock-Guttman B, Lema J, Bastianello S, Bergamaschi R, Cosi V, Benedict RH, Munschauer FE 3rd, Zivadinov R. Neuroimage. 2007 Jul 1;36(3):746-54. Epub 2007 Apr 10.	RP***

JACOBS
NEUROLOGICAL
INSTITUTE

UB
University at Buffalo
The State University of New York

5	**Application of hidden Markov random field approach for quantification of perfusion/diffusion mismatch in acute ischemic stroke.** Dwyer MG, Bergsland N, Saluste E, Sharma J, Jaisani Z, Durfee J, Abdelrahman N, Minagar A, Hoque R, Munschauer FE 3rd, Zivadinov R. Neurol Res. 2008 Oct;30(8):827-34. doi: 10.1179/174313208X340987.	RP**
6	**A sensitive, noise-resistant method for identifying focal demyelination and remyelination in patients with multiple sclerosis via voxel-wise changes in magnetization transfer ratio.** Dwyer MG, Bergsland N, Hussein S, Durfee JE, Wack DS, Zivadinov R. J Neurol Sci. 2009 Jul 15;282(1-2):86-95. doi: 10.1016/j.jns.2009.03.016. Epub 2009 Apr 22.	RP**
7	**Relationship of optic nerve and brain conventional and non-conventional MRI measures and retinal nerve fiber layer thickness, as assessed by OCT and GDx: a pilot study.** Frohman EM, Dwyer MG, Frohman T, Cox JL, Salter A, Greenberg BM, Hussein S, Conger A, Calabresi P, Balcer LJ, Zivadinov R. J Neurol Sci. 2009 Jul 15;282(1-2):96-105. doi: 10.1016/j.jns.2009.04.010. Epub 2009 May 12.	RP***
8	**Signal abnormalities on 1.5 and 3 Tesla brain MRI in multiple sclerosis patients and healthy controls. A morphological and spatial quantitative comparison study.** Di Perri C, Dwyer MG, Wack DS, Cox JL, Hashmi K, Saluste E, Hussein S, Schirda C, Stosic M, Durfee J, Poloni GU, Nayyar N, Bergamaschi R, Zivadinov R. Neuroimage. 2009 Oct 1;47(4):1352-62. doi: 10.1016/j.neuroimage.2009.04.019. Epub 2009 Apr 14.	RP***
9	**Abnormal subcortical deep-gray matter susceptibility-weighted imaging filtered phase measurements in patients with multiple sclerosis: a case-control study.** Zivadinov R, Heininen-Brown M, Schirda CV, Poloni GU, Bergsland N, Magnano CR, Durfee J, Kennedy C, Carl E, Hagemeier J, Benedict RH, Weinstock-Guttman B, Dwyer MG. Neuroimage. 2012 Jan 2;59(1):331-9. doi: 10.1016/j.neuroimage.2011.07.045. Epub 2011 Jul 27.	RP***
10	**Improved assessment of multiple sclerosis lesion segmentation agreement via detection and outline error estimates.** Wack DS, Dwyer MG, Bergsland N, Di Perri C, Ranza L, Hussein S, Ramasamy D, Poloni G, Zivadinov R. BMC Med Imaging. 2012 Jul 19;12:17. doi: 10.1186/1471-2342-12-17.	RP***
11	**Improved longitudinal gray matter atrophy assessment via a combination of SIENA and a 4-dimensional hidden Markov random field model.** Dwyer MG, Bergsland N, Zivadinov R. 28th European Committee for Treatment and Research in Multiple Sclerosis, Lyon, France, October 10-13, 2012:P840.	N-P**
12	**Improved longitudinal gray and white matter atrophy assessment via application of a 4-dimensional hidden Markov random field model.** Dwyer MG, Bergsland NP, Zivadinov R. Neuroimage [in revision][*] [*]This work has been submitted to Neuroimage, and is currently in the second round of revisions after positive peer review. The manuscript is included in Appendix B.	

Please do not hesitate to contact me for further confirmation or additional clarification.

Sincerely yours,

Robert Zivadinov, MD PhD FAAN
Director, BNAC

63

Appendix B – Published works for consideration

This appendix collects the documents listed in Table 1. For copyright reasons, only the first page of each document, including abstract, is shown here. The full versions of each work can be retrieved from the appropriate academic journals.

Semi-automatic brain region extraction (SABRE) reveals superior cortical and deep gray matter atrophy in MS

D.A. Carone,[a,b,c] R.H.B. Benedict,[a,b,c,*] M.G. Dwyer,[a] D.L. Cookfair,[c] B. Srinivasaraghavan,[a] C.W. Tjoa,[a] and R. Zivadinov[a,b,c]

[a]Buffalo Neuroimaging Analysis Center, Buffalo, NY 14203, USA
[b]The Jacobs Neurological Institute, Buffalo, NY 14203, USA
[c]Department of Neurology, State University of New York (SUNY) at Buffalo, School of Medicine and Biomedical Sciences, Buffalo, NY 14203, USA

In multiple sclerosis (MS), atrophy occurs in various cortical and subcortical regions. However, it is unclear whether this is mostly due to gray (GM) or white matter (WM) loss. Recently, a new semi-automatic brain region extraction (SABRE) technique was developed to quantify parenchyma volume in 13 hemispheric regions. This study utilized SABRE and tissue segmentation to examine whether regional brain atrophy in MS is mostly due to GM or WM loss, correlated with disease duration, and moderated by disease course. We studied 68 MS patients and 39 normal controls with 1.5 T brain MRI. As expected, MS diagnosis was associated with significantly lower ($P < 0.001$) regional brain parenchymal fractions (RBPFs). While significant findings emerged in 11 GM comparisons, only four WM comparisons were significant. The largest mean RBPF percent differences between groups (MS < NC) were in the posterior basal ganglia/thalamus region (-19.3%), superior frontal (-15.7%), and superior parietal (-14.3%) regions. Logistic regression analyses showed GM regions were more predictive of MS diagnosis than WM regions. Eight GM RBPFs were significantly correlated ($P < 0.001$) with disease duration compared to only one WM region. Significant trends emerged for differences in GM, but not WM between secondary progressive (SP) and relapsing–remitting MS patients. Percent differences in GM between the two groups were largest in superior frontal (-9.9%), medial superior frontal (-6.5%), and superior parietal (-6.1%) regions, with SP patients having lower volumes. Overall, atrophy in MS is diffuse and mostly related to GM loss particularly in deep GM and superior frontal–parietal regions.
© 2005 Elsevier Inc. All rights reserved.

Keywords: Multiple sclerosis; Regional brain atrophy; Quantified magnetic resonance imaging; Gray matter

Introduction

Multiple sclerosis (MS) is a demyelinating and degenerative disease of the central nervous system (CNS) characterized by lesion formation and atrophy of the brain and spinal cord. Brain lesions were described in a series of apparent MS cases by Cruveilhier (1835), and atrophy was later documented by Charcot (1877). Subsequent autopsy studies revealed ventriculomegaly in a significant number of MS patients (Barnard and Triggs, 1974; Brownell and Hughes, 1962; Friedman and Davidson, 1945). With advances in neuroimaging technology, it is now well established that brain atrophy occurs in approximately 50% of MS patients studied in vivo (Chard et al., 2002b; Kassubek et al., 2003; Miller et al., 2002; Pelletier et al., 2003; Zivadinov and Bakshi, 2004). Atrophy of gray (GM) and white matter (WM) occurs early in the disease (Chard et al., 2002b; Chen et al., 2004; Ge et al., 2001; Quarantelli et al., 2003), increases with disease progression, and reflects widespread loss of myelin, axons, glial cells, and neuronal cell bodies (Minagar et al., 2004; Pelletier et al., 2004; Silber and Sharief, 1999). It is likely that clinical impairment results once brain atrophy reaches a critical threshold (Zivadinov et al., 2004a), making this an important variable to measure.

Few studies have compared measurements of WM and GM atrophy in MS. Chard et al. (2002b) found greater WM than GM atrophy in early relapsing–remitting (RR) disease, although lesion load correlated only with GM loss. In another recent study (Ge et al., 2001), RR patients had more WM atrophy than controls but the same GM volume. These authors reported a significant correlation between total lesion volume and WM but not GM volume. Conversely, a more recent study with 50 RR patients found lower GM, but not WM volume, when compared to normal controls (Quarantelli et al., 2003). In this study, it was also found that total lesion volume was associated with GM but not WM loss. It has been demonstrated that over a 3-year period, patients converting to clinically definite MS were those who also developed significant GM but not WM atrophy (Dalton et al., 2004).

Effect of MRI coregistration on serial short-term brain volume changes in multiple sclerosis

David A. Fritz*, Michael G. Dwyer*, Francesca Bagnato[†], Kelly L. Watts*, Alessio Bratina[‡], Marino Zorzon[‡], Valentina Durastanti[§], Laura Locatelli[‡], Enrico Millefiorini[§] and Robert Zivadinov*

*Buffalo Neuroimaging Analysis Center, Department of Neurology, The Jacobs Neurological Institute, University at Buffalo, State University of New York, Buffalo, NY, USA
[†]Neuroimmunology Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA
[‡]Department of Clinical, Morphological and Technological Sciences, University of Trieste, Trieste, Italy
[§]Department of Neurological Sciences, University of Rome La Sapienza, Rome, Italy

Objective: To test the effect of serial magnetic resonance (MR) coregistration on short-term brain volume changes using different semiautomated and automated brain volume techniques in patients with relapsing-remitting (RR) multiple sclerosis (MS). Coregistration is frequently used to increase precision in serial MR imaging (MRI) analyses. However, the effect of coregistration on measurement of whole brain volume changes from serial scans in the short term has not been tested in MS patients.
Methods: Twenty-eight patients with RR MS [mean disease duration: 4.9 years, mean age: 34.4 years and mean expanded disability status scale (EDSS): 1.4] were scanned at baseline and monthly for a period of 3 months with 2D spin-echo T1-weighted sequences obtained with non-gapped 3 mm axial slices. Percent brain parenchymal fraction change (PBPFC) was calculated by a semiautomated (Buffalo) and, separately, by two automated (Buffalo automated and SIENAX) techniques, whereas percent brain volume change (PBVC) was calculated by the SIENA technique. For coregistration of serial images we used a robust, fully automated linear image coregistration tool. PBPFC and PBVC were calculated before and after coregistration, comparing scans from the following time periods: (1) baseline to month 3; (2) baseline to month 1; (3) month 1 to 2 and (4) month 2 to 3.
Results: The highest median PBPFCs measured on non-coregistered images were detected for the baseline-to-month-3 time period and ranged from -0.11% for Buffalo semiautomated to -0.45% for Buffalo automated (p=ns). On coregistered images, the highest PBPFCs were detected for the baseline-to-month-3 time period and ranged from 0.3% for Buffalo semiautomated, -0.3% for Buffalo automated, 0.02% for SIENAX and -0.02% for SIENA (PBVC). At all time points of the study, no significant differences of median volume changes were measured on coregistered and non-coregistered images when comparing the results among the segmentation algorithms.
Conclusions: Over a 3 month period we did not detect short-term changes in normalized brain volumes using different measurement techniques. A longer observation period is needed to assess whether coregistration can affect the measurement of long-term brain volume changes.
[Neurol Res 2006; **28**: 275–279]

Keywords: Magnetic resonance imaging; coregistration; brain atrophy; multiple sclerosis

INTRODUCTION

Whole brain atrophy measures are used to evaluate macroscopic neurodegenerative disease progression in patients with multiple sclerosis (MS)[1]. Longitudinal studies have demonstrated the ability of brain parenchymal fraction (BPF), a self-normalized whole brain atrophy measure, to detect significant brain volume change over both the short and long term[2,3]. Because these changes can be very small, accurate detection requires precise and reproducible techniques[4].

One particular hindrance to measurement of such highly reproducible volume changes is inaccurate patient repositioning during follow-up magnetic resonance imaging (MRI) sessions. Follow-up scans acquired in a slightly different plane or at a different angle from baseline scans may result in artificial

Neocortical Atrophy, Third Ventricular Width, and Cognitive Dysfunction in Multiple Sclerosis

Ralph H. B. Benedict, PhD; Jared M. Bruce, PhD; Michael G. Dwyer, BS; Nadir Abdelrahman, MD;
Sarah Hussein, BS; Bianca Weinstock-Guttman, MD; Neeta Garg, MD;
Frederick Munschauer, MD; Robert Zivadinov, MD

Background: Cognitive dysfunction is common in multiple sclerosis (MS). Correlations are reported between atrophy and neuropsychological test results.

Objective: To determine if neocortical volume would supplant or supplement third ventricular width and other magnetic resonance imaging measures when predicting neuropsychological impairment.

Design: Cross-sectional study.

Setting: University MS clinic.

Participants: Seventy-seven patients with relapsing-remitting MS, 42 patients with secondary progressive MS, and 27 healthy control subjects.

Main Outcome Measures: Brain atrophy and lesion burden measures were obtained in all patients. A subset of 82 patients and all controls underwent neuropsychological testing.

Results: Patients with secondary progressive MS had more atrophy than patients with relapsing-remitting MS and controls. Neocortical volume was significantly correlated with all neuropsychological measures, with *r* values ranging from 0.29 to 0.58. Third ventricular width was retained in most stepwise regression analyses predicting cognitive impairment in patients with MS and distinguishing secondary progressive from relapsing-remitting courses of MS.

Conclusions: We confirm an association between neocortical volume and multiple cognitive domains in MS, although neocortical volume did not explain significantly more variance than other magnetic resonance imaging measures. Of the magnetic resonance imaging variables studied, third ventricular width was retained in most regression models.

Arch Neurol. 2006;63:1301-1306

Author Affiliations:
Department of Neurology, School of Medicine, State University of New York at Buffalo, and Jacobs Neurological Institute (Drs Benedict and Zivadinov) and Buffalo Neuroimaging Analysis Center (Drs Benedict, Bruce, Abdelrahman, Weinstock-Guttman, Garg, Munschauer, and Zivadinov, Mr Dwyer, and Ms Hussein).

APPROXIMATELY 50% OF PATIENTS with multiple sclerosis (MS) exhibit cognitive impairment[1] that adversely affects employability[2] and social skills.[3] Research findings highlight the value of magnetic resonance (MR) imaging in predicting cognitive dysfunction in MS. Magnetization transfer ratio,[4] whole brain atrophy,[5,6] cortical atrophy,[7] and lesion volume[8] correlate with cognitive function. Arguably, the simple measure of third ventricular width (TVW) has shown the highest correlation with cognitive dysfunction compared with other MR imaging measures.[9] In that study, TVW was significantly correlated with a wide range of tests measuring verbal memory, visuospatial memory, and processing speed; however, the regression models in the study would have been enhanced by newer semiautomated and automated MR imaging techniques that measure neocortical atrophy.

There is evidence of neocortical pathologic features in MS,[10-13] but only 1 study[7] has assessed the relationship between neocortical volume (NCV) and cognition, to our knowledge. That study showed significant correlation between NCV and measures of auditory and verbal memory, verbal fluency, and attention in 41 patients. In the present study, we aimed to replicate this work in a larger sample and to determine whether NCV would supplant or supplement TVW and other MR imaging measures in regression models predicting cognitive impairment.

METHODS

SUBJECTS

Patients with MS (n=119) provided informed consent and met diagnostic criteria for MS[14] and for relapsing-remitting (RR) or secondary progressive (SP) disease.[15] Exclusion criteria were

67

Quantitative diffusion weighted imaging measures in patients with multiple sclerosis

Eleonora Tavazzi,[a,b] Michael G. Dwyer,[a] Bianca Weinstock-Guttman,[c] Jordan Lema,[a] Stefano Bastianello,[b] Roberto Bergamaschi,[b] Vittorio Cosi,[b] Ralph H.B. Benedict,[a,c] Frederick E. Munschauer III,[c] and Robert Zivadinov[a,c,*]

[a]Buffalo Neuroimaging Analysis Center, Department of Neurology, University at Buffalo, State University of New York, Buffalo, NY, USA
[b]Department of Neurology, IRCCS, C. Mondino, University of Pavia, Pavia, Italy
[c]The Jacobs Neurological Institute, Department of Neurology, University at Buffalo, State University of New York, Buffalo, NY, USA

Diffusion-weighted imaging (DWI) has been proposed as a sensitive measure of disease severity capable of detecting subtle changes in gray matter and white matter brain compartments in patients with multiple sclerosis (MS). However, DWI has been applied to the study of MS clinical subtypes in only a few studies. The objective of this study was to demonstrate the validity of a novel, fully automated method for the calculation of quantitative DWI measures. We also wanted to assess the correlation between whole brain (WB)-DWI variables and clinical and MRI measures of disease severity in a large cohort of MS patients. For this purpose we studied 432 consecutive MS patients (mean age 44.4 ± 10.2 years), 16 patients with clinically isolated syndrome (CIS) and 38 normal controls (NC) using 1.5 T brain MRI. Clinical disease subtypes were as follows: 294 relapsing–remitting (RR), 123 secondary-progressive (SP) and 15 primary-progressive (PP). Mean disease duration was 12 ± 10 years. Mean Expanded Disability Status Scale (EDSS) was 3.3 ± 2.1. Brain parenchymal fraction (BPF), gray matter fraction (GMF) and white matter fraction (WMF) were calculated using a fully automated method. Mean parenchymal diffusivity (MPD) maps were created. DWI indices of peak position (PP), peak height (PH), MPD and entropy (ENT) were obtained. T2- and T1-lesion volumes (LV), EDSS, ambulation index (AI) and nine-hole peg test (9-HPT) were also assessed. MS patients had significantly lower BPF ($d=1.26$; $p<0.001$) and GMF ($d=0.61$; $p=0.003$), and higher ENT ($d=1.2$; $p<0.0001$), MPD ($d=1.04$; $p<0.0001$) and PH ($d=0.47$; $p=0.045$) than NC subjects. A GLM analysis, adjusted for age and multiple comparisons, revealed significant differences between different clinical subtypes for BPF, GMF, ENT, PH, PP, T2-LV and T1-LV ($p<0.0001$), WMF ($p=0.001$) and MPD ($p=0.023$). In RR and SP MS patients, ENT showed a more robust correlation with other MRI ($r=0.54$ to 0.67, $p<0.0001$) and clinical ($r=0.31$ to 0.36, $p<0.0001$) variables than MPD ($r=0.23$ to 0.41, $p<0.001$ for MRI

and $r=0.13$ to 0.18; $p=0.006$ to $p<0.001$ for clinical variables). The GMF and BPF showed a slightly stronger relationship with all clinical variables ($r=0.33$ to 0.48; $p<0.0001$), when compared to both lesion and DWI measures. ENT ($R^2=0.28$; $p<0.0001$) and GMF ($R^2=0.26$; $p<0.001$) were best related with SP disease course. This study highlights the validity of DWI in discerning differences between NC and MS patients, as well as between different MS subtypes. ENT is a sensitive marker of overall brain damage that is strongly related to clinical impairment in patients with SP MS.
© 2007 Elsevier Inc. All rights reserved.

Introduction

Diffusion weighted imaging (DWI) is a magnetic resonance imaging (MRI) technique that measures tissue water diffusional motion and, as a consequence, provides information about orientation, size and geometry of the tissue. The mobility of water molecules is reduced in highly organized tissue-like white matter (WM) and gray matter (GM) because of interactions with cellular and tissue structures, so the apparent diffusion coefficient (ADC) is lower in those tissues than in pure water. Conventionally, the average ADC is calculated from three orthogonal directions that provide information about the overall diffusivity in the tissue.

Pathological processes that modify tissue organization can cause abnormal water motion, with the consequence of altered ADC values. In multiple sclerosis (MS), the two main pathological processes affecting the brain are demyelination and neurodegeneration; they can alter the ability of tissues to restrict water motion, resulting in an increase of water diffusivity measurable with different DWI indices.

* Corresponding author. Department of Neurology, School of Medicine and Biomedical Sciences, The Jacobs Neurological Institute, 100 High St., Buffalo, NY 14203, USA. Fax: +1 716 859 7874.
E-mail address: rzivadinov@thejni.org (R. Zivadinov).

Available online on ScienceDirect (www.sciencedirect.com).

Application of hidden Markov random field approach for quantification of perfusion/diffusion mismatch in acute ischemic stroke

Michael G. Dwyer*, Niels Bergsland*, Erik Saluste†, Jitendra Sharma‡,
Zeenat Jaisani‡, Jacqueline Durfee*, Nadir Abdelrahman*, Alireza Minagar§,
Romy Hoque§, Frederick E. Munschauer III‡ and Robert Zivadinov*‡

*Buffalo Neuroimaging Analysis Center, Department of Neurology, State University of New York, Buffalo,
NY, USA
†Kaleida Health, Buffalo, NY, USA
‡The Jacobs Neurological Institute, Department of Neurology, State University of New York, Buffalo, NY,
USA
§Department of Neurology, Louisiana State University Health Sciences Center, Shreveport, LA, USA

The perfusion/diffusion 'mismatch model' in acute ischemic stroke provides the potential to
more accurately understand the consequences of thrombolytic therapy on an individual patient
basis. Few methods exist to quantify mismatch extent (ischemic penumbra) and none have
shown a robust ability to predict infarcted tissue outcome. Hidden Markov random field (HMRF)
approaches have been used successfully in many other applications. The aim of the study was to
develop a method for rapid and reliable identification and quantification of perfusion/diffusion
mismatch using an HMRF approach. An HMRF model was used in combination with automated
contralateral identification to segment normal tissue from non-infarcted tissue with perfusion
abnormality. The infarct was used as a seed point to initialize segmentation, along with the
contralateral mirror tissue. The two seeds were then allowed to compete for ownership of all
unclassified tissue. In addition, a novel method was presented for quantifying tissue
salvageability by weighting the volume with the degree of hypoperfusion, allowing the
penumbra voxels to contribute unequal potential damage estimates. Simulated and in vivo
datasets were processed and compared with results from a conventional thresholding approach.
Both simulated and in vivo experiments demonstrated a dramatic improvement in accuracy with
the proposed technique. For the simulated dataset, the mean absolute error decreased from
171.9% with conventional thresholding to 2.9% for the delay-weighted HMRF approach. For the
in vivo dataset, the mean absolute error decreased from 564.6% for thresholding to 34.2% for the
delay-weighted HMRF approach. The described method represents a significant improvement
over thresholding techniques. [Neurol Res 2008; 30: 827–834]

Keywords: Perfusion-weighted imaging; diffusion-weighted imaging; acute ischemic stroke;
ischemic penumbra; stroke quantification; hidden Markov random field approach

INTRODUCTION

The use of tissue plasminogen activator (tPA) thrombolytic therapy in acute stroke treatment has dramatically improved clinical outcomes for many patients[1,2]. However, its administration has been primarily limited to a 3 hour time window, with the majority of patients arriving too late for treatment[3]. Unfortunately, the substantial risk of increased hemorrhagic transformation that accompanies tPA therapy greatly limits its further use[4]. Despite this risk, increasing evidence shows that there is a large subset of patients who may benefit from therapy even after the 3 hour window[5,6]. A more accurate understanding of potential benefit and risk that

is patient-specific can greatly assist clinicians and patients in making an informed decision with regard to this treatment[5].

The MRI perfusion/diffusion mismatch model has shown promise as a means of identifying tissue that is functionally impaired and at risk of infarction, but may be salvaged via timely reperfusion[7–9]. Diffusion-weighted imaging (DWI) measures the mobility of water molecules within tissue, showing acute infarction as hyperintense (due to decreased water motion)[10,11]. Perfusion-weighted imaging (PWI) measures the flow of blood to the brain and is very sensitive to abnormal blood flow or arrival times[12]. Thus, DWI is capable of imaging infarcted tissue and PWI is capable of imaging tissue with abnormal blood supply. By combining these two techniques, areas that are not yet infarcted but are at risk due to decreased blood supply can be directly

A sensitive, noise-resistant method for identifying focal demyelination and remyelination in patients with multiple sclerosis via voxel-wise changes in magnetization transfer ratio

Michael Dwyer [a,*], Niels Bergsland [a], Sara Hussein [a], Jackie Durfee [a], David Wack [a], Robert Zivadinov [a,b]

[a] Buffalo Neuroimaging Analysis Center, State University of New York, Buffalo, NY, USA
[b] The Jacobs Neurological Institute, Department of Neurology, University at Buffalo, State University of New York, Buffalo, NY, USA

ARTICLE INFO

ABSTRACT

Magnetization transfer imaging (MTI) provides a reliable and histopathologically validated means for identifying important tissue changes in multiple sclerosis (MS), including demyelination and remyelination. However, most approaches to date have been based on a priori regions of interest (ROIs) and have been relatively insensitive to small focal changes or competing processes. More recent techniques have sought to address this through a voxel-wise approach, but have been limited in their detection capabilities by the amount of noise in standard MTR images. To address this issue while remaining sensitive to local changes, we propose the use of the recently introduced threshold-free cluster enhancement (TFCE) technique in combination with a Monte Carlo estimation approach. TFCE is first applied to enhance individual voxels based on their level of local cluster support, and then Monte Carlo estimation is performed to allow meaningful statistical interpretation of the resulting TFCE values. We validated this technique in three complementary ways: healthy control scan–rescan analysis, analysis of a "gold standard" simulated dataset, and analysis of a group of MS patients and healthy volunteers with 1-year longitudinal MRI scans. Scan–rescan analysis demonstrated a very low false-positive rate (1.44 mL increasing and 1.48 mL decreasing at the optimal detection threshold). Simulated dataset analysis yielded an area under receiver-operating characteristic curve of 0.942 (compared to 0.801 for a more conventional voxel-wise thresholding analysis). Finally, analysis of the real subject population showed highly significant differences ($p<0.001$) in volume of decreasing MTR between patients and controls. The proposed method provides a valuable means for quantifying MS-related tissue changes, particularly demyelination and remyelination, in vivo and without the use of highly complex or experimental MRI acquisition techniques. It improves on the sensitivity of other approaches, and may increase the statistical power of studies investigating the effects of therapy on MRI outcomes in MS.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Multiple sclerosis (MS) is a chronic, immune-mediated disease of the central nervous system (CNS). It is a complex, multi-factorial disease that includes inflammatory and neurodegenerative processes manifesting both focally in the form of lesions and diffusely in otherwise normal-appearing brain tissue (NABT) [10,13,15,25,26]. The key hallmarks of this disease are the ongoing processes of demyelination and remyelination [3,5,32,50]. Studies using magnetic resonance imaging (MRI) [3,12,21,28,31,46], histopathological analysis [3,7,32,39], and animal models [11] have demonstrated that these processes occur in lesions as well as in NABT. The ability to detect both demyelination and remyelination in vivo could provide substantial benefits, including

better tracking of individual patient responses to therapy and better overall understanding of the effects of various therapies on disease progression.

Unfortunately, myelin itself cannot currently be directly and specifically imaged via MRI because its relaxation time is too short. However, magnetization transfer imaging (MTI) has been proposed as a sensitive and reliable surrogate measure [13,14,18,50]. The magnetization transfer contrast (MTC) technique uses an off-resonance radio-frequency (RF) pulse to selectively saturate protons bound to macromolecules (including myelin). Because bound protons have a very wide resonance range while free water protons have a very narrow resonance range, a sufficiently offset pulse is able to excite the bound protons without significantly affecting free water protons. When these different protons are in contact with each other, magnetization can be exchanged from the bound protons to the free protons, indirectly causing a signal change from the free protons. The net result is a loss of signal from the free protons that becomes greater

* Corresponding author.
E-mail address: mgdwyer@bnac.net (M. Dwyer).

Relationship of optic nerve and brain conventional and non-conventional MRI measures and retinal nerve fiber layer thickness, as assessed by OCT and GDx: A pilot study

Elliot M. Frohman [a,b,*], Michael G. Dwyer [c], Teresa Frohman [a], Jennifer L. Cox [c], Amber Salter [a], Benjamin M. Greenberg [a], Sara Hussein [c], Amy Conger [a], Peter Calabresi [d], Laura J. Balcer [e], Robert Zivadinov [c,f]

[a] Department of Neurology, University of Texas Southwestern Medical Center at Dallas, USA
[b] Department of Ophthalmology, University of Texas Southwestern Medical Center at Dallas, USA
[c] Buffalo Neuroimaging Analysis Center, State University of New York, Buffalo, NY, USA
[d] Department of Neurology, Johns Hopkins Hospital, USA
[e] Department of Neurology and Epidemiology, University of Pennsylvania School of Medicine, USA
[f] The Jacobs Neurological Institute, State University of New York, Buffalo, NY, USA

ARTICLE INFO

ABSTRACT

Background: Measurement of retinal nerve fiber layer (RNFL) thickness in multiple sclerosis (MS) is gaining increasing attention.
Objectives: To explore the relationship between RNFL thickness as measured by optical coherence tomography (OCT) and scanning laser polarimetry with variable corneal compensation (GDx), and conventional and non-conventional optic nerve and brain MRI measures.
Methods: Twelve relapsing–remitting (RR) MS patients (12 affected and 12 unaffected eyes) and 4 age- and sex-matched normal controls (NC) (8 unaffected eyes) were enrolled. Four MS patients had a history of bilateral optic neuritis (ON), four had a history of unilateral ON, and 4 had no history of ON. Optic nerve MRI measurements included the length of T2 lesions, measurement of optic nerve atrophy, magnetization transfer ratio (MTR) and diffusion tensor imaging (DTI) measures. Optic nerve atrophy was measured by a novel method with high reproducibility. Brain MRI measurements included T1 and T2 lesion volumes (LVs) and their relative MTRs, and tissue class specific atrophy, MTR and DTI measures. Measures of RNFL were evaluated with OCT and GDx. We also evaluated both high and low contrast letter acuities (LCLA) in order to determine the relationship between vision, MRI metrics, and retinal structural architecture.
Results: LCLA, RNFL-OCT and optic nerve radius measures showed more robust differences between NC and MS patients, and between MS patients with affected and unaffected eyes. T2-LV and T1-LV, as well as gray matter atrophy, DTI and MTR measures were related to LCLA and RNFL thickness. Unique additive variance regression models showed that both brain and optic nerve MRI measures independently accounted for about 50% of the variance in LCLA and RNFL thickness. In reverse models, about 20% of the additional independent variance was explained by optic nerve or brain MRI metrics.
Conclusions: Measurement of RNFL thickness and radius of the optic nerve should be preferred to the other optic nerve MRI measures in clinical studies. Whole brain lesion and GM measures are predictive of impaired visual function with corresponding structural concomitants.

1. Introduction

A major objective in multiple sclerosis (MS) therapeutics is to develop strategic targeting of discrete central nervous system anatomic pathways, in order to precisely model and confirm neuroprotective and potentially even restorative properties of novel treatments [1]. Conventional and non-conventional MRI techniques are cur-

rently considered the most sensitive and reliable surrogate markers for assessing diffuse inflammatory and axonal pathology in patients with MS [2].

Measurement of retinal fiber layer thickness (RNFL) by optical coherence tomography (OCT) and scanning laser polarimetry with variable corneal compensation (GDx) is emerging as a promising tool for characterizing the amount of axonal loss in patients with MS [3,4]. RNFL thickness principally reflects axonal density, given that there is no myelin in the retina (myelination begins at the lamina cribrosa). Patients with complete recovery following monosymptomatic acute optic neuritis (ON) demonstrate significant thinning of the RNFL and

* Corresponding author. Department of Neurology and Opthalmology, University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75235, USA.
E-mail address: elliot.frohman@utsouthwestern.edu (E.M. Frohman).
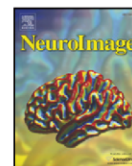
Signal abnormalities on 1.5 and 3 Tesla brain MRI in multiple sclerosis patients and healthy controls. A morphological and spatial quantitative comparison study

Carol Di Perri [a,b], Michael G. Dwyer [a], David S. Wack [a], Jennifer L. Cox [a], Komal Hashmi [a], Erik Saluste [c], Sara Hussein [a], Claudiu Schirda [a], Milena Stosic [a], Jacqueline Durfee [a], Guy U. Poloni [b], Navdeep Nayyar [a], Roberto Bergamaschi [b], Robert Zivadinov [a,*]

[a] Buffalo Neuroimaging Analysis Center, The Jacobs Neurological Institute, Department of Neurology, State University of New York, Buffalo, NY, USA
[b] Department of Neuroradiology, IRCCS, C. Mondino, University of Pavia, Pavia, Italy
[c] Buffalo Niagara MRI Center, Buffalo General Hospital, Kaleida Health, Buffalo, NY, USA

ARTICLE INFO

ABSTRACT

Previous studies in patients with multiple sclerosis (MS) revealed increased lesion count and volume on 3 T compared to 1.5 T. Morphological and spatial lesion characteristics between 1.5 T and 3 T have not been examined. The aim of this study was to investigate the effect of changing from a 1.5 T to a 3 T MRI scanner on the number, volume and spatial distribution of signal abnormalities (SA) on brain MRI in a sample of MS patients and normal controls (NC), using pair- and voxel-wise comparison procedures. Forty-one (41) MS patients (32 relapsing-remitting and 9 secondary-progressive) and 38 NC were examined on both 1.5 T and 3 T within one week in random order. T2-weighted hyperintensities (T2H) and T1-weighted hypointensities (T1H) were outlined semiautomatically by two operators in a blinded fashion on 1.5 T and 3 T images. Spatial lesion distribution was assessed using T2 and T1 voxel-wise SA probability maps (SAPM). Pair-wise analysis examined the proportion of SA not simultaneously outlined on 1.5 T and 3 T. A posteriori unblinded analysis was conducted to examine the non-overlapping identifications of SA between the 1.5 T and 3 T. For pair-wise T2- and T1-analyses, a higher number and individual volume of SA were detected on 3 T compared to 1.5 T ($p<0.0001$) in both MS and NC. Logistic regression analysis showed that the likelihood of missing SA on 1.5 T was significantly higher for smaller SA in both MS and NC groups. SA probability map (SAPM) analysis revealed significantly more regionally distinct spatial SA differences on 3 T compared to 1.5 T in both groups ($p<0.05$); these were most pronounced in the occipital, periventricular and cortical regions for T2H. This study provides important information regarding morphological and spatial differences between data acquired using 1.5 T and 3 T protocols at the two scanner field strengths.

© 2009 Elsevier Inc. All rights reserved.

Introduction

Magnetic resonance imaging (MRI) has a unique sensitivity for detecting tissue abnormalities in the central nervous system (CNS). Therefore, its use has been progressively increased in recent decades, enabling better diagnosis and prognosis of several neurological diseases.

In particular, MRI is the most sensitive diagnostic method for detection of inflammatory lesions in the CNS in patients with multiple sclerosis (MS) (Zivadinov, 2007). MRI plays an important role in diagnosis and prognosis of MS (McDonald et al., 2001) and is also commonly used as a surrogate marker to monitor disease activity in clinical trials (Miller, 1995; Paty et al., 1994).

The MRI criteria for MS focus on evidence for dissemination of lesions in time and space (McDonald et al., 2001), but several factors can affect the number and volume of MS lesions that can be identified on serial MRI scans. These include: the choice of pulse sequence (Filippi et al., 1996; Patola et al., 2001; Yousry et al., 1997), slice thickness (Dolezal et al., 2007; Filippi et al., 1998b; Molyneux et al., 1998), repositioning errors (Filippi et al., 1997), spatial resolution (Molyneux et al., 1998), differences among types of scanners (Filippi et al., 1999) and magnetic field strengths (Schima et al., 1993; Sicotte et al., 2003).

As high-field imaging becomes increasingly available in clinical routine care, it is more important than ever to understand the impact on SA detection rate of a change to higher field, including commonly associated changes in sequence parameters and/or resolution.

Higher field strength has been shown to detect more SA in clinically defined MS (Fischbach and Bruhn, 2008; Keiper et al., 1998; Lee et al., 1995; Sicotte et al., 2003) and in clinically isolated syndrome (Wattjes et al., 2006a, 2008).

* Corresponding author. School of Medicine and Biomedical Sciences, State University of New York Director, Buffalo Neuroimaging Analysis Center, The Jacobs Neurological Institute, 100 High St., Buffalo, NY 14203, USA. Fax: +1 716 859 7874.
E-mail address: rzivadinov@bnac.net (R. Zivadinov).
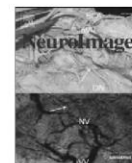
Abnormal subcortical deep-gray matter susceptibility-weighted imaging filtered phase measurements in patients with multiple sclerosis
A case-control study

Robert Zivadinov [a,b,*], Mari Heininen- Brown [a], Claudiu V. Schirda [a], Guy U. Poloni [a], Niels Bergsland [a], Christopher R. Magnano [a], Jacqueline Durfee [a], Cheryl Kennedy [a], Ellen Carl [a], Jesper Hagemeier [a], Ralph H.B. Benedict [b], Bianca Weinstock-Guttman [b], Michael G. Dwyer [a]

[a] *Buffalo Neuroimaging Analysis Center, University at Buffalo, Buffalo, NY, USA*
[b] *The Jacobs Neurological Institute, University at Buffalo, Buffalo, NY, USA*

ARTICLE INFO

ABSTRACT

Objective: To investigate abnormal phase on susceptibility-weighted imaging (SWI)-filtered phase images indicative of iron content, in subcortical deep-gray matter (SDGM) of multiple sclerosis (MS) patients and healthy controls (HC), and to explore its relationship with MRI outcomes.

Methods: 169 relapsing–remitting (RR) and 64 secondary-progressive (SP) MS patients, and 126 age- and sex-matched HC were imaged on a 3 T scanner. Mean phase of the abnormal phase tissue (MP-APT), normal phase tissue volume (NPTV) and normalized volume were determined for total SDGM, caudate, putamen, globus pallidus, thalamus, pulvinar nucleus of thalamus (PVN), hippocampus, amygdala, nucleus accumbens, red nucleus and substantia nigra. 63 HC were used for establishment of normal reference phase values, while additional 63 HC were used for blinded comparisons with MS patients.

Results: Increased MP-APT, decreased normalized volume and decreased NPTV were detected in total SDGM, caudate, putamen, globus pallidus, thalamus and PVN in MS patients compared to HC ($p < .0004$). MS patients also showed decreased volume in hippocampus ($< .0001$) and decreased NPTV in the hippocampus, amygdala and accumbens ($< .0004$). SPMS patients had increased MP-APT, decreased volume and decreased NPTV in total SDGM, caudate and amygdala compared to RRMS ($p < .005$), while individual measure differences were also detected in putamen, thalamus, hippocampus and accumbens ($p < .006$). RRMS patients showed a significant relationship between increased MP-APT and increased lesion burden and more advanced brain atrophy ($p < .004$).

Conclusions: Abnormal phase, indicative of higher iron content was significantly increased in MS patients compared to HC, and was related to more severe lesion burden and brain atrophy.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Increased iron deposition has been described previously in multiple sclerosis (MS) (Adams, 1988; Bakshi et al., 2000; Craelius et al., 1982; LeVine and Chakrabarty, 2004; LeVine et al., 1999; Mehindate et al., 2001; Valberg et al., 1989). However, the precise role of increased iron is not clear (Grimaud et al., 1995). Previous imaging approaches designed to study the presence of abnormal iron deposits in brain parenchyma have focused on T2- and T2*-weighted imaging (WI). Using these techniques, investigators have shown correlations of increases in putative iron content with clinical progression (Neema et al., 2009a; Tjoa et al., 2005), cognitive impairment (Brass et al., 2006) and brain atrophy (Bakshi et al., 2001; Bakshi et al., 2002; Bermel et al., 2005; Khalil et al., 2009). Nonetheless, direct interpretation of these results has suffered from the non-specificity of T2-based approaches, which are highly sensitive to a range of tissue changes other than iron (Haacke et al., 2004; Neema et al., 2007).

More recently, greater attention has been paid to the phase component of MRI acquisition (Haacke et al., 2009b). The phase values of individual voxels can in fact provide more information about the presence of substances with different magnetic properties than normal tissue (Haacke et al., 2007). It is known that paramagnetic substances such as deoxyhemoglobin and ferritin change the local magnetic field and thus influence the frequency or "phase" of proton spin isochromats. Tissues differ in their susceptibility to phase effects, making possible a form of contrast enhancement called susceptibility-weighted imaging (SWI). As the constituents of iron markedly

BMC
Medical Imaging

TECHNICAL ADVANCE

Improved assessment of multiple sclerosis lesion segmentation agreement via detection and outline error estimates

David S Wack[1,2,5*], Michael G Dwyer[1], Niels Bergsland[1], Carol Di Perri[3], Laura Ranza[3], Sara Hussein[1],
Deepa Ramasamy[1], Guy Poloni[1] and Robert Zivadinov[1,4]

Abstract

Background: Presented is the method "Detection and Outline Error Estimates" (DOEE) for assessing rater agreement in the delineation of multiple sclerosis (MS) lesions. The DOEE method divides operator or rater assessment into two parts: 1) Detection Error (DE) -- rater agreement in detecting the same regions to mark, and 2) Outline Error (OE) -- agreement of the raters in outlining of the same lesion.

Methods: DE, OE and Similarity Index (SI) values were calculated for two raters tested on a set of 17 fluid-attenuated inversion-recovery (FLAIR) images of patients with MS. DE, OE, and SI values were tested for dependence with mean total area (MTA) of the raters' Region of Interests (ROIs).

Results: When correlated with MTA, neither DE ($\rho = .056$, p=.83) nor the ratio of OE to MTA ($\rho = .23$, p=.37), referred to as Outline Error Rate (OER), exhibited significant correlation. In contrast, SI is found to be strongly correlated with MTA ($\rho = .75$, $p < .001$). Furthermore, DE and OER values can be used to model the variation in SI with MTA.

Conclusions: The DE and OER indices are proposed as a better method than SI for comparing rater agreement of ROIs, which also provide specific information for raters to improve their agreement.

Keywords: Multiple sclerosis, Detection and outline error estimates, Rater agreement, Operator agreement, Metric, Jaccard Index, Similarity index, Measure, Index, Kappa, Lesion, MRI, ROI

Background

Multiple operators are often used to draw regions of interest (ROIs) on medical images when the workload would be too great for a single operator. When using multiple operators, it is desirable to have the ROIs from each to be similar. There are multiple measures available to assess inter-rater variability, such as Kappa, Jaccard's Index (JI), Similarity Index (SI), Hausdorff Distances, Conformity and Sensibility, etc. [1-6]. We want to be able to assess an operator's (or automated method's) ability to create lesion ROIs,

using the ROIs they created. However, for any assessment we should consider whether some test scans are easier or harder than others to achieve good measured agreement on. An ideal measure would solely reflect the operator's ability, and not the difficulty of the underlying test scans.

One of the original and common results of multiple sclerosis lesion segmentation is the determination of the total lesion volume for an individual subject. A center may validate operators by their ability to draw ROIs that are in agreement with the overall lesion volume of a gold standard analysis. Fortunately, this intra-observer agreement was not found to be significantly correlated with lesion volume [7]. However, this only assesses an operator's ability to calculate total lesion volume; it does not make a strong statement about the ability of the operator to produce ROIs

* Correspondence: dswack@buffalo.edu
[1]Buffalo Neuroimaging Analysis Center, Dept. of Neurology, University at Buffalo, State University of New York at Buffalo, Buffalo, NY, USA
[2]Department of Nuclear Medicine, University at Buffalo, State University of New York at Buffalo, Buffalo, NY, USA
Full list of author information is available at the end of the article

Improved longitudinal gray matter atrophy assessment via a combination of SIENA and a 4-dimensional hidden Markov random field model

Michael Dwyer[1], Niels Bergsland[1], Robert Zivadinov[1,2]

[1]Buffalo Neuroimaging Analysis Center, State University of New York, Buffalo, NY, USA; [2]The Jacobs Neurological Institute, Department of Neurology, University at Buffalo, State University of New York, Buffalo, NY, USA

Background

- Brain atrophy is an important component of many neurologic diseases, including multiple sclerosis, Alzheimer's, Parkinson's, and epilepsy.
- Many investigators are now focusing specifically on gray matter atrophy as a potentially more sensitive and/or specific marker of disease progression
- Unfortunately, tissue-specific atrophy is also more difficult to measure from a methodological perspective. The gray/white border as seen on conventional MRI is generally not as clearly defined as the brain/CSF border.
- Small errors in estimating the precise boundary can have a very large impact on the final volumes estimated.
- One of the most common approaches to measuring tissue-specific atrophy over time is independent tissue segmentation at both baseline and follow-up time points (potentially with some spatial normalization), and then calculation of volumetric changes via simple arithmetic. While this approach is straightforward, it unfortunately incurs different measurement errors for each of the two sub-analyses. Absent a direct comparison, it is quite likely that the two segmentations will come to slightly different conclusions about the precise spatial and intensity distributions of the tissue classes involved.
- The "direct" SIENA approach is favorable, but is only currently fully workable for whole brain change measurements, not for tissue specific assessment.
- To address these issues, we developed a novel SIENAX-like multi-timepoint technique (SIENAX-MTP) by augmenting FMRIB's FAST tissue segmentation algorithm with a 4-dimensional hidden markov random field (HMRF). Additionally, we incorporated two of the ideas from SIENA into our tissue-specific analysis technique: skull-constrained scaling factor estimation and uniform brain extraction.

Methods

Inclusion of a temporal component in the HMRF

- The HMRF framework employed by SIENAX's FAST tool is designed to use spatial neighborhood information to elegantly mitigate the noise and homogeneity problems inherent in MRI-based tissue segmentation. Intuitively, when assigning tissue class labels to voxels, the algorithm attempts to minimize a global cost function that penalizes two separate things: selecting class labels whose mean intensities don't match well with the individual voxels (e.g., labeling a relatively bright voxel gray matter), and creating spatially isolated labels (e.g., a few gray matter voxels completely surrounded by white matter).
- More rigorously, the class labeling in standard FAST is selected iteratively according to:

$$\mathbf{x}^{(t)} = arg \max_{\mathbf{x} \in X} \{ P(\mathbf{y} | \mathbf{x}, \theta^{(t)}) + P(\mathbf{x}) \}$$

where x is the class labels vector, y is the observed intensity vector, and θ provides the class intensity parameters (means and standard deviations). The key component in this context is the $P(\mathbf{x})$ term, which expresses the overall a-priori probability of finding a three-dimensional class configuration matching x. Internally, FAST calculates this term via a conversion to "MRF weights".

- From a local perspective, the total label weight for a voxel i and label l is decided by iterating over its neighbors $j \in N_i$ (where N_i is the neighborhood of the 26 voxels surrounding i), and is calculated as:

$$w_{i,l} \sum_{j \in N_i} \frac{1}{d(i,j)} p(x_j = l)$$

where $d(i,j)$ is the distance between voxels i and j, and $p(x_j=l)$ is the current iteration a posteriori probability of classification l for voxel j.

All of the mathematics behind the HMRF model in FAST can be very naturally extended from 3 dimensions to 4 with minimal modification. Recognizing this and reorganizing terms, we modified the above calculation to be:

$$w_{i,l} \sum_{j \in N_i} \frac{1}{d(i,j)} p(x_j = l) + \frac{1}{t} p(x_i' = l)$$

where $p(x_i'=l)$ is the current iteration a posteriori probability of classification l for voxel i in the same physical position as i but at the other time-point. Analytically, the 4th dimension can be included just like the other spatial ones. Intuitively, this calculation now includes a regularization term for maintaining constant voxel classification over time as well as for the usual agreement with neighboring classifications.

Skull-constrained scaling factor

- To avoid potential scanner-related scaling issues, we employed a similar skull-constrained scaling approach to that used in SIENA. Baseline and follow-up images were individually de-skulled using BET, and both brain and skull images were retained. A skull-constrained 12-parameter affine registration was then performed with FLIRT, resulting in halfway transforms and scaling factors. We used these scaling factors to correct the final volume change estimates.

Uniform brain extraction

- Another potentially confounding factor in longitudinal measurement is differing brain extractions. Because we transformed both images into a common halfway space, we were also able to combine the individual brain masks via a union operation. After masking with valid voxel masks from each time point, we were thus able to isolate a common set of voxels and to avoid spurious volume changes measurements due to differing BET extractions.
- Figure 1 represents schematic view of the SIENAX-MTP processing pipeline.

Validation

- To evaluate the characteristics of our proposed approach, we used both simulation and testing on a real clinical dataset of patients with MS and matched healthy controls.

Simulation

- To provide images with known atrophy over time, we artificially introduced volume changes in a set of images. We began with a scan-rescan dataset of 5 healthy volunteers, all scanned twice within a one-week period. We then performed brain/skull extractions on all images using BET and manually correcting where appropriate. Next, we generated transformation matrices preserving order and position but scaling by a variety of known factors in the range of 0-10% reduction. We applied these transformations to the follow-up brain images, retaining the original follow-up skull masks. We then performed the above-described SIENAX-MTP technique on the various pairs of images.

Clinical evaluation

- To evaluate whether SIENAX-MTP can better measure clinically meaningful differences, we analyzed a cohort of 64 RRMS patients who did not develop disability progression and 64 RRMS patients who did develop disability progression over 5 years of follow-up. We applied SIENA, SIENAX, and SIENAX-MTP, and tested the effect sizes between the groups. Additionally, we examined correlations between the various measures.

Results

- Simulation results are shown in Figures 2 and 3. The SIENAX-MTP technique resulted in better correlation with the actual scale values applied to the images (R=0.83 compared to R=0.23 for SIENAX). In addition, variance within the cases for each scaling value was reduced in SIENAX-MTP results as compared to standard SIENAX.
- Clinical results are shown in Tables 1 and 2. SIENAX MTP showed reduced variance and a larger effect size than SIENAX for all measures evaluated. For GM, SIENAX-MTP differences reached statistical significance, while SIENAX differences did not. SIENAX-MTP measures also correlated better with SIENA than did standard SIENAX measures.

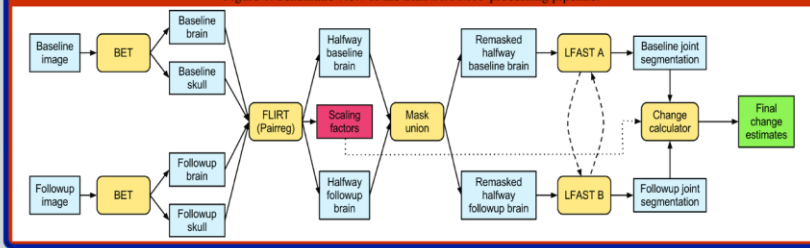Figure 1. Schematic view of the SIENAX-MTP processing pipeline.

Table 2. Spearman correlations between whole brain, gray matter and white matter volume estimation changes measures over 5-year follow-up in 64 relapsing-remitting (RR) multiple sclerosis (MS) patients with no disability progression and 64 RRMS patients with disability progression (the values are shown as r coefficients and p).

Whole brain measures	SIENA PBVC	SIENAX PBVC	SIENAX-MTP PBVC	SIENAX PGMVC	SIENAX-MTP PGMVC	SIENAX PWMVC	SIENAX-MTP PWMVC
SIENA PBVC	NA	0.55 <0.0001	0.67 <0.0001	0.46 <0.0001	0.58 <0.0001	0.29 <0.0001	0.36 <0.0001
SIENAX PBVC	0.55 <0.0001	NA	0.65 <0.0001	0.56 <0.0001	0.46 <0.0001	0.35 <0.0001	0.33 <0.0001
SIENAX-MTP PBVC	0.67 <0.0001	0.65 <0.0001	NA	0.68 <0.0001	0.64 <0.0001	0.48 <0.0001	0.61 <0.0001
SIENAX GMVC	0.46 <0.0001	0.56 <0.0001	0.68 <0.0001	NA	0.26 0.003	0.73 <0.0001	0.67 <0.0001
SIENAX-MTP PGMVC	0.58 <0.0001	0.46 <0.0001	0.64 <0.0001	0.26 0.003	NA	0.17 0.061	0.09 0.334
SIENAX PWMVC	0.29 <0.0001	0.35 <0.0001	0.48 <0.0001	0.73 <0.0001	0.17 0.061	NA	0.92 <0.0001
SIENAX-MTP PWMVC	0.36 <0.0001	0.33 <0.0001	0.61 <0.0001	0.67 <0.0001	0.09 0.334	0.92 <0.0001	NA

Legend: NA - non available; HC - healthy controls; RRMS - relapsing-remitting multiple sclerosis; PBVC - percent brain volume change; MTP - multi-timepoint; 2C - two component; 3C - three component; PGMVC - percent gray matter volume change; PWMVC - percent white matter volume change

Figure 2. Representative results from scan-rescan analysis (with 0 scaling). Differences, which should be negligible, are quite noticeable for SIENAX but not for SIENAX-MTP. These visible difference voxels are sources of error in volumetric change measurements.
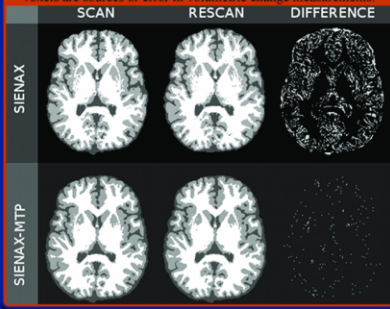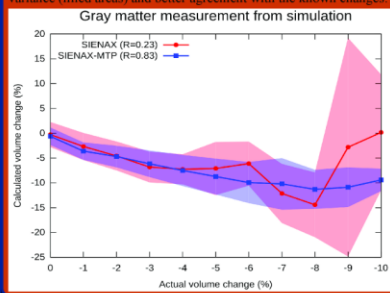
SCAN | RESCAN | DIFFERENCE

Figure 3. Results of simulated analysis. Scan-rescan data was artificially resized to simulate atrophy, and both SIENAX and SIENAX-MTP were run on the image pairs. SIENAX-MTP shows lower variance (filled areas) and better agreement with the known changes.

Gray matter measurement from simulation

Conclusions

- The proposed SIENAX-MTP technique can provide significantly improved measurement of GM atrophy over time in MS by reducing error-related variance.
- The increased statistical power gained with SIENAX-MTP can potentially benefit research studies and clinical trials by detecting changes earlier or in smaller cohorts, and/or by detecting more subtle changes.

Disclosures

References

1. S.M. Smith, N. De Stefano, M. Jenkinson, and P.M. Matthews. Normalised accurate measurement of longitudinal brain change. Journal of Computer Assisted Tomography, 25(3):466-475, May/June 2001.
2. S.M. Smith, Y. Zhang, M. Jenkinson, J. Chen, P.M. Matthews, A. Federico, and N. De Stefano. Accurate, robust and automated longitudinal and cross-sectional brain change analysis. NeuroImage, 17(1):479-489, 2002.
3. Winkler, G. Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction (2nd ed.). Springer, 2006.
4. Zhang, Y. and Brady, M. and Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imag, 20(1):45-57, 2001.

Table 1. Five year follow-up differences in whole brain, gray matter and white matter volume estimation change measures in 64 relapsing-remitting (RR) multiple sclerosis (MS) patients without disability progression and 64 RRMS patients with disability progression.

Whole brain measures	No disability progression MS group (n=64) Mean (SD) median (min/max)	Disability progression MS group (n=64) Mean (SD) median (min/max)	Cohen's d	p-value
SIENA PBVC	-2.72 (1.94) -2.34 (-8.95/1.18)	-4.85 (3.49) -3.97 (-17.21/0.79)	0.75	<0.0001
SIENAX PBVC	-5.56 (3.25) -4.86 (-16.12/-0.79)	-6.45 (3.82) -6.89 (-14.63/3.41)	0.25	0.160
SIENAX-MTP PBVC	-2.08 (1.21) -1.88 (-5.72/0.35)	-3.51 (2.23) -3.10 (-10.37/0.24)	0.81	<0.0001
GM measures				
SIENAX PGMVC	-4.18 (2.89) -4.11 (-16.42/4.82)	-5.41 (4.28) -5.01 (-18.06/4.14)	0.34	0.058
SIENAX-MTP PGMVC	-3.92 (1.81) -4.37 (-7.87/0.46)	-5.26 (2.78) -4.45 (-14.61/-1.42)	0.57	0.002
WM measures				
SIENAX PWMVC	1.47 (6.81) 2.75 (-21.87/13.84)	-2.49 (8.24) -2.42 (-24.91/19.58)	0.52	0.004
SIENAX-MTP PWMVC	0.40 (2.61) 0.88 (-9.13/4.82)	-1.52 (3.29) -0.75 (-11.22/4.35)	0.64	<0.0001

Legend: HC - healthy controls; RRMS - relapsing-remitting multiple sclerosis; SD - standard deviation; min - minimum; max - maximum; PBVC - percent brain volume change; GM - gray matter; MTP - multi-timepoint; 2C - two component; 3C - three component; PGMVC - percent GM volume change; PWMVC - percent WM volume change

Cohen's d represents the effect size. The differences between the MS groups were tested by using Student's t-test.

Appendix C – Supporting documents

This appendix collects the supporting documents published by the author. In contrast to Appendix B, the full documents have not been included in order to save space. Instead, the first page including the abstract has been provided. Full documents are available upon request to the author or directly from the associated journals. However, the unpublished manuscript related to work 11 has been reproduced in full.

DETECTION OF CORTICAL LESIONS IS DEPENDENT ON CHOICE OF SLICE THICKNESS IN PATIENTS WITH MULTIPLE SCLEROSIS

Ondrej Dolezal,*,[†] Michael G. Dwyer,* Dana Horakova,[†] Eva Havrdova,[†] Alireza Minagar,[‡] Srivats Balachandran,* Niels Bergsland,* Zdenek Seidl,[§] Manuela Vaneckova,[§] David Fritz,* Jan Krasensky,[§] and Robert Zivadinov*

*Department of Neurology, State University of New York at Buffalo Buffalo Neuroimaging Analysis Center, The Jacobs Neurological Institute Buffalo, New York 14203, USA
[†]Department of Neurology, 1st School of Medicine Charles University, Prague, Czech Republic
[‡]Department of Neurology, Louisiana State University Health Sciences Center, Shreveport Louisiana 71103, USA
[§]Department of Radiology, 1st School of Medicine Charles University, Prague, Czech Republic

Understanding the importance of cortical lesions in MS pathogenesis has changed. Histopathologic studies using new immunohistochemical methods show that cortical lesions can be detected more frequently than previously reported. Newer MRI sequences also detect cortical lesions more accurately. The aim of this study was to evaluate whether the effect of slice thickness (th) is an important factor for detection of cortical lesions in patients with multiple sclerosis (MS). We aimed also to investigate the relationship of cortical lesions with clinical status or other MRI variables. Forty-one patients with relapsing-remitting (RR) MS (11 males,

30 females with mean EDSS 2.3) underwent scans of Two-dimensional (2D)-fluid-attenuated inversion recovery (FLAIR) and 3D-T1-WI at 1.5-, 3-, and 5-mm slice thicknesses on 1.5-T MRI. Cortical and juxtacortical lesions were volumetrically assessed using a semiautomated method. 2D-FLAIR and 3D-T1-WI were coregistered and the matrix of the neocortical volume (NCV) segmentation mask (SIENAX-generated) was used to classify the location of the cortical–subcortical lesions. Cortical lesions fell into three classes: (1) class 1 were defined as lesions located in the NCV, (2) class 2 were juxtacortical lesions in contact with the NCV mask, and (3) class 3 were cortical–juxtacortical lesions situated in both regions. We measured NCV and normalized gray matter (GM) volume as well. We used partial correlation and multiple regressions to investigate the relationship between cortical lesions and other clinical and MRI variables. Of the total T2-lesion volume (T2-LV) measured on 1.5-mm th scans (mean 16108 mm^3), cortical lesions represented 2.4% (276 mm^3), juxtacortical lesions 6.1% (760 mm^3), and cortical–juxtacortical 3.7% (491 mm^3). Compared to 1.5-mm th scan, cortical LV was reduced by -28.3%, $p < 0.001$ on 3-mm th and by -40.78%, $p < 0.001$ on 5-mm th scans. Results for juxtacortical LV were for 3-mm th scans (-17.9%, $p < 0.01$) and for 5-mm th scans (-30.3%, $p < 0.01$). The figures for cortical–juxtacortical LV were also for 3-mm th scans (-16.2%, $p < 0.01$) and for 5-mm th scans (-26.7%, $p < 0.01$). We observed a significant correlation between T2-LV and GM atrophy in all slice thickness ($r = -0.4$ to -0.48, $p = 0.001$–0.003) and a modest relationship between cortical and cortical–juxtacortical LVs and disability, especially at 1.5-mm slice thickness ($r = 0.35$, $p = 0.025$). Use of thinner slices (1.5mm) on 2D-FLAIR images can significantly increase the sensitivity and precision of detecting cortical and juxtracotical lesions in patients with MS. Cortical and juxtacortical lesions contribute more to disability development than total T2-LV alone.

I. Introduction

Multiple sclerosis (MS) is an autoimmune demyelinating disease of the central nervous system (CNS), leading to disability in young adults. Historically, MS has been considered primarily a disease affecting the white matter (WM) of the CNS. Evidence is mounting that MS is also a gray matter (GM) disease (Bo *et al.*, 2003; Brownell and Hughes, 1962; Kutzelnigg and Lassmann, 2005; Peterson *et al.*, 2001; Wegner *et al.*, 2006). GM lesions are similar to those in the WM; they are described as clearly defined areas of demyelination within the cerebral cortex, basal ganglia, and GM of the spinal cord and brain stem that reflect an ongoing

Diffusion-weighted imaging predicts cognitive impairment in multiple sclerosis

Ralph HB Benedict[1,2], Jared Bruce[2,3], Michael G Dwyer[2], Bianca Weinstock-Guttman[1,2], Chris Tjoa[1,2], Eleonora Tavazzi[2], Frederick E Munschauer[1] and Robert Zivadinov[1,2]

Following a previous study with diffusion tensor imaging, we investigated the correlation between diffusion-weighted imaging (DWI) and cognitive dysfunction in multiple sclerosis (MS). We studied 60 MS patients (mean age 45.8 ± 9.0 years) using 1.5-T MRI. Disease course was RR $=40$ and SP $= 20$. Mean disease duration was 12.8 ± 8.7 years. Mean EDSS was 3.4 ± 1.7. Whole brain, gray and white matter normalized volumes were calculated on 3D SPGR T1-WI using a fully automated Hybrid SIENAX method. Parenchymal mean diffusivity (PMD) maps were created after automated segmentation of the brain parenchyma and cerebrospinal fluid using T2-WI and DW images. Histogram analysis was performed and DWI indices of peak position (PP), peak height (PH), mean parenchymal diffusivity (MPD) and entropy were obtained. Neuropsychological (NP) evaluation emphasized auditory/verbal and visual/spatial memory, as well as processing speed and executive function. We found significant correlations between DWI and performance in all cognitive domains. Overall, stronger correlations emerged for MPD and entropy than other DWI measures, although all correlations were in the expected direction. The strongest association was between DWI entropy and performance on the Symbol Digit Modalities Test, which assesses processing speed and working memory ($r = -0.54$). Fisher r to z transformations revealed that DWI, gray matter (GMF) and whole brain (BPF) atrophy, T1-lesion volume (LV) and T2-LV all accounted for similar amounts of variance in NP testing. Stepwise regression models determined whether multiple MRI measures predicted unique additive variance in test performance. GMF ($R^2 = 0.35$, $F = 30.82$, $P < 0.01$) and entropy ($\Delta R^2 = 0.06$, $\Delta F = 5.47$, $P < 0.05$) both accounted for unique variance in processing speed. Our data make a stronger case for the clinical validity of DWI in MS than heretofore reported. DWI has very short acquisition times, and the segmentation method applied in the present study is reliable and fully automated. Given its overall simplicity and *moderate* correlation with cognition, DWI may offer several logistic advantages over more traditional MRI measures when predicting the presence of NP impairment. *Multiple Sclerosis* 2007; **13**: 722–730. http://msj.sagepub.com

Key words: cognition; diffusion-weighted imaging; magnetic resonance imaging; multiple sclerosis; neuropsychology

Introduction

Recent years have witnessed increasing interest in the prediction of neuropsychological (NP) impairment in patients with multiple sclerosis (MS). Approximately 50% of MS patients exhibit some degree of NP impairment [1,2]. Deficits in processing speed [3,4], memory [5,6] and higher executive function [7,8] are particularly common, affecting quality of life and employment [9]. Identifying patients at risk for NP impairment on the basis of MRI would thus enhance quality of care. A number of MRI measures are correlated with NP impairment in MS, including lesion volume (LV) [10], ventricle size [11,12], cortical sulcal enlargement [13], cortical volume [14] and whole brain volume [15,16]. Recent publications showed that atrophy accounts for more NP variance than does lesion burden [11,12,17].

[1] Department of Neurology, State University of New York (SUNY) at Buffalo, School of Medicine, Buffalo, NY, USA
[2] Jacobs Neurological Institute, Buffalo, NY, USA
[3] Healthcare System, Buffalo VA Medical Center, Buffalo, NY, USA
Author for correspondence: Ralph HB Benedict, Neurology, D-6, Buffalo General Hospital, 100 High Street, Buffalo, NY 14203, USA. E-mail: benedict@buffalo.edu

79

ORIGINAL
RESEARCH

R. Zivadinov
A.C. Banas
V. Yella
N. Abdelrahman
B. Weinstock-Guttman
M.G. Dwyer

Comparison of Three Different Methods for Measurement of Cervical Cord Atrophy in Multiple Sclerosis

BACKGROUND AND PURPOSE: Evidence is mounting that spinal cord atrophy significantly correlates with disability in patients with multiple sclerosis (MS). The purpose of this work was to validate 3 different measures for the measurement of cervical cord atrophy on high-resolution MR imaging in patients with MS and in normal control subjects (NCs). We also wanted to evaluate the relationship between cervical cord atrophy and clinical disability in the presence of other conventional and nonconventional brain MR imaging metrics by using a unique additive variance regression model.

MATERIALS AND METHODS: We studied 66 MS patients (age, 41.2 ± 12.4 years; disease duration, 11.8 ± 10.7 years; Expanded Disability Status Scale, 3.1 ± 2.1) and 19 NCs (age, 30.4 ± 12.0 years). Disease course was relapsing-remitting (34), secondary-progressive (14), primary-progressive (7), and clinically isolated syndrome (11). The cervical cord absolute volume (CCAV) in cubic millimeters and 2 normalized cervical cord measures were calculated as follows: cervical cord fraction (CCF) = CCAV/thecal sac absolute volume, and cervical cord to intracranial volume (ICV) fraction (CCAV/ICV). Cervical and brain lesion volume measures, brain parenchyma fraction (BPF), and mean diffusivity were also calculated.

RESULTS: CCAV ($P < .0001$) and CCF ($P = .007$) showed the largest differences between NCs and MS patients and between different disease subtypes. In regression analysis predicting disability, CCAV was retained first ($R^2 = 0.498$; $P < .0001$) followed by BPF ($R^2 = 0.08$; $P = .08$). Only 8% of the variance in disability was explained by brain MR imaging measures when coadjusted for the amount of cervical cord atrophy.

CONCLUSIONS: 3D CCAV measurement showed the largest differences between NCs and MS patients and between different disease subtypes. Cervical cord atrophy measurement provides valuable additional information related to disability that is not obtainable from brain MR imaging metrics.

MR imaging of the brain is a sensitive tool for making a diagnosis of multiple sclerosis (MS). Abnormalities of brain MR imaging are present in more than 95% of patients with clinically definite MS; however, there is poor correlation between disability and the number and volume of focal brain lesions visible on MR imaging.[1]

Involvement of the spinal cord, especially of the cervical cord,[2,3] is of particular significance in the development of physical disability in patients with MS.[4,5] During the course of their disease, approximately 80% of patients with MS present with spinal cord symptoms.[6] Conventional T2-weighted spinal cord imaging is sensitive in detecting spinal cord lesions and their changes over time.[7,8] However, measures of cord T2 lesion number and volume failed to show a significant relationship with disability and have poor prognostic value for disability accumulation over the mid-to-long term.[2,3] Evidence is mounting that spinal cord atrophy significantly correlates with disability.[5,9-11]

Atrophy of the spinal cord in MS is thought to reflect inflammatory tissue injury, demyelination, and axonal loss. Postmortem pathologic studies have documented spinal cord axonal loss in MS.[12,13] However, whereas the correlation between central nervous system atrophy and disability has been interpreted as a reflection of axonal loss in pre-existing lesions,[14-16] axonal loss does not appear to directly affect the cross-sectional cord area in pathologic studies.[2] Measurement of spinal cord atrophy has demonstrated value in the clinical realm. Serial MR imaging of the spinal cord has shown evidence of disease activity undetectable on clinical examination, thereby increasing the diagnostic sensitivity of MR imaging for patients with suspected MS.[17] Spinal cord abnormalities on MR imaging are not restricted only to patients presenting with spinal cord symptoms, because changes suggestive of atrophy may be seen before any manifestation of clinical symptoms. It has been shown that atrophy of the cervical spinal cord is a useful measure for determining clinical disability[10,15,18] and monitoring disease progression,[19] as well as therapeutic drug effects in MS.[20]

Key problems in the evaluation of spinal cord atrophy have been related to poor resolution of MR imaging, small size of the cord, and surrounding fat, bone, and CSF that can cause artifacts and, as a consequence, compromise the final image quality. Indeed, artifacts related to pulsation and respiratory cardiac motion have also been considered.[2,3] This led in most of the earlier studies to unacceptable error in manual delineation of the cord/CSF interface.[2] The technical challenges of spinal cord imaging posed by the size and anatomy of the cord and by its surrounding structures have been addressed in recent years by improved receiver coils, fast imaging, 3D imaging, motion suppression, and cardiac gating. Subsequently, interest has emerged in a reproducible semiautomated measurement of the cord cross-sectional area[21] and its improved measurement by reduction of partial volume effect,[22] as well as by 3D extraction of the cord surface area.[23]

The goal of the present study was to investigate whether

SPINE

ORIGINAL RESEARCH

80

Robert Zivadinov
Milena Stosic
Jennifer L. Cox
Deepa P. Ramasamy
Michael G. Dwyer

The place of conventional MRI and newly emerging MRI techniques in monitoring different aspects of treatment outcome

■ **Abstract** Magnetic resonance imaging (MRI) is the most important paraclinical measure for assessing and monitoring the pathologic changes implicated in the onset and progression of multiple sclerosis (MS). Conventional MRI sequences, such as T1-weighted gadolinium (Gd) enhanced and spin-echo T2-weighted imaging, only provide an incomplete picture of the degree of inflammation and underlying neurodegenerative changes in this disease. Two- and three-dimensional fluid-attenuated inversion recovery and double inversion recovery sequences allow better identification of cortical, periventricular and infratentorial lesions. Ultra-high field strength MRI has the potential to detect subpial cortical and deep gray matter lesions. Unenhanced T1-weighted imaging can reveal hypointense black holes, a measure of chronic neurodegeneration. Magnetization transfer imaging (MTI) is increasingly used to characterize the evolution of MS lesions and normal-appearing brain tissue. Evidence suggests that the dynamics of magnetization transfer changes correlate with the extent of demyelination and remyelination. Magnetic resonance spectroscopy, which provides details on tissue biochemistry, metabolism, and function, also has the capacity to reveal neuroprotective mechanisms. By measuring the motion of water, diffusion imaging can provide information about the orientation, size, and geometry of tissue damage in white and gray matter. These advanced non-conventional MRI techniques relate better to clinical impairment, disease progression, accumulation of disability, and have the potential to detect neuroprotective effects of treatment. Although detecting the status of neuronal integrity using MRI techniques continues to improve, a "gold standard" model remains to be established.

■ **Key words** multiple sclerosis · magnetic resonance imaging · high field imaging · magnetization transfer imaging · magnetic resonance spectroscopy · diffusion imaging

R. Zivadinov (⊠) · M. Stosic · J. L. Cox · D. P. Ramasamy · M. G. Dwyer
Dept. of Neurology
School of Medicine and Biomedical Sciences
State University of New York
Director, Buffalo Neuroimaging Analysis Center
The Jacobs Neurological Institute
100 High St.
Buffalo, NY 14203, USA
Tel.: +1-716/859-7031
Fax: +1-716/859-7874
E-Mail: rzivadinov@thejni.org

Introduction

Over the last two decades, magnetic resonance imaging (MRI) has revolutionized the diagnosis and monitoring of patients with multiple sclerosis (MS). Over this period, MRI technology has continually evolved with more and more sophisticated techniques becoming available that may visualize more accurately the disease processes.

For much of this period, presence and accumulation of T2-weighted and gadolinium (Gd) enhancing T1-weighted lesions have represented the MRI gold standard for making the diagnosis and for evaluating long-term prognosis in MS. These measures have also been used for a long time as principal MRI outcomes in clinical trials. The principal limitation of these techniques is that they reveal only incompletely the pathophysiologic process in this disease. In particular, more than twenty clinical trials have demonstrated very pronounced inhibition of these inflammatory MRI measures without concurrent clinical benefit over the long term. A meta-analysis published in 1999 [33] analyzed data from five

Use of perfusion- and diffusion-weighted imaging in differential diagnosis of acute and chronic ischemic stroke and multiple sclerosis

Robert Zivadinov*[†], Niels Bergsland*, Milena Stosic*, Jitendra Sharma[†], Fernando Nussenbaum*, Jacqueline Durfee*, Nima Hani*, Nadir Abdelrahman*, Zeenat Jaisani[†], Alireza Minagar[‡], Romy Hoque[‡], Frederick E. Munschauer III[†] and Michael G. Dwyer*

*Buffalo Neuroimaging Analysis Center and [†]The Jacobs Neurological Institute, Department of Neurology, State University of New York, Buffalo, NY, USA
[‡]Department of Neurology, Louisiana State University Health Sciences Center, Shreveport, LA, USA

Objective: To investigate differences in lesions and surrounding normal appearing white matter (NAWM) by perfusion-weighted imaging (PWI) and diffusion-weighted imaging (DWI) in patients with acute and chronic ischemic stroke and multiple sclerosis (MS).
Methods: Study subjects included 45 MS patients, 22 patients with acute ischemic stroke and 20 patients with chronic ischemic stroke. All subjects underwent T_2-weighted imaging (WI), flair attenuated inversion recovery (FLAIR), DWI and dynamic contrast enhanced PWI. Apparent diffusion coefficient (ADC) and mean transit time (MTT) maps were generated and values were calculated in the acute and chronic ischemic and demyelinating lesions, and in NAWM for distances of 5, 10 and 15 mm. Fifty-three acute ischemic and 33 acute demyelinating lesions, and 775 chronic ischemic and 998 chronic demyelinating lesions, were examined. Univariate, multivariate and data mining analyses were used to examine the feasibility of a prediction model between different lesion types. Correctly and incorrectly classified lesions, true positive (TP), false positive (FP) and precision rates were calculated.
Results: Patients with acute ischemic lesions presented more prolonged mean MTT values in lesions (p=0.002) and surrounding NAWM for distances of 5, 10 and 15 mm (all p<0.0001) than those with acute demyelinating lesions. In multinomial logistic regression analysis, 65 of 86 acute lesions were correctly classified (75.6%). The TP rates were 81.1% for acute ischemic lesions and 66.7% for acute demyelinating lesions. The FP rates were 33.3% for acute ischemic and 18.9% for acute demyelinating lesions. The precision was 79.6% for classification of acute ischemic lesions and 68.8% for prediction of acute demyelinating lesions. The logistic model tree decision algorithm revealed that prolonged MTT of surrounding NAWM for a distance of 15 mm ($\geq$7459.2 ms) was the best classifier of acute ischemic versus acute demyelinating lesions. Patients with chronic ischemic lesions presented higher mean ADC (p<0.0001) and prolonged MTT (p=0.013) in lesions, and in surrounding NAWM for distances of 5, 10 and 15 mm (all p<0.0001), compared to the patients with chronic demyelinating lesions. Data mining analyses did not show reliable predictability for correctly discerning between chronic ischemic and chronic demyelinating lesions. The precision was 56.7% for classification of chronic ischemic and 58.9% for prediction of chronic demyelinating lesions.
Discussion: We found prolonged MTT values in lesions and surrounding NAWM of patients with acute and chronic ischemic stroke when compared to MS patients. The use of PWI is a promising tool for differential diagnosis between acute ischemic and acute demyelinating lesions. The results of this study contribute to a better understanding of the extent of hemodynamic abnormalities in lesions and surrounding NAWM in patients with MS. [Neurol Res 2008; **30**: 816–826]

Keywords: Perfusion-weighted imaging; diffusion-weighted imaging; acute ischemic stroke; chronic ischemic stroke; multiple sclerosis; demyelinating lesions

INTRODUCTION

Quantitative neuroimaging indices have advanced our understanding of the underlying pathophysiologic mechanisms of many neurological diseases. Differentiating acute and subacute ischemic stroke lesions from acute demyelinating lesions of multiple

Correspondence and reprint requests to: Robert Zivadinov, MD, PhD, Department of Neurology, School of Medicine and Biomedical Sciences, The Jacobs Neurological Institute, 100 High St., Buffalo, NY 14203, USA. [rzivadinov@bnac.net]

Gray matter atrophy and disability progression in patients with early relapsing–remitting multiple sclerosis
A 5-year longitudinal study

Dana Horakova [a,*], Michael G. Dwyer [b], Eva Havrdova [a], Jennifer L. Cox [b], Ondrej Dolezal [a], Niels Bergsland [b], Brett Rimes [b], Zdeněk Seidl [c], Manuela Vaneckova [c], Robert Zivadinov [b]

[a] Department of Neurology, Charles University in Prague, First Faculty of Medicine, Czech Republic
[b] Buffalo Neuroimaging Analysis Center, The Jacobs Neurological Institute, Dept. of Neurology, Buffalo, NY, USA
[c] Department of Radiology, Charles University in Prague, First Faculty of Medicine, Czech Republic

ARTICLE INFO

ABSTRACT

We assessed the relationship between gray matter (GM) and white matter (WM) atrophy and clinical status in early relapsing–remitting multiple sclerosis (MS) patients over 5 years. A group of 181 patients who participated in the ASA (Avonex–Steroid–Azathioprine) study and had complete clinical and MRI assessments over 2 and 5 years was investigated. One hundred seventy (170) patients completed the 12-month follow-up, 147 the 24-month, 98 the 36-month, 65 the 48-month and 47 the 60-month. Changes in GM (GMV), WM (WMV) and peripheral GM (PGV) volumes, whole brain volume (percentage brain volume change PBVC), lateral ventricle volume (LVV), third ventricle width (3VW) and T2-lesion volume (T2-LV) were measured. Patients were assigned according to their clinical status to one of two groups: the Stable group, and the Reached Confirmed Sustained Progression (RCSP) group (24-week interval). At 0–6 months PBVC and GMV, at 0–12 months PBVC, GMV and T2-LV, at 0–24 months PBVC and GMV, at 0–36 months PBVC, GMV and T2-LV, and at 0–48 PBVC predicted the differences between the RCSP and Stable groups. PBVC and LVV showed the strongest ability to differentiate patients who presented 0 or ≥3 relapses in the Stable group. Decline in PBVC and GMV were predictive markers of disability deterioration. Correlation of T2-LV with clinical status was weaker and decreased over time. Higher number of relapses was associated with faster decline in whole brain volume.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Multiple sclerosis (MS) is a chronic inflammatory disease of the central nervous system (CNS) characterized by early axonal damage that, in a substantial number of patients, leads to irreversible disability. It is now widely believed that this permanent neurological disability develops when a threshold of cumulative axonal loss is reached and CNS compensatory resources are exhausted [1]. Despite known clinical patterns [2], there is considerable individual variation in the course of MS and its clinical characteristics, and newer mathematical prognostic models [3] still do not provide satisfactory answers regarding long-term predictability of clinical disability in patients with MS.

Among different surrogate markers proposed for this disease, magnetic resonance imaging (MRI) is currently one of the most explored and utilized. Technical advances in MRI during recent years have led to increasingly accurate measurements of subclinical disease activity, along with improved understanding of the mechanisms underlying the inflammatory and neurodegenerative phases of this disease [4–6]. MRI plays an essential role in the diagnosis of MS and studies have shown that inflammatory lesion measures may accurately predict conversion to clinically definite MS in the mid-to-long term [7–9]. The correlation are less satisfactory, however, between these inflammatory MRI markers and long-term clinical status in patients with relapsing–remitting (RR), secondary-progressive or primary-progressive MS. Furthermore, an important question still awaits complete clarification, i.e., whether newer MRI techniques (such as brain atrophy) may provide an advantage in predicting long-term clinical status.

Numerous studies have underscored the usefulness of MRI in assessing brain atrophy and its relationship to long-term neurodegeneration and disability progression [4]. Brain atrophy is a modest-to-strong correlate of clinical disability (including both cognitive and physical components) and has moderate predictive value for subsequent development of neurological impairment. On the other hand, the association between brain atrophy and disability is independent of the effect of conventional MRI lesions. Moreover, carefully conducted long-term serial MRI studies,

* Corresponding author. Department of Neurology, Charles University in Prague, First Faculty of Medicine, Katerinska 30, Prague 128 08, Czech Republic. Tel.: +420 224966515; fax: +420 224917907.
E-mail address: dana.horak@post.cz (D. Horakova).

Voxel-wise magnetization transfer imaging study of effects of natalizumab and IFNβ-1a in multiple sclerosis

R Zivadinov[1,2], MG Dwyer[1], S Hussein[1], E Carl[1], C Kennedy[1], M Andrews[1], D Hojnacki[1], M Heininen-Brown[1], L Willis[1], M Cherneva[1], N Bergsland[1] and B Weinstock-Guttman[1,2]

Abstract

Objective: To determine the effects of intravenous natalizumab and intramuscular interferon beta-1a (IFNβ-1a) on the volume of white-matter (WM) lesions and normal appearing brain tissue (NABT) undergoing voxel-wise (VW) increases in magnetization transfer ratio (MTR) suggestive of remyelination in patients with relapsing multiple sclerosis.

Methods: This prospective, open-label, single-blinded study enrolled patients with relapsing–remitting multiple sclerosis (RRMS) and relapsing secondary progressive multiple sclerosis (RSPMS) as well as a group of age/sex-matched healthy controls (n=22). Patients with multiple sclerosis were assigned to receive natalizumab monotherapy (n=77; RRMS/RSPMS) or intramuscular IFNβ-1a (n=26) as either monotherapy (RRMS) or combined with pulsed i.v. methylprednisolone, as needed (RSPMS). The primary endpoint was the two-year change in volume of NABT VWMTR, by quantifying the number of voxels that increased (suggesting remyelination) or decreased (suggesting demyelination) in their MTR value.

Results: The volume of tissue undergoing increases in VWMTR was significantly larger in natalizumab compared with IFNβ-1a-treated patients (year 1: $p=0.001$ in NABT and $p<0.006$ in WM lesions; year 2: $p=0.008$ in NABT) and compared with healthy control subjects (year 1: $p=0.05$ and year 2: $p=0.007$ in NABT). The larger volume within NABT undergoing decreases in VWMTR was detected in multiple sclerosis patients compared with healthy controls ($p<0.001$), and in the IFNβ-1a group compared with the natalizumab group (year 1: $p=0.05$; year 2: $p=0.002$). One patient on natalizumab died from progressive multifocal leukoencephalopathy eight months after completing the study.

Conclusion: Natalizumab may promote remyelination and stabilize demyelination in lesions and NABT in relapsing multiple sclerosis, compared with intramuscular IFNβ-1a.

Introduction

The underlying pathology of multiple sclerosis (MS) is characterized by alternating cycles of inflammation associated with demyelination that manifests both as focal lesions and more diffusely in otherwise normal appearing brain tissue (NABT), followed by remyelination.[1,2] Magnetization transfer imaging (MTI) may be used to detect variations in myelin content within MS lesions and NABT. Decrease in the magnetization transfer ratio (MTR) has been reported to correlate with clinical decline assessed by the Expanded Disability Status Scale (EDSS), MS Functional Composite (MSFC), and other neuropsychological tests.[3-6] In addition, it is a non-specific sign of demyelination, macrophage infiltration and axonal damage.[7] Increased MTR has been shown to be related to remyelination in animal[7] and post-mortem studies.[8,9]

[1]Buffalo Neuroimaging Analysis Center, State University of New York, Buffalo, USA.
[2]The Jacobs Neurological Institute, Department of Neurology, University at Buffalo, State University of New York, Buffalo, USA.

Corresponding author:
Robert Zivadinov, MD, PhD, FAAN, Buffalo Neuroimaging Analysis Center, Jacobs Neurological Institute, Department of Neurology, University at Buffalo, State University of New York, E-2 Buffalo General Hospital, 100 High Street, Buffalo, NY 14203, USA.
Email: rzivadinov@bnac.net

84

Iron Deposition on SWI-Filtered Phase in the Subcortical Deep Gray Matter of Patients with Clinically Isolated Syndrome May Precede Structure-Specific Atrophy

ORIGINAL RESEARCH

J. Hagemeier
B. Weinstock-Guttman
N. Bergsland
M. Heininen-Brown
E. Carl
C. Kennedy
C. Magnano
D. Hojnacki
M.G. Dwyer
R. Zivadinov

BACKGROUND AND PURPOSE: Increasing evidence suggests that iron deposition is present in the later stages of MS. In this study we examined abnormal phase values, indicative of increased iron content on SWI-filtered phase images of the SDGM in CIS patients and HC. We also examined the association of abnormal phase with conventional MR imaging outcomes at first clinical onset.

MATERIALS AND METHODS: Forty-two patients with CIS (31 female, 11 male) and 65 age and sex-matched HC (41 female, 24 male) were scanned on a 3T scanner. Mean age was 40.1 (SD = 10.4) years in patients with CIS, and 42.8 (SD = 14) years in HC, while mean disease duration was 1.2 years (SD = 1.3) in patients with CIS. MP-APT, NPTV, and normalized volume measurements were derived for all SDGM structures. Parametric and nonparametric group-wise comparisons were performed, and associations were determined with other MR imaging metrics.

RESULTS: Patients with CIS had significantly increased MP-APT ($P = .029$) and MP-APT volume ($P = .045$) in the pulvinar nucleus of the thalamus compared with HC. Furthermore, the putamen ($P = .004$), caudate ($P = .035$), and total SDGM ($P = .048$) displayed significant increases in MP-APT volume, while MP-APT was also significantly increased in the putamen ($P = .029$). No global or regional volumetric MR imaging differences were found between the study groups. Significant correlations were observed between increased MP-APT volumes of total SDGM, caudate, thalamus, hippocampus, and substantia nigra with white matter atrophy and increased T2 lesion volume ($P < .05$).

CONCLUSION: Patients with CIS showed significantly increased content and volume of iron, as determined by abnormal SWI-phase measurement, in the various SDGM structures, suggesting that iron deposition may precede structure-specific atrophy.

ABBREVIATIONS: CIS = clinically isolated syndrome; EDSS = Expanded Disability Status Scale; ETL = echo-train length; FIRST = fMRI-integrated registration and segmentation tool; Gd = gadolinium; GM = gray matter; HC = healthy controls; LV = lesion volume; MP-APT = mean phase of the abnormal phase tissue; NBV = normalized brain volume; NGMV = normalized gray matter volume; NLVV = normalized lateral ventricle volume; NPTV = normal phase tissue volume; NWMV = normalized white matter volume; pFOV = phase FOV; RRMS = relapsing-remitting MS; SDGM = subcortical deep GM

It has become increasingly clear that GM damage is present in patients with MS and involves both the cortical GM and SDGM structures.[1] GM damage is most extensive in patients with progressive MS,[2] but is already present in those with CIS[3-10] and can predict conversion to clinically definite MS.[11-13]

Most authors investigating iron deposition in MS have used imaging techniques such as T2 hypointensity,[14-16] relaxometry,[17-19] magnetic field correlation,[20] and SWI.[21-23] Accumulation of iron in the SDGM has also been observed histologically in MS,[24,25] and iron deposition has been detected in the same SDGM structures where the volume loss also occurred.[23] It is thought that excessive levels of iron in the brain may lead to tissue damage by the generation of reactive oxygen species, most probably through the Fenton reaction.[26,27] In this process, ferrous iron (Fe^{2+}) donates an electron to hydrogen peroxide, with the resulting toxic products being ferric iron (Fe^{3+}), hydroxyl anion (OH^-), and the highly reactive hydroxyl radical (OH).

Few studies have investigated iron deposition in the earliest stages of MS, such as in those with CIS.[28] Ceccarelli et al[29] reported increased T2 hypointensity in the head of the left caudate nucleus of patients with CIS compared with HC. However, a recent study using relaxometry did not show differences between patients with CIS and HC.[19] Moreover, patients with RR-MS appear to show even higher iron content in the SDGM than patients with CIS, suggesting that disease progression may be related to pronounced iron deposition.[18,19,30] If increased levels of iron are causally related to SDGM damage and disease development, iron deposition must be present at the earliest stages of the disease, such as in patients with CIS.

This study used a SWI-filtered phase approach to characterize possible iron deposition in the SDGM of patients with CIS. Paramagnetic substances within the brain, mostly in the

ORIGINAL RESEARCH

N. Bergsland
D. Horakova
M.G. Dwyer
O. Dolezal
Z.K. Seidl
M. Vaneckova
J. Krasensky
E. Havrdova
R. Zivadinov

Subcortical and Cortical Gray Matter Atrophy in a Large Sample of Patients with Clinically Isolated Syndrome and Early Relapsing-Remitting Multiple Sclerosis

BACKGROUND AND PURPOSE: Recent studies have shown that selective regional, but not global, GM atrophy occurs from clinical onset to conversion to clinically definite MS. Our aim was to investigate the difference in the extent of SDGM and cortical atrophy in a large sample of patients with CIS and early RRMS and to explore the relationship between SDGM and cortical atrophy and other MR imaging and clinical outcomes.

MATERIALS AND METHODS: Two hundred twelve patients with CIS recruited at the first clinical event (mean age, 29.3 years; median EDSS, 1.5; median disease duration, 3 months) and 177 patients with early RRMS (mean age, 30.7 years; median EDSS, 2.0; median disease duration, 47 months) were imaged on a 1.5T scanner by using a high-resolution 3D T1 spoiled gradient-recalled sequence. Volumetric data for SDGM structures were obtained by using FSL FIRST, while whole-brain, GM, white matter, cortical, and lateral ventricle volumes were estimated by using SIENAX software. Comparisons between the groups were adjusted for age and sex.

RESULTS: Patients with early RRMS showed significantly lower SDGM but not cortical volumes compared with patients with CIS. The most apparent SDGM differences were evident in the caudate and thalamus ($P < .0001$), total SDGM ($P = .0001$), and globus pallidus ($P = .01$). Patients with CIS with a median T2 lesion volume >4.49 mL showed lower total SDGM, caudate, thalamus ($P < .001$), globus pallidus ($P = .007$), hippocampus ($P = .004$), and putamen ($P = .01$) volumes and higher lateral ventricle volume ($P = .001$) than those with a median T2 lesion volume <4.49 mL. Decreased thalamic volume showed the most consistent relationship with MR imaging outcomes ($P < .0001$) in patients with CIS.

CONCLUSIONS: Significant SDGM, but not cortical, atrophy develops during the first 4 years of the RRMS. GM atrophy is relevant for disease progression from the earliest clinical stages.

ABBREVIATIONS: ASA = Avonex-Steroid-Azathioprine; CIS = clinically isolated syndrome; EDSS = Expanded Disease Status Scale; FSL = FMRIB Software Library; GM = gray matter; NBV = normalized brain volume; NCV = normalized cortical volume; NGMV = normalized gray matter volume; NLVV = normalized lateral ventricle volume; NWMV = normalized white matter volume; RRMS = relapsing remitting MS; SDGM = subcortical deep gray matter; SET = Study of Early Interferon β 1a Treatment in High Risk Subjects after CIS

MR imaging is a vital tool enabling clinicians to diagnose, monitor, and predict the progression of MS. Conventional MR imaging has proved to be very sensitive for detecting focal changes in the WM, yet it is relatively insensitive to involvement of the GM in MS. GM pathology in MS is quite different from that in WM, with only mild blood-brain barrier disruption and little-to-no inflammation due to minimal T-cell infiltration.[1] Hence, the difference between lesion and

normal GM relaxation times on MR imaging is less than that seen between lesion and normal white matter.[2]

In the past decade, continuous effort has been made to develop novel MR imaging techniques that are able to quantitatively measure a wide spectrum of GM pathology.[1-3] The introduction of double inversion recovery sequences has advanced the ability to capture part of GM pathology in vivo,[4] and measurement of GM atrophy has become one of the important outcomes in MS studies.[5-7] Other nonconventional MR imaging techniques have also contributed to determining the extent of GM damage in patients with MS.[2] The role of GM pathology in MS has come under increasingly close scrutiny, especially after it was shown that GM atrophy predicts clinical outcomes better than WM damage.[5-7]

More recently, research has focused on determining the extent of GM pathology at the first clinical event in patients presenting with CIS[8-16] or on its evolution with conversion to clinically definite MS.[17-20] It has been reported that global GM volume measures are not sensitive enough to detect GM atrophy at the time of the initial attack.[11] Consequently, GM atrophy studies in patients with CIS have

Research Article

Bimonthly Evolution of Cortical Atrophy in Early Relapsing-Remitting Multiple Sclerosis over 2 Years: A Longitudinal Study

Robert Zivadinov,[1] Carmen Tekwe,[2] Niels Bergsland,[1] Ondrej Dolezal,[1,3] Eva Havrdova,[3] Jan Krasensky,[4] Michael G. Dwyer,[1] Zdeněk Seidl,[4] Deepa P. Ramasamy,[1] Manuela Vaneckova,[4] and Dana Horakova[3]

[1] *Department of Neurology, Buffalo Neuroimaging Analysis Center, The Jacobs Neurological Institute, Buffalo, NY 14203, USA*

[2] *Department of Biostatistics, Buffalo Neuroimaging Analysis Center, Buffalo, NY 14203, USA*

[3] *Department of Neurology and Center of Clinical Neuroscience, First Faculty of Medicine and General Teaching Hospital, Charles University in Prague, 121 08 Prague, Czech Republic*

[4] *Department of Radiology, First Faculty of Medicine and General Teaching Hospital, Charles University in Prague, 121 08 Prague, Czech Republic*

Correspondence should be addressed to Robert Zivadinov; rzivadinov@bnac.net

Received 2 May 2012; Revised 1 December 2012; Accepted 15 December 2012

Academic Editor: Caterina Mainero

We investigated the evolution of cortical atrophy in patients with early relapsing-remitting (RR) multiple sclerosis (MS) and its association with lesion volume (LV) accumulation and disability progression. 136 of 181 RRMS patients who participated in the Avonex-Steroids-Azathioprine study were assessed bimonthly for clinical and MRI outcomes over 2 years. MS patients with disease duration (DD) at baseline of ≤24 months were classified in the early group (DD of 1.2 years, $n = 37$), while patients with DD > 24 months were classified in the late group (DD of 7.1 years, $n = 99$). Mixed effect model analysis was used to investigate the associations. Significant changes in whole brain volume (WBV) ($P < 0.00$), cortical volume (CV) ($P < 0.00$), and in T2-LV ($P < 0.00$) were detected. No significant MRI percent change differences were detected between early and late DD groups over 2 years, except for increased T2-LV accumulation between baseline and year 2 in the early DD group ($P < 0.0$). No significant associations were found between changes in T2-LV and CV over the followup. Change in CV was related to the disability progression over the 2 years, after adjusting for DD ($P = 0.01$). Significant cortical atrophy, independent of T2-LV accumulation, occurs in early RRMS over 2 years, and it is associated with the disability progression.

1. Introduction

Multiple sclerosis (MS) is an autoimmune disease of the central nervous system (CNS) that affects both white matter (WM) and gray matter (GM).

In the last decade, there has been increased interest in studying GM damage in MS, especially in the cortical regions [1]. Advances in both MRI acquisition and analysis techniques have enabled better detection of changes in GM morphology [2]. The MRI assessments included measurements of cortical atrophy, cortical thinning, and cortical lesions [1–4].

Because imaging techniques are still unable to adequately detect GM lesions, especially in the cortex [5–9], measurement of cortical atrophy is gaining increasing attention in the literature [1], in order to assess the real extent of cortical pathology in vivo in patients with MS [2].

Recent studies have established that subcortical, but not cortical, atrophy is present at the earliest clinical stages of the disease [10–15]. However, most of these studies had a cross-sectional design, and only a few longitudinal studies investigated possible associations between GM atrophy and clinical outcomes in patients with MS [11, 15–18].

Gray matter SWI-filtered phase and atrophy are linked to disability in MS

Jesper Hagemeier[1], Bianca Weinstock-Guttman[2], Mari Heininen-Brown[1], Guy U. Poloni[1], Niels Bergsland[1], Claudiu Schirda[1], Christopher R. Magnano[1], Cheryl Kennedy[1], Ellen Carl[1], Michael G. Dwyer[1], Alireza Minagar[3], Robert Zivadinov[1,2]

[1]Buffalo Neuroimaging Analysis Center, University at Buffalo, Buffalo, NY, USA [2]The Jacobs Neurological Institute, University at Buffalo, Buffalo, NY, USA, [3]Department of Neurology, Louisiana State University Health Sciences Center, Shreveport, LA

TABLE OF CONTENTS

1. ABSTRACT

The association between clinical outcomes and abnormal susceptibility-weighted imaging (SWI)-filtered phase, indicative of increased iron content, as well as atrophy, was investigated in the subcortical deep-gray matter (SDGM) of multiple sclerosis (MS) patients. 149 relapsing-remitting (RR) and 61 secondary-progressive (SP) MS patients underwent SWI on a 3T scanner. Mean phase of the abnormal phase tissue (MP-APT) and normalized volumes were determined for the total and region-specific SDGM structures. In an age- and gender-adjusted regression model, total SDGM volume was the strongest predictor of Expanded Disability Status Scale (EDSS) (beta = -.224, $p <$.001), followed by total SDGM MP-APT (beta = -.168, $p <$.019). This model accounted for 30.4% of the variance in EDSS. Only SDGM MP-APT added additional variance in predicting EDSS, compared to conventional MRI metrics. Caudate and red nucleus MP-APT and amygdala volume were associated with EDSS. Our findings suggest that disability in MS patients is associated better with SDGM pathology, as indicated by increased iron content and atrophy, than with lesion burden or white matter and cortical volumes

2. INTRODUCTION

Although multiple sclerosis (MS) is considered historically as a white matter (WM) disease, abnormalities in the gray matter (GM) are consistently reported in both the cerebral cortex and subcortical deep GM (SDGM) brain structures. (1) Mounting evidence suggests that GM pathology may play a more important role in predicting clinical outcomes in MS patients than WM damage. (2, 3)

Over the last decade, efforts have been made to develop novel MRI techniques that are able to quantitatively measure a wide spectrum of GM pathology. Histopathological and MRI studies have found increased iron deposition in the SDGM of individuals with any of several neurodegenerative disorders, including MS. (4-8) Studies have shown a relationship between increased level of iron content and clinical progression, cognitive impairment, and brain atrophy in MS patients. (5, 9-12) The underlying pathological mechanisms of iron deposition in MS patients are unknown; however, it is thought that iron may be derived from myelin/oligodendrocyte debris, destroyed macrophages, or it can be the product of hemorrhaging from damaged brain vessels. (8, 13, 14)

Improved longitudinal gray and white matter atrophy assessment
via application of a 4-dimensional hidden Markov random field model

Michael G. Dwyer, Niels Bergsland, Robert Zivadinov

Buffalo Neuroimaging Analysis Center, Department of Neurology, University at Buffalo, Buffalo, NY, USA

Abstract:

SIENA and similar techniques have demonstrated the *utility of performing "direct" measurements as opposed* to post-hoc comparison of cross-sectional data for the measurement of whole brain (WB) atrophy over time. However, gray matter (GM) and white matter (WM) atrophy are now widely recognized as important components of neurological disease progression, and are being actively evaluated as secondary endpoints in clinical trials. Direct measures of GM/WM change with advantages similar to SIENA have been lacking. We created a robust and easily-implemented method for direct longitudinal analysis of GM/WM atrophy, SIENAX multi-time-point (SIENAX-MTP). We built on the basic halfway-registration and mask composition components of SIENA to improve the raw output of FMRIB's FAST tissue segmentation tool. In addition, we created LFAST, a modified version of FAST incorporating a 4th dimension in its hidden Markov random field model in order to directly represent time. The method was validated by scan-rescan, simulation, comparison with SIENA, and two clinical effect size comparisons. All validation approaches demonstrated improved longitudinal precision with the proposed SIENAX-MTP method compared to SIENAX. For GM, simulation showed better correlation with experimental volume changes (r=0.99 vs. 0.75), scan-rescan showed lower absolute deviations (1.5% vs. 2.5%), correlation with SIENA was more robust (r=0.70 vs 0.53), and effect sizes were improved by up to 68%. Statistical power estimates indicated a potential drop of 55% in the number of subjects required to detect the same treatment effect with SIENAX-MTP vs. SIENAX. The proposed direct GM/WM method significantly improves on the standard SIENAX technique, and may provide more precise data and additional statistical power in longitudinal studies.

1. Introduction:

Brain atrophy measurement has become a key analysis in basic neuroimaging science, aging research, and research into pathologic conditions including multiple sclerosis (MS) (Zivadinov and Bakshi, 2004; Bermel and Bakshi, 2006), Alzheimer's disease (AD) (Sluimer et al., 2008), and Parkinson's disease (Burton et al., 2005). It is also becoming an important component of modern MS and AD clinical trials (Thal et al., 2006; Zivadinov et al., 2008). Readily available segmentation tools such as FMRIB's Structural Image Evaluation, using Normalisation, of Atrophy (SIENA) (Smith et al., 2001) allow atrophy measurements to be both reliable (Sormani et al., 2004) and highly standardized across studies and research centers (Jasperse et al., 2007).

A better understanding of the specific mechanisms of atrophy has led many researchers to focus on separate quantification of gray matter (GM) and white matter (WM) atrophy (Karas et al., 2003; Thompson et al., 2003; Chard et al., 2004; Sanfilipo et al., 2006; Zivadinov and Minagar, 2009; Hulst and Geurts, 2011; Zivadinov and Pirko, 2012) Although a variety of measurement approaches are available (Chen et al., 2004; Nakamura and Fisher, 2009; Derakhshan et al., 2010; Nakamura et al., 2011), one used by a number of groups is to perform independent tissue segmentations at both baseline and follow-up time points (potentially with some spatial normalization), and then calculate the changes via simple subtraction of the relevant total volumes (Oreja-Guevara et al., 2005; Valsasina et al., 2005; Healy et al., 2009; Horakova et al., 2009). Unfortunately, although this approach is straightforward, intuitive, and easily implemented, it is considerably less reproducible than a direct measurement like SIENA. In fact, even whole brain measures from SIENAX (the cross-sectional variant of SIENA) (Smith et al., 2002) are less reproducible than SIENA change measures (Cover et al., 2011), and GM/WM-specific measures are even more difficult. This reduction in precision can have serious consequences for the statistical power of

planned studies, resulting in the need for either very large subject groups or in the inability to detect real changes (Anderson et al., 2007; Healy et al., 2009).

There are a number of potential reasons that reproducibility issues arise from this approach. First, and fundamentally, since two independent cross-sectional measures are used rather than a single direct measurement, there are two sources of measurement error. Without a direct comparison, it is likely that the two segmentations will come to slightly different conclusions about the precise spatial and intensity distributions of the tissue classes involved. In particular, voxels of relatively ambiguous intensity (i.e., halfway between GM and WM) will often be classified differently at baseline and follow-up, despite a lack of change in actual tissue morphometry. Although these differences may ultimately cancel out in aggregate, they add to the overall variance of the measurement. This problem is also aggravated by the fact that the GM/WM border (as seen on conventional MRI) is generally not as clearly defined as the brain/cerebrospinal fluid (CSF) border, and the absolute intensity contrast between the two tissues is usually considerably lower than between GM and CSF.

Second, scanner drift and differences in positioning can lead to minor geometric distortions in the acquired images that can change volumetric measurements (Freeborough, 1996). Even when subtle, these changes can dwarf small clinical changes that are the target of studies and clinical trials. Although their nonlinear nature can make them challenging to completely correct (Caramanos et al., 2010), they can be at least somewhat ameliorated by improved co-registration with full affine parameters.

Third, brain extraction can have a significant effect on measured tissue volumes (Battaglini et al., 2008; Keihaninejad et al., 2010; Leung et al., 2011; Popescu et al., 2012), so inconsistent brain extraction at baseline and follow-up can lead to tissue volume changes that do not reflect actual atrophy.

To address these issues, we developed a novel technique that augments FMRIB's Automated Segmentation Tool (FAST) algorithm with a 4-dimensional hidden Markov random field (HMRF)

to ensure more consistent classification. In the current version of FAST, a 3-dimensional HMRF is used to impose local spatial constraints on the segmentation process. Essentially, this HMRF penalizes discrepancies in tissue classification for isolated voxels, but allows for contiguous areas of change (Zhang et al., 2001). So, for example, a shift in the WM/GM border when moving from slice to slice is not significantly penalized. By extending the basic spatial model to a full spatio-temporal (4-dimensional) model, we hoped to allow for the same sort of shift in border over the time dimension – atrophy/growth – instead of the slice direction dimension (or any other spatial dimension), while simultaneously penalizing small, localized discontinuities.

Additionally, we incorporated two of the key elements of SIENA into our tissue-specific analysis technique: skull-constrained halfway-space co-registration to address positioning and scanner drift issues, and uniform brain extraction to reduce extraction-related variance.

2. Materials and methods:

2.1 Inclusion of a temporal component in the HMRF:

The HMRF framework employed by FAST (Zhang et al., 2001) is designed to use spatial neighborhood information to elegantly mitigate the noise and homogeneity problems inherent in MRI-based tissue segmentation. Intuitively, when assigning tissue class labels to voxels, the algorithm attempts to minimize a global cost function that penalizes two separate elements: selection of class labels whose mean intensities do not match well with the labeled voxels (e.g., labeling a relatively bright voxel GM on a T1-weighted image), and creation of spatially isolated labels (e.g., a single GM voxel completely surrounded by WM). It is the tension and balance between these intensity-matching and spatial-homogeneity-preserving goals that allows for the quality of tissue segmentations achievable with FAST.

More rigorously, the class labeling prior probability in standard FAST is updated iteratively according to

$$P^{(t)}(l|y_i) \propto \exp \beta w_{i,l} - \left(\log(\sqrt{2\pi\sigma_l^2}) + \frac{(y_i - \mu_l)^2}{2\sigma_l^2} \right)$$

Where t is the iteration, l is the label of a specific class, i is the voxel index, β is the neighborhood weighting factor, w is a neighborhood weighting function, and the entire second term is the standard log Gaussian penalizing deviations of the voxel intensity from the proposed class mean. The proportionality rather than equivalence reflects the fact that FAST allows for the incorporation of spatial prior probabilities, and also for the fact that the probabilities are normalized. The key component in this context is the weighting function w, which expresses the overall a-priori probability of finding a three-dimensional class configuration where voxel i's class is l, given its already-classified neighbors. Internally, FAST calculates this term via a conversion to "MRF weights".

From a local perspective, the total label weight for a voxel i *i* and label *l* l is determined by iterating over its neighbors j in neighborhood N (where N is the neighborhood of the up to 26 voxels surrounding voxel i), and is calculated as:

$$w_{i,l} = \sum_{j \in N_i} \frac{1}{d(i,j)} \cdot p^{(t)}(l|y_j)$$

where d(i,j) is the distance between voxels *i* i and j (diagonal neighbors are farther than horizontal neighbors, and the slice resolution is often lower than the in-plane resolution), and the right side is the current iteration's a posteriori probability of classification l for voxel j.

However, it is important to note that the mathematics behind the general HMRF model are not limited to the usual 3 spatial dimensions (Winkler, 2003), and in fact the implementation in FAST can be very naturally extended from 3 dimensions to 4 with minimal modification. As noted above, tissue atrophy or growth can be considered as a change in the border between tissues or between tissue and CSF when moving along the time dimension, and is analogous to the shifts that occur when moving from slice to slice. The main distinction is that for a two-point tissue change analysis, each voxel has only one temporal neighbor whereas there are usually 26 spatial neighbors.

Recognizing this, we modified the above calculation to be:

$$w_{i,l} = \sum_{j \in N_i} \frac{1}{d(i,j)} \cdot p^{(t)}(l|y_j) + z p^{(t-1)}(l|y_j')$$

where the second term is the prior iteration a posteriori probability of classification l for voxel i in the same physical position as i but at the other time-point, multiplied by a factor z. Thus, this calculation now includes a regularization term for maintaining constant voxel classification over time as well as for the usual agreement with spatially neighboring classifications. The z factor is a weighting coefficient to control the importance of temporal consistency in the model, and can be specified either as a constant or as a constant divided by the length of time between scans.

We implemented the above scheme, called LFAST (for longitudinal FAST), by modifying FAST to be multithreaded, and having the two segmentations proceed in parallel. We used a producer/consumer semaphore system to cause each thread of FAST to operate in symmetric lock-step, moving on to the next iteration only when the total cost function could be calculated by each side comparison. Furthermore, this was implemented in an unbiased way such that neither time point was processed "first" or influenced the other time point in a non-reciprocal manner.

2.2 Skull-constrained co-registration:
To avoid potential scanner- and position-related scaling issues, we employed a skull-constrained co-registration approach similar to that used in SIENA (Smith et al., 2001). Baseline and follow-up images were individually de-skulled using BET (Smith, 2002), and both brain and skull images were retained. A skull-constrained 12-parameter affine registration was then performed with FLIRT in each direction. The resulting transforms were then manipulated to determine unbiased halfway transforms for each image into a space directly between the two. Both images were resampled into this space using tri-linear interpolation before proceeding with analysis.

2.3 Uniform brain extraction:

Another potentially confounding factor in longitudinal measurement is differing brain extractions. Because we transformed both images into a common halfway space, we were also able to combine the individual brain masks via a union operation. First, to prevent scanning volume differences from being reflected as atrophy, we created a joint mask of valid voxels. Briefly, we took each time point's input image in native space and coded all voxels as 1. Then, we transformed these from each time-point into the halfway space, padding with zeros (the default for FLIRT). This resulted in maps of which voxels were actually present in the original scanned volumes, and by taking the intersection we were able to create a joint map of voxels present in both original scanning volumes. Second, we transformed both time points' original brain masks into the common space, took the union of these, and then masked the result with the joint valid voxels mask. In the end, this left us with a mask of voxels present in both raw scanning volumes and included in either brain mask. We used this mask to define the voxels for our modified FAST to classify.

2.4 Complete pipeline:

A schematic view of the combination of these elements into a single processing pipeline called SIENAX-MTP (for multiple time points) is given in Figure 1. Note that both modified FAST components run in parallel. The final output was similar to the results of running two individual SIENAX analyses, including standard partial volume estimate (PVE) maps. For all subsequent analysis, cross-sectional volumes and longitudinal changes were calculated based on summation of the individual voxel-wise tissue type PVE estimates.

2.5 Validation:

To evaluate the characteristics of our proposed approach, we used simulation, scan-rescan, and testing on a real clinical dataset of patients with MS and matched healthy controls (HCs).

2.5.1 Simulation

To better understand the impact of the inclusion of the temporal factor in LFAST, as well as to empirically determine the optimal z weighting value,

we created a semi-realistic controlled simulation framework (Fig 2). Balancing simplicity against the need for a reasonably complex shape with gyrus-like folds, we opted to use a set of concentric spherical shells distorted by periodic polar coordinate sin^2 waves in the x-y plane. The model consisted of five such shells, representing (from the inside out) ventricular CSF, deep GM, white matter, peripheral GM, and sulcal CSF. Intensities were set to correspond to a conventional T1-weighted image, with ventricular and sulcal CSF set to be dark, peripheral GM set to be intermediate, and WM set to be bright. Deep GM was set to be midway between peripheral GM and WM, as is often the case.

Given this model, peripheral and deep GM were iteratively thinned to yield 13 sample volumes with varying levels of simulated atrophy. To simulate partial volume effect, the sample volume were created in a high-resolution space (512x512x256) and then down-sampled. The down-sampled images were subsequently corrupted by randomly generated multiplicative bias field based on Legendre polynomials (Styner and Brechbuhler, 2000). Finally, the images were additionally corrupted by Gaussian random noise to yield a signal to noise ratio of approximately 15.

To evaluate LFAST's performance and to compare it against FAST, we ran FAST from baseline to each of the subsequent 12 simulated images. We then ran LFAST in the same way, and repeated with 10 different z factors ranging from 0 to 256. Finally, we repeated this entire process 5 times using different random seeds for noise and bias field calculation.

2.5.2 Scan-rescan:

For scan-rescan, we acquired a dataset of 5 HC, all scanned twice within a maximum of one-week. Axial 3D-SPGR T1-WI scans were acquired with field of view (FOV) 25.6x25.6 cm, matrix 256 x 256, percent phase FOV 0.75, 1mm thickness (th), 184 slices, no gap, TE 2.8 ms, TR 5.9 ms, NEX 1, flip angle 10° on a 3-tesla Signa MRI scanner (General Electric, Milwaukee, WI, USA). We then performed brain/skull extractions on all images using BET and manually correcting where appropriate using fslview (less than 5% of cases). We made simple modifications to the SIENAX and SIENA scripts to allow for the input of brain and

skull images. Subsequently, we ran the above-described SIENAX-MTP technique with a z factor of 32, traditional SIENAX (version 2.6), and SIENA (version 2.6) on each pair of images and masks. Based on the assumption that no real atrophy should occur within HCs during this time period, we compared the resulting outcomes to the expected value of 0% change. To ensure that opposite errors did not cancel, we considered the absolute value of each scan pair's observed deviation from the expected 0%.

2.5.3 Clinical evaluation:

Two clinical cohorts were used for clinical validation of the WB, GM and WM volume change measures over a variable period of follow-up. The first cohort included 67 relapsing-remitting (RR) MS patients and 34 age- and sex-matched HCs who obtained their follow-up scan 2 years from baseline. This cohort was used to test differences in WB, GM and WM volume change measures between MS patients and HC. The second cohort included 64 RRMS patients who did not develop disability progression and 64 patients who did develop disability progression after 5 years of follow-up, and received MRI scans at baseline and 5 years. This cohort was used to test differences in WB, GM and WM volume estimation change measures between MS patients with or without disability progression after 5 years. Both cohorts were randomly selected from a larger cohort of RRMS patients and HC participating in a 2- and 5-year follow-up study in our center. The two cohorts did not overlap with respect to participating subjects. Inclusion criteria for MS patients were RR disease course (Lublin and Reingold, 1996), age 18-65 years, and Expanded Disability Status Scale (EDSS) (Kurtzke, 1983) between 0 and 5.5. All HC subjects fulfilled the health screen questionnaire requirements containing information regarding medical history (illnesses, surgeries, vascular and environmental risk factors, medications) and health screen requirements on physical and neurologic examination. Patients with a relapse and/or steroid treatment in the 30 days preceding study entry and pre-existing medical conditions known to be associated with brain pathology e.g., neurodegenerative disorder, cerebrovascular disease, positive history of alcohol dependence, traumatic brain injury with loss of consciousness > 5 min, cognitive impairment, history of psychiatric disorders and seizures) or

pregnant subjects were excluded. The original follow-up studies from which the subjects were randomly selected were approved by an internal Institutional Review Board and written informed consent was obtained from all subjects. All image analysis was blinded to patients' demographic, clinical and disease group characteristics.

2.5.4 Clinical MRI acquisition:

MRI assessments were performed at baseline, 2 years (34 HC and 67 RRMS patients), and 5 years (128 RRMS patients) using the same 1.5-tesla Signa unit (General Electric, Milwaukee, WI, USA). For each session, 3D high resolution T1-weighted imaging (WI) using a spoiled gradient echo (SPGR), 2D multiplanar dual spin-echo (SE) proton density (PD) and T2-WI and Fluid Attenuated Inversion Recovery (FLAIR) were obtained using the same protocol between the various time points.

The axial 3D-SPGR T1-WI scans were acquired with FOV 25x25 cm, matrix 256 x 256, 1.5mm thickness (th), 128 slices, no gap, TE 7 ms, TR 24 ms, NEX 1, flip angle 30°. The axial dual SE sequence was acquired with TE 12/90 ms, TR 3000 ms, NEX 1, ETL 14, FOV 25x25 cm, matrix 256 x 256, 3mm slice th with a total of 48 slices, no gap, whereas axial FLAIR with FOV 25x25 cm, matrix 256 x 256, 48 slices, 3 mm th, no gap, TE 128 ms, TI 2000 ms, TR 8002 ms, ETL 22, NEX 1. Patients and controls were positioned in the magnet according to commonly accepted international guidelines.

2.5.5 Clinical MRI analysis: Lesion analysis: T2-lesion volumes (LV) were calculated using a semi-automated edge-detection contouring-thresholding technique, as previously described (Zivadinov et al., 2001). The lesion maps were then used to inpaint the 3D-SPGR T1-WI scans via an in-house dilation-based method (Zivadinov et al., 2012).

Volume change analysis: For each subject's pair of inpainted 3D-SPGR T1-WI scans (baseline and follow-up), we performed three analyses. SIENAX (version 2.6) was used on each individual T1-weighted 3D SPGR image. Normalized WB volume, normalized GM volume, and normalized WM volume were measured as previously described

(Zivadinov et al., 2012). For longitudinal changes of the WB volume, we applied the SIENA 2.6 method (Smith et al., 2002) to calculate the percentage brain volume change (PBVC). In addition, SIENAX and SIENAX-MTP approaches were used to estimate WB volume changes. To better understand the specific contribution of the LFAST component, we also ran conventional SIENAX on registered halfway-space images with unified brain masks (SIENAX-HW).

2.6 Quality control

All analyses were reviewed by a trained operator (MD/NB) at all critical points, and either corrected if possible or excluded from further analysis. Brain and skull extraction errors were corrected manually or by adjusting appropriate BET parameters. For all analyses (SIENA, SIENAX, SIENAX-HW, and SIENAX-MTP), the same corrected brain extractions were used. In addition, we also marked as failures any longitudinal analysis pairs producing biologically implausible percent changes in GM, WM, or WB. Although the precise value of biological implausibility is not known and may vary by disease, for the work described here we adopted a standardized cutoff of 5% per year.

2.7 Statistical analyses:

All data analyses were performed using SPSS version 16.0 (SPSS, Inc., Chicago IL). For simulation, correspondence with experimental volume change was evaluated using Pearson product-moment correlation. A Fisher r-to-z transform was then used to compare the correlation coefficients. For scan-rescan, differences in mean absolute PBVC, percent GM volume change (PGMVC), and percent WM volume change (PWMVC) between measurements were evaluated using Student's t-test for paired samples. Two-year comparisons between HC and MS groups were performed using the Pearson chi-squared test for categorical values and Student's t-test for continuous variables (percent change from baseline). The 5-year MS cohort was selected based on disability progression at the 5-year of follow-up. The first group contained patients who, at the end of the 5-year period, presented with disability progression. This was defined as a ≥1.0-point increase in EDSS score in patients who had a baseline EDSS score of

≥1.0, or a ≥1.5-point increase in EDSS score in patients who had a baseline EDSS score of 0.0. The second group contained patients without disability progression. Two-tailed comparisons between these two groups were performed using the Pearson chi-squared test for categorical values and Student's t-test for continuous variables (percent change from baseline). Spearman rank correlation was performed within the cohorts to test associations between different WB, GM and WM volume estimation change measures. Fisher's exact test was also applied to compare the rate of quality control failures for each analysis. The nominal p-value <0.01 was considered significant, while p-value <0.05 was considered a trend, using two-tailed tests. In addition, we used the 5-year dataset to perform basic power calculations to determine required sample sizes to reliably detect GM and WM treatment effects in atrophy reduction of varying degree, as described in (Fox et al., 2000).

3. Results

Our method was able to run successfully on all 234 acquired MRI image pairs in the study with no quality control failures for the SIENAX MTP approach, compared to 12 quality control failures for SIENAX (2 for visual segmentation and 10 for "cut-off" acceptable range values, comparison p < 0.001). Provided inpainted and properly de-skulled images, it took approximately 20 minutes per case on a modern system with 8GB RAM and an Intel Core-i5 2500 CPU. In particular, our multithreaded approach allowed the LFAST component to take only marginally longer than a single original FAST run.

3.1 Simulation results:

Simulation results for representative z factors are shown in Figure 3. For all z factors above 0 and below 128, LFAST agreed better with ground truth than conventional FAST. Additionally, for those z factors, LFAST showed lower variation between runs due to noise and/or bias field. In particular, with a z factor of 32, LFAST showed a correlation of r=0.99 compared to r=0.75 for conventional FAST (p < 0.001), and a mean absolute deviation (MAD) of 1.3% compared to 5.6% for conventional FAST. Median absolute deviations were similar, at 1.2% and 5.5%, respectively. Below a z factor of 128, LFAST did not appear significantly biased toward no change; in fact, conventional FAST appeared to

somewhat overestimate the degree of change for atrophy amounts less than 8%.

3.2 Scan-rescan results:

For the raw scan-rescan cases, SIENAX showed a MAD of 2.5% for GM volume measurements, 3.4% for WM measurements, and 1.5% for WB measurements between acquisitions (averaged over subjects). In contrast, our SIENAX-MTP method had a MAD of 1.5% for GM, 2.4% for WM, and 0.39% for WB (Figure 2). Most likely due to the very small sample size (5 subjects), these individual differences showed only statistical trends. When pooling all measures together, though, results were significant (p < 0.01). For comparison, SIENA had a MAD of 0.36% for WB.

3.3 Brain volume estimation change measure method differences in the examined cohorts:

The mean age of the 67 RRMS patients in the first cohort at baseline was 34.1 ± 7.6 years and 35.0 ± 8.3 years in the 34 HC (p=0.595). There were 57 females in MS and 27 in HC groups (p=0.575). The mean T2-LV at baseline in MS patients was 9.032 ± 13.232 cm^3. The HC did not present T2 lesions.

The mean age of the second cohort of 128 RRMS patients was 36.7 ± 11.1 years in the group without disability progression and 38.5 ± 12.3 years in the disability progression group (p=0.386). There were no significant gender differences between the two groups (0.709). T2-LV was not significantly different at baseline between the 2 groups (15.723 ± 10.9 vs. 12.201 ± 5.9 cm^3, p=0.227). However, there was significantly decreased normalized WB volume at baseline in the 5-year disability progressed group (1489.7 ± 81.9 vs. 1525.8 ± 72.7 cm^3, p=0.003), in normalized WM volume (691.9 ± 47.0 vs. 709.6 ± 36.8 cm^3, p=0.008), and a trend in normalized GM volume (797.8 ± 54.4 vs. 816.3 ± 50.7 cm^3, p=0.018).

The PBVC differed significantly between HC and MS patients over the 2-year period for all 3 methods (p<0.0001), but the highest effect size was detected for SIENAX-MTP PBVC (d=1.59), followed by SIENA, SIENAX-HW and SIENAX PBVCs (Table 1). The PGMVC also differed significantly between HC and MS patients over the

2-year period for SIENAX-MTP and SIENAX PGMVCs, and again the highest effect size was detected with SIENAX-MTP PGMVC (d=0.75, Table 1). Interestingly, the SIENAX-HW difference did not reach significance. The PWMVC was significantly different between HC and MS patients over the 2-year period for both methods, with the highest effect size detected by SIENAX-MTP PWMVC (d=1.15, Table 1). The SIENAX-MTP method also showed the lowest measurement SD for WB, GM and WM percent volume changes, for both HC and RRMS patients.

When the cohort of 128 RRMS patients with or without disability progression over 5 years was compared with regard to PBVCs, all methods yielded significant differences (p<0.0001, Table 2), except SIENAX PBVC. Again, the highest effect size was for SIENAX-MTP PBVC (d=0.81) followed by SIENA and then SIENAX-HW. For PGMVC, the 2 groups differed significantly only with SIENAX-MTP (d=0.57, p=0.002, Table 2). The PWMVC was significantly different between the 2 MS groups for both methods with the highest effect size for SIENAX-MTP PWMVC (d=0.64, p<0.0001, Table 2). Again, the SIENAX-MTP method showed the lowest measurement SD for WB, GM and WM percent volume changes, for both groups of RRMS patients.

In both cohorts, the SIENAX-MTP PBVC showed higher correlation with SIENA PBVC (r=0.70 and r=0.67, respectively) than SIENAX PBVC (r=0.53 and r=0.55) or SIENAX-HW PBVC (r=0.61 and r=0.55).

3.4 Power estimation:

Taking the second cohort MS patients' GM atrophy rates and variances, we found that using SIENAX-MTP to detect a reduction in five-year GM atrophy with a significance level of 0.05 and a power level of 90% would require approximately 55% less subjects than for traditional SIENAX. For example, detecting a 30% reduction would require 131 subjects with SIENAX, compared to only 59 with SIENAX-MTP. Even detecting a more pronounced 50% reduction would require 47 subjects with SIENAX compared to 21 with SIENAX-MTP. When evaluating WM, both methods were less powerful than for GM due to the low amount of observed atrophy in WM, but our

results again indicated that SIENAX-MTP would require 35% fewer subjects than for SIENAX (793 subject vs. 1211 subjects to detect a 50% reduction in atrophy).

4. Discussion

This study describes a novel and efficient technique for evaluating tissue-specific atrophy via longitudinal MRI data. It is a direct and relatively straightforward extension of the widely used SIENAX and SIENA techniques, building upon the software tools used in constructing them (particularly FAST). As such, it retains many of the advantages of those programs while providing improved precision in longitudinal tissue-specific change measurement.

It is difficult to produce a gold standard for MRI-based longitudinal brain atrophy estimation, since per se it requires in vivo measurements. Because of this, many alternative validation approaches have been proposed (Sharma et al., 2009). Here, we have used four such approaches, including experimental simulation, scan-rescan, patient/control separation, and clinical outcome separation. In all four cases, SIENAX-MTP performed better than standard SIENAX. In our simulation, SIENAX-MTP measurements correlated significantly better with experimental volumetric changes with far less case-to-case variance. For scan-rescan, SIENAX-MTP provided a 40% reduction in GM variability, a 29% reduction in WM variability, and a 74% reduction in WB variability (measured by mean absolute deviation). For patient/control separation, effect sizes were consistently larger for SIENAX-MTP than SIENAX, and were even marginally better than SIENA in this particular group. For clinical progression separation, SIENAX-MTP effect sizes were also consistently higher in all tissue measures, and significant GM volume change differences between the groups were detected with SIENAX-MTP, while this was not the case with SIENAX.

These improvements in precision were realized through a combination of three variance-reducing techniques – skull constrained co-registration, uniform brain extraction, and a four-dimensional extension to FAST's MRF model. The first two modifications are relatively straightforward, and are already part of the standard SIENA approach. However, the LFAST component of SIENAX-MTP is to our knowledge completely novel. The general HMRF approach allows for controlled Bayesian analysis in the face of contaminated (noisy) data. The existing FAST program uses this to substantially improve its segmentation results by incorporating prior knowledge about the spatial smoothness of brain tissues – put simply, most voxels are the same tissue as their spatial neighbors. Our LFAST approach takes this one step further by incorporating the prior knowledge that, despite atrophy, the majority of voxels do not actually change classification across time points. Furthermore, when atrophy does occur, it is almost never single voxels changing, but rather large groups of voxels along edges. These two properties are highly amenable to HMRF model analysis.

As with FAST, the optimal degree of regularization to use is an important question. It is dependent on a number of factors, including signal to noise ratio (SNR), contrast to noise ratio (CNR), and bias field strength, and is largely empirically derived. For example, the commonly used value for the β factor in FAST is 0.1, based on its experimental success. The arbitrariness of this selection is largely ameliorated by the model's robustness across a wide range of parameter values. To better understand the behavior of our time-specific weighting, z, we analyzed a large amount of simulation data and found that for z values above 0 and below 128, LFAST consistently performed better than FAST across a wide range of atrophy amounts. Based on this, our recommendation is to use a constant z factor of 32 for most datasets. In most studies, the time duration between scans is well matched between subjects, so the z factor can be kept constant. If this is not the case, then the z factor can be weighted as a function of the individual subject's fraction of the average time, according to:

$$z_s = \frac{t_s}{\mu_t} \cdot z$$

Where z_s is the subject-specific z factor, t_s is the time difference in scans for the specific subject, μ_t is the mean time difference across all subjects or scan pairs, and z is the reference z factor (e.g., 32).

One potential confounding issue with our regularization approach is the fact that atrophy may result in some degree of mis-registration between scans, or may produce deformations that cannot be corrected without more error-prone non-linear warping approaches. However, the nature of such mis-registrations and/or deformations implies that the discontinuities will be distributed over many voxels representing a clear edge in 4-dimensional space. Therefore, they will not be significantly penalized by our additional 4-dimensional regularization, as is indicated by our experimental results.

Another important consideration with the novel LFAST component is the degree to which such a regularization-based technique introduces bias into the analysis. In this case, there are three mitigating factors that should be considered. First, this approach builds on an already-existing regularization scheme, simply extending the pre-existing spatial framework to the temporal domain. The pre-existing FAST approach makes the assumption that tissues are generally contiguous while still allowing for non-isolated continuities. Our work extends this to atrophy, incorporating the assumption that it generally occurs over connected groups of voxels. Second, any bias introduced is toward maintaining consistent classification over time. This is both conservative (avoiding type I error) and also more closely mimics the behavior of a human expert in that it avoids classification changes without clear intensity differences. Third, and perhaps most important in practice, the actual degree of regularization as specified is very small compared to the contribution of the intensity models, serving primarily to remove ambiguity due to noise or intermediate intensities and having little effect on areas of strong intensity change. Taking these factors into account, and considering that our simulation results do not show a significant bias, it appears unlikely that our modifications have any meaningfully detrimental effect on accuracy to offset the clear gains in precision.

Given the robust results from various validation findings, all pointing in favor of the SIENAX MTP approach, we conclude that our direct GM/WM method is indeed able to meaningfully improve measurement precision. Overall, our power analysis based on actual clinical data shows that these improvements in precision from SIENAX to SIENAX-MTP can result in an up to 55% decrease in the number of subjects required to reliably detect a given treatment effect. This is a substantial difference, indicating that SIENAX-MTP can be used to significantly reduce patient risk and/or costs for tissue-specific studies and clinical trials and/or allow for the detection of more subtle differences for the same cost. It may also make possible more studies from difficult-to-acquire groups (such as rare disease variants). Furthermore, to be as objective as possible we did not take scan failure rate into account (Fox et al., 2000), but this would theoretically result in an additional 5.1% increase in required subjects for SIENAX.

Furthermore, the SIENAX MTP approach was robust to segmentation of images of considerably varying image quality, such as those obtained in multi-center clinical trials. Although the present study only included two different MRI scanners of different field strength to study robustness, the SIENAX MTP method was also recently applied to a large number of MRI scans deriving from multiple scanners in an ongoing multi-center clinical trials (Zivadinov, Bergsland, et al., 2013; Zivadinov, Havrdová, et al., 2013). Preliminary data from these multi-center studies confirm that SIENAX MTP can be readily applied to MRI scans with varying image quality, and results in minimal failures. Moreover, the present study showed that there were 12 SIENAX GM/WM failures on total of 234 MRI pair analyses (5.1%). Therefore, another important advantage of our newly developed direct GM/WM method is a decrease of scan failures due to inadequate GM/WM segmentation and/or biologically implausible percentage changes outside acceptable ranges.

As discussed previously, ours is not the only approach to improvement of longitudinal tissue-specific atrophy measurement (Karas et al., 2003; Thompson et al., 2003; Chen et al., 2004; Battaglini et al., 2009; Nakamura and Fisher, 2009; Derakhshan et al., 2010; Nakamura et al., 2011). However, we believe it has a number of comparative advantages. First, it is a natural extension of already existing methods that have been previously used with success in various studies and clinical trials, and may therefore be both easier for other to implement and have less overall risk while still

significantly reducing measurement errors. Also, it is capable of measuring atrophy in the entire brain, including both deep GM and cortical GM and is not restricted to specific areas as previously reported (Chen et al., 2004). Moreover, unlike voxel-wise group approaches (Douaud et al., 2007; Battaglini et al., 2009), it makes no assumptions about the spatial homogeneity of atrophy between patients – an important concern in highly variable diseases like MS. Deformation-based mapping approaches in general suffer from a similar issue in that they are generally designed to find spatially consistent changes across subjects (Thompson et al., 2004; Studholme and Drapaca, 2006; Pieperhoff et al., 2008). Although subject-specific volume changes can be calculated via techniques such as whole-brain Jacobian integration (Boyes et al., 2006; Cardenas et al., 2007), they are highly dependent on the underlying accuracy of the non-linear normalization technique used (Camara et al., 2008). These techniques are a complex area of active research, can still be relatively failure-prone, and are complicated by the fact that there many potential warp fields can be used to produce the same image. Furthermore, the Jacobian integration technique requires a-priori ROIs so that for accurate GM volumetry, an accurate map of baseline or follow-up GM would be required. A final advantage of SIENAX-MTP is that, just as for SIENAX, it produces subject-specific voxel-wise maps that can be subsequently used for assessing other non-conventional metrics in specific tissue compartments.

4.1 Limitations and future work:

Currently, experiments with SIENAX-MTP have been limited to two time-points. However, the mathematical framework naturally extends to an arbitrary number of scans, and may therefore be useful in frequent serial analyses of the type described in (Valsasina et al., 2005). By using this approach and processing all scans at once in a unified model, significant reductions in the impact of noise may be realizable.

One limitation of the current study is that we did not investigate the effect of slice thickness, which may be another confounding factor. However, we used images of both 1.5mm and 1mm thickness, which we believe represent the current standard for tissue atrophy analysis in most clinical studies. For example, it is comparable to the recommendations of

the Alzheimer's Disease Neuroimaging Initiative (ADNI) (Jack et al., 2010). Furthermore, it seems likely that increases in slice thickness would either affect both methods equally or further argue in favor of SIENAX-MTP (since regularization would become more important when faced with sparser data).

Also, in the current study we used the relatively simple power analysis approach of (Fox et al., 2000) rather than a more accurate but complex modeling technique (Healy et al., 2009). We believe this approach is justified here both because it is more straightforward and because we are primarily concerned with relative power between SIENAX and SIENAX-MTP. Although this may affect the absolute number of patients calculated, it should not have a significant impact on the relative values between SIENAX and SIENAX-MTP.

There were also a few unexpected findings. First, we observed increases in PBVC and PBWMC in HC over two years. It is doubtful that this change was "real" and not artifactual. Many factors can systematically affect brain volume measurements, including scanner drift, subject positioning, hydration and hardware maintenance (Zivadinov et al., 2008). However, SIENAX-MTP provided the closest change to zero in addition to the best differentiation between groups, and the groups were both equally subject to any systematic artifactual change. Second, SIENAX-HW in some cases performed worse than SIENAX on the raw images. It is possible that linear co-registration and/or the smoothing entailed by that co-registration introduces some small errors that in well-controlled scanning environments may outweigh small positioning errors.

5. Conclusions:
The proposed SIENAX-MTP technique can provide significantly improved measurement of GM/WM atrophy over time and significantly fewer scan failures by reducing error-related variance. It is also relatively simple to implement and builds on a widely recognized and thoroughly tested method. The increased statistical power gained with SIENAX-MTP can potentially benefit research studies and clinical trials by detecting changes earlier or in smaller cohorts, or by detecting more subtle changes.

References

Anderson, V.M., Bartlett, J.W., Fox, N.C., Fisniku, L., Miller, D.H., 2007. Detecting treatment effects on brain atrophy in relapsing remitting multiple sclerosis: sample size estimates. Journal of neurology 254, 1588–94.

Battaglini, M., Giorgio, A., Stromillo, M.L., Bartolozzi, M.L., Guidi, L., Federico, A., De Stefano, N., 2009. Voxel-wise assessment of progression of regional brain atrophy in relapsing-remitting multiple sclerosis. Journal of the neurological sciences 282, 55–60.

Battaglini, M., Smith, S.M., Brogi, S., De Stefano, N., 2008. Enhanced brain extraction improves the accuracy of brain atrophy estimation. NeuroImage 40, 583–9.

Bermel, R.A., Bakshi, R., 2006. The measurement and clinical relevance of brain atrophy in multiple sclerosis. Lancet neurology 5, 158–70.

Boyes, R.G., Rueckert, D., Aljabar, P., Whitwell, J., Schott, J.M., Hill, D.L.G., Fox, N.C., 2006. Cerebral atrophy measurements using Jacobian integration: comparison with the boundary shift integral. NeuroImage 32, 159–69.

Burton, E.J., McKeith, I.G., Burn, D.J., O'Brien, J.T., 2005. Brain atrophy rates in Parkinson's disease with and without dementia using serial magnetic resonance imaging. Movement disorders : official journal of the Movement Disorder Society 20, 1571–6.

Camara, O., Schnabel, J.A., Ridgway, G.R., Crum, W.R., Douiri, A., Scahill, R.I., Hill, D.L.G., Fox, N.C., 2008. Accuracy assessment of global and local atrophy measurement techniques with realistic simulated longitudinal Alzheimer's disease images. NeuroImage 42, 696–709.

Caramanos, Z., Fonov, V.S., Francis, S.J., Narayanan, S., Pike, G.B., Collins, D.L., Arnold, D.L., 2010. Gradient distortions in MRI: characterizing and correcting for their effects on SIENA-generated measures of brain volume change. NeuroImage 49, 1601–11.

Cardenas, V., Studholme, C., Gazdzinski, S., 2007. Deformation-based morphometry of brain changes in alcohol dependence and abstinence. Neuroimage.

Chard, D.T., Griffin, C.M., Rashid, W., Davies, G.R., Altmann, D.R., Kapoor, R., Barker, G.J., Thompson, a J., Miller, D.H., 2004. Progressive grey matter atrophy in clinically early relapsing-remitting multiple sclerosis. Multiple sclerosis (Houndmills, Basingstoke, England) 10, 387–91.

Chen, J.T., Narayanan, S., Collins, D.L., Smith, S.M., Matthews, P.M., Arnold, D.L., 2004. Relating neocortical pathology to disability progression in multiple sclerosis using MRI. NeuroImage 23, 1168–75.

Cover, K.S., Van Schijndel, R.A., Van Dijk, B.W., Redolfi, A., Knol, D.L., Frisoni, G.B., Barkhof, F., Vrenken, H., 2011. Assessing the reproducibility of the SienaX and Siena brain atrophy measures using the ADNI back-to-back MP-RAGE MRI scans. Psychiatry research 193, 182–90.

Derakhshan, M., Caramanos, Z., Giacomini, P.S., Narayanan, S., Maranzano, J., Francis, S.J., Arnold, D.L., Collins, D.L., 2010. Evaluation of automated techniques for the quantification of grey matter atrophy in patients with multiple sclerosis. NeuroImage 52, 1261–7.

Douaud, G., Smith, S., Jenkinson, M., Behrens, T., Johansen-Berg, H., Vickers, J., James, S., Voets, N., Watkins, K., Matthews, P.M., James, A., 2007. Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. Brain : a journal of neurology 130, 2375–86.

Fox, N.C., Cousens, S., Scahill, R., Harvey, R.J., Rossor, M.N., 2000. Using Serial Registered Brain Magnetic Resonance Imaging to Measure Disease Progression in Alzheimer Disease 57.

Freeborough, P., 1996. Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders. Journal of computer ....

Healy, B., Valsasina, P., Filippi, M., Bakshi, R., 2009. Sample size requirements for treatment effects using gray matter, white matter and whole brain volume in relapsing-remitting multiple sclerosis. Journal of neurology, neurosurgery, and psychiatry 80, 1218–23.

Horakova, D., Dwyer, M.G., Havrdova, E., Cox, J.L., Dolezal, O., Bergsland, N., Rimes, B., Seidl, Z., Vaneckova, M., Zivadinov, R., 2009. Gray matter atrophy and disability progression in patients with early relapsing-remitting multiple sclerosis: a 5-year longitudinal study. Journal of the neurological sciences 282, 112–9.

Hulst, H.E., Geurts, J.J.G., 2011. Gray matter imaging in multiple sclerosis: what have we learned? BMC neurology 11, 153.

Jack, C.R., Bernstein, M.A., Borowski, B.J., Gunter, J.L., Fox, N.C., Thompson, P.M., Schuff, N., Krueger, G., Killiany, R.J., Decarli, C.S., Dale, A.M., Carmichael, O.W., Tosun, D., Weiner, M.W., 2010. Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. Alzheimer's & dementia : the journal of the Alzheimer's Association 6, 212–20.

Jasperse, B., Valsasina, P., Neacsu, V., Knol, D.L., De Stefano, N., Enzinger, C., Smith, S.M., Ropele, S., Korteweg, T., Giorgio, A., Anderson, V., Polman, C.H., Filippi, M., Miller, D.H., Rovaris, M., Barkhof, F., Vrenken, H., 2007. Intercenter agreement of brain atrophy measurement in multiple sclerosis patients using manually-edited SIENA and SIENAX. Journal of magnetic resonance imaging : JMRI 26, 881–5.

Karas, G.., Burton, E.., Rombouts, S.A.R.., Van Schijndel, R.., O'Brien, J.., Scheltens, P.., McKeith, I.., Williams, D., Ballard, C., Barkhof, F., 2003. A comprehensive study of gray matter loss in patients with Alzheimer's disease using optimized voxel-based morphometry. NeuroImage 18, 895–907.

Keihaninejad, S., Heckemann, R.A., Fagiolo, G., Symms, M.R., Hajnal, J. V, Hammers, A., 2010. A robust method to estimate the intracranial volume across MRI field strengths (1.5T and 3T). NeuroImage 50, 1427–37.

Kurtzke, J.F., 1983. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). Neurology 33, 1444–1444.

Leung, K.K., Barnes, J., Modat, M., Ridgway, G.R., Bartlett, J.W., Fox, N.C., Ourselin, S., 2011. Brain MAPS: an automated, accurate and robust brain extraction technique using a template library. NeuroImage 55, 1091–108.

Lublin, F.D., Reingold, S.C., 1996. Defining the clinical course of multiple sclerosis: Results of an international survey. Neurology 46, 907–911.

Nakamura, K., Fisher, E., 2009. Segmentation of brain magnetic resonance images for measurement of gray matter atrophy in multiple sclerosis patients. NeuroImage 44, 769–76.

Nakamura, K., Fox, R., Fisher, E., 2011. CLADA: cortical longitudinal atrophy detection algorithm. NeuroImage 54, 278–89.

Oreja-Guevara, C., Rovaris, M., Iannucci, G., Valsasina, P., Caputo, D., Cavarretta, R., Sormani, M.P., Ferrante, P., Comi, G., Filippi, M., 2005. Progressive gray matter damage in patients with relapsing-remitting multiple sclerosis: a longitudinal diffusion tensor magnetic resonance imaging study. Archives of neurology 62, 578–84.

Pieperhoff, P., Südmeyer, M., Hömke, L., Zilles, K., Schnitzler, A., Amunts, K., 2008. Detection of structural changes of the human brain in longitudinally acquired MR images by deformation field morphometry: methodological analysis, validation and application. NeuroImage 43, 269–87.

Popescu, V., Battaglini, M., Hoogstrate, W.S., Verfaillie, S.C.J., Sluimer, I.C., Van Schijndel, R.A., Van Dijk, B.W., Cover, K.S., Knol, D.L., Jenkinson, M., Barkhof, F., De Stefano, N., Vrenken, H., 2012. Optimizing parameter choice for FSL-Brain Extraction Tool (BET) on 3D T1 images in multiple sclerosis. NeuroImage 61, 1484–94.

Sanfilipo, M.P., Benedict, R.H.B., Weinstock-Guttman, B., Bakshi, R., 2006. Gray and white matter brain atrophy and neuropsychological impairment in multiple sclerosis. Neurology 66, 685–92.

Sharma, S., Noblet, V., Rousseau, F., Heitz, F., Rumbach, L., Armspach, J.-P., 2009. Use of simulated atrophy for performance analysis of brain atrophy estimation approaches. Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention 12, 566–74.

Sluimer, J.D., Vrenken, H., Blankenstein, M.A., Fox, N.C., Scheltens, P., Barkhof, F., Van der Flier, W.M., 2008. Whole-brain atrophy rate in Alzheimer disease: identifying fast progressors. Neurology 70, 1836–41.

Smith, S., De Stefano, N., Jenkinson, M., Matthews, P., 2001. Normalized accurate measurement of longitudinal brain change. Journal of computer ... 25, 466–475.

Smith, S.M., 2002. Fast robust automated brain extraction. Human brain mapping 17, 143–55.

Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, A., De Stefano, N., 2002. Accurate, Robust, and Automated Longitudinal and Cross-Sectional Brain Change Analysis. NeuroImage 17, 479–489.

Sormani, M.P., Rovaris, M., Valsasina, P., Wolinsky, J.S., Comi, G., Filippi, M., 2004. Measurement error of two different techniques for brain atrophy assessment in multiple sclerosis. Neurology 62, 1432–1434.

Studholme, C., Drapaca, C., 2006. Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change. ... IEEE Transactions on.

Styner, M., Brechbuhler, C., 2000. Parametric estimate of intensity inhomogeneities applied to MRI. Medical Imaging, IEEE ....

Thal, L., Kantarci, K., Reiman, E., Klunk, W., Weiner, M., Zetterberg, H., Galasko, D., Pratico, D., Griffin, S., Schenk, D., Siemers, E., 2006. The role of biomarkers in clinical trials for Alzheimer disease. Alzheimer disease ... 20, 6–15.

Thompson, P., Hayashi, K., Sowell, E., 2004. Mapping cortical change in Alzheimer's disease, brain development, and schizophrenia. Neuroimage.

Thompson, P.M., Hayashi, K.M., De Zubicaray, G., Janke, A.L., Rose, S.E., Semple, J., Herman, D., Hong, M.S., Dittmer, S.S., Doddrell, D.M., Toga, A.W., 2003. Dynamics of Gray Matter Loss in Alzheimer's Disease. J. Neurosci. 23, 994–1005.

Valsasina, P., Benedetti, B., Rovaris, M., Sormani, M.P., Comi, G., Filippi, M., 2005. Evidence for progressive gray matter loss in patients with relapsing-remitting MS. Neurology 65, 1126–8.

Winkler, G., 2003. Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction (Google eBook). Springer.

Zhang, Y., Brady, M., Smith, S., 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE transactions on medical imaging 20, 45–57.

Zivadinov, R., Bakshi, R., 2004. Brain And Spinal Cord Atrophy In Multiple Sclerosis. Nova Science Pub Inc.

Zivadinov, R., Bergsland, N., Dolezal, O., Hussein, S., Seidl, Z., Dwyer, M.G., Vaneckova, M., Krasensky, J., Potts, J.A., Kalincik, T., Havrdová, E., Horáková, D., 2013. Evolution of Cortical and Thalamus Atrophy and Disability Progression in Early Relapsing-Remitting MS during 5 Years. AJNR. American journal of neuroradiology ajnr.A3503–.

Zivadinov, R., Havrdová, E., Bergsland, N., Tyblova, M., Hagemeier, J., Seidl, Z., Dwyer, M.G., Vaneckova, M., Krasensky, J., Carl, E., Kalincik, T., Horáková, D., 2013. Thalamic Atrophy is Associated with Development of Clinically Definite Multiple Sclerosis. Radiology.

Zivadinov, R., Heininen-Brown, M., Schirda, C. V, Poloni, G.U., Bergsland, N., Magnano, C.R., Durfee, J., Kennedy, C., Carl, E., Hagemeier, J., Benedict, R.H.B., Weinstock-Guttman, B., Dwyer, M.G., 2012. Abnormal subcortical deep-gray matter susceptibility-weighted imaging filtered phase measurements in patients with multiple sclerosis: a case-control study. NeuroImage 59, 331–9.

Zivadinov, R., Minagar, A., 2009. Evidence for gray matter pathology in multiple sclerosis: a neuroimaging approach. Journal of the neurological sciences 282, 1–4.

Zivadinov, R., Pirko, I., 2012. Advances in understanding gray matter pathology in multiple sclerosis: are we ready to redefine disease pathogenesis? BMC neurology 12, 9.

Zivadinov, R., Reder, a T., Filippi, M., Minagar, a, Stüve, O., Lassmann, H., Racke, M.K., Dwyer, M.G., Frohman, E.M., Khan, O., 2008. Mechanisms of action of disease-modifying agents and brain volume changes in multiple sclerosis. Neurology 71, 136–44.

Zivadinov, R., Rudick, R.A., De Masi, R., Nasuelli, D., Ukmar, M., Pozzi-Mucelli, R.S., Grop, A., Cazzato, G., Zorzon, M., 2001. Effects of IV methylprednisolone on brain atrophy in relapsing-remitting MS. Neurology 57, 1239–1247.

Table 1. Differences in whole brain, gray matter and white matter volume estimation change measures over 2-year follow-up in 34 healthy controls and 67 relapsing-remitting multiple sclerosis patients.

Whole brain measures	HC (n=34) Mean (SD) median (min/max)	RRMS (n=67) Mean (SD) median (min/max)	d	p value
SIENA PBVC	-0.30 (0.46) -0.31 (-0.99/1.08)	-1.71 (1.4) -1.52 (-7.61/2.71)	1.36	<0.0001
SIENAX PBVC	-0.22 (1.63) -0.42 (-4.22/2.94)	-2.78 (2.35) -2.55 (-9.3/2.13)	1.26	<0.0001
SIENAX-HW PBVC	-0.19 (0.93) -0.12 (-1.71/2.31)	-1.87 (1.56) -.1.66 (-6.77/1.63)	1.31	<0.0001
SIENAX-MTP PBVC	0.19 (0.59) -0.15 (-1.18/1.27)	-1.84 (1.33) -1.66 (-6.45/0.82)	1.59	<0.0001
GM measures				
SIENAX PGMVC*	-2.26 (1.87) -2.61 (-4.8/2.45)	-4.04 (3.14) -4.18 (-9.92/3.21)	0.69	0.003
SIENAX-HW PGMVC	-1.76 (1.87) -1.93 (-3.91/2.39)	-2.98 (2.45) -3.24 (-7.48/4.23)	0.35	0.082
SIENAX-MTP PGMVC	-0.64 (0.83) -0.78 (-1.96/1.66)	-1.87 (2.18) -1.48 (-9.76/3.16)	0.75	0.002
WM measures				
SIENAX PWMVC*	2.04 (2.34) 2.31 (-4.98/5.89)	-1.27 (3.77) -0.83 (-13.74/10.50)	1.08	<0.0001
SIENAX-HW PWMVC	2.06 (1.90) 1.89 (-2.82/6.41)	-0.48 (3.97) -0.27 (-13.18/5.28)	0.86	0.001
SIENAX-MTP PWMVC	1.86 (1.56) 1.89 (-2.18/6.41)	-0.56 (2.49) -0.41 (-7.07/5.28)	1.15	<0.0001

Legend: HC - healthy controls; RRMS - relapsing-remitting multiple sclerosis; SD - standard deviation; min - minimum; max - maximum; PBVC - percent brain volume change; GM - gray matter; MTP - multi-timepoint; HW - halfway; PGMVC - percent GM volume change; WM - white matter; PWMVC - percent WM volume change. d- Cohen's d represents the effect size. The differences between HC and MS groups were tested using Student's t-test.

* 4 SIENAX longitudinal analyses were excluded during quality control due to inadequate GM/WM segmentation. The analyses were performed with and without these cases excluded for SIENA, SIENAX-HW, and SIENAX-MTP measures. The results were nearly identical. Therefore, the results are presented with SIENAX failed cases included, because one of the main points of the study was to test whether SIENAX-MTP will result in fewer failures.

Table 2. Five year follow-up differences in whole brain, gray matter and white matter volume estimation change measures in 64 relapsing-remitting (RR) multiple sclerosis (MS) patients without disability progression and 64 RRMS patients with disability progression.

Whole brain measures	No disability progression MS group (n=64) Mean (SD) median (min/max)	Disability progression MS group (n=64) Mean (SD) median (min/max)	d	p value
SIENA PBVC	-2.72 (1.94) -2.34 (-8.95/1.18)	-4.85 (3.49) -3.97 (-17.21/0.79)	0.75	<0.0001
SIENAX PBVC	-5.56 (3.25) -4.86 (-16.12/-0.79)	-6.45 (3.82) -6.89 (-14.63/3.41)	0.25	0.160
SIENAX-HW PBVC	-2.56 (1.61) -2.31 (-8.15/0.99)	-3.79 (2.37) -3.42 (-9.98/0.19)	0.61	0.001
SIENAX-MTP PBVC	-2.08 (1.21) -1.88 (-5.72/0.35)	-3.51 (2.23) -3.10 (-10.37/0.24)	0.81	<0.0001
GM measures				
SIENAX PGMVC*	-4.18 (2.89) -4.11 (-16.42/4.82)	-5.41 (4.28) -5.01 (-18.06/4.14)	0.34	0.058
SIENAX-HW PGMVC	-4.60 (2.37) -4.58 (-8.96/0.51)	-5.11 (2.83) -4.77 (-13.21/-0.83)	0.20	0.261
SIENAX-MTP PGMVC	-3.92 (1.81) -4.37 (-7.87/0.46)	-5.26 (2.78) -4.45 (-14.61/-1.42)	0.57	0.002
WM measures				
SIENAX PWMVC*	1.47 (6.81) 2.75 (-21.87/13.84)	-2.49 (8.24) -2.42 (-24.91/19.58)	0.52	0.004
SIENAX-HW PWMVC	-0.09 (3.55) 0.48 (-14.26/5.70)	-2.14 (4.21) -1.39 (-13.81/5.86)	0.53	0.003
SIENAX-MTP PWMVC	0.40 (2.61) 0.88 (-9.13/4.82)	-1.52 (3.29) -0.75 (-11.22/4.35)	0.64	<0.0001

Legend: HC - healthy controls; RRMS - relapsing-remitting multiple sclerosis; SD - standard deviation; min - minimum; max - maximum; PBVC - percent brain volume change; GM - gray matter; MTP - multi-timepoint; HW - halfway; PGMVC - percent GM volume change; WM - white matter; PWMVC - percent WM volume change. **d -** Cohen's d represents the effect size. The differences between the MS groups were tested by using Student's t-test.

* 8 SIENAX longitudinal analyses were excluded during quality control due to inadequate GM/WM segmentation. The analyses were performed with and without these cases excluded for SIENA, SIENAX-HW, and SIENAX-MTP measures. The results were nearly identical. Therefore, the results are presented with SIENAX failed cases included, because one of the main points of the study was to test whether SIENAX-MTP will result in fewer failures.

Figure 1. Schematic view of the SIENAX-MTP processing pipeline. The arrows between the two LFAST components indicate communication of HMRF costs, allowing for more consistent tissue segmentation between time points.
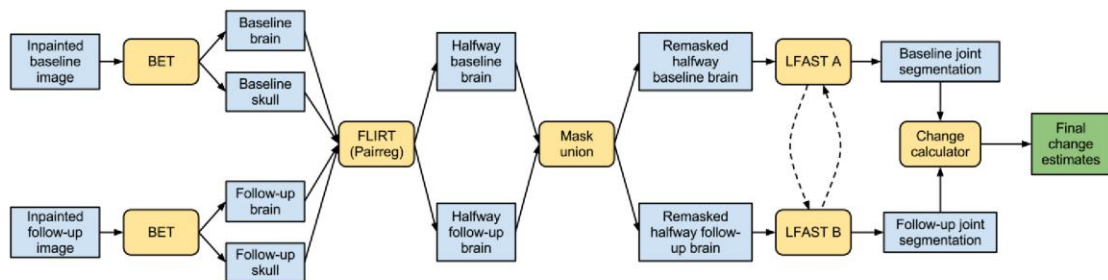
Figure 2. Representative images of the experimental simulation process, showing the 3D shape used to provide a parametric volume with gyrus-like ridges (left), a mid-level cross-section of a sample set of concentric "tissues" (middle), and a final bias field and noise corrupted image used for FAST and LFAST analysis (right).
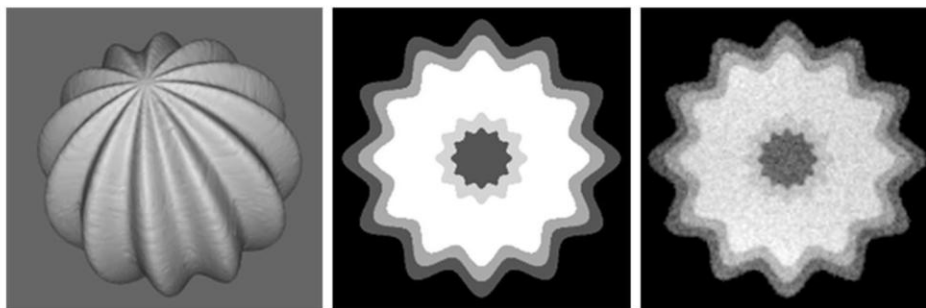
Figure 3. Simulation results. The top panel shows the results for FAST and some representative LFAST z factors across a variety of atrophy levels, averaged across five random runs each. The middle panel shows the associated mean absolute deviations (MADs) between runs for each atrophy level. The bottom panel shows a graph of the relationship between z factor and agreement with ground truth, and indicates a relatively stable plateau between 32 and 64 for which results are substantially better than conventional FAST.
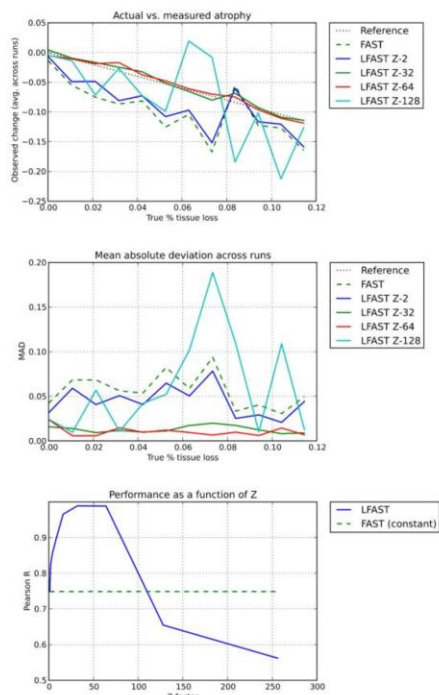
Figure 4. Representative results from scan-rescan analysis. Differences, which should be negligible, are quite noticeable for SIENAX but not for SIENAX-MTP. These visible difference voxels are sources or error in volumetric change measurements. Note that standard SIENAX assigns different classifications for tissue that has actually remained the same – e.g., in the deep gray matter.
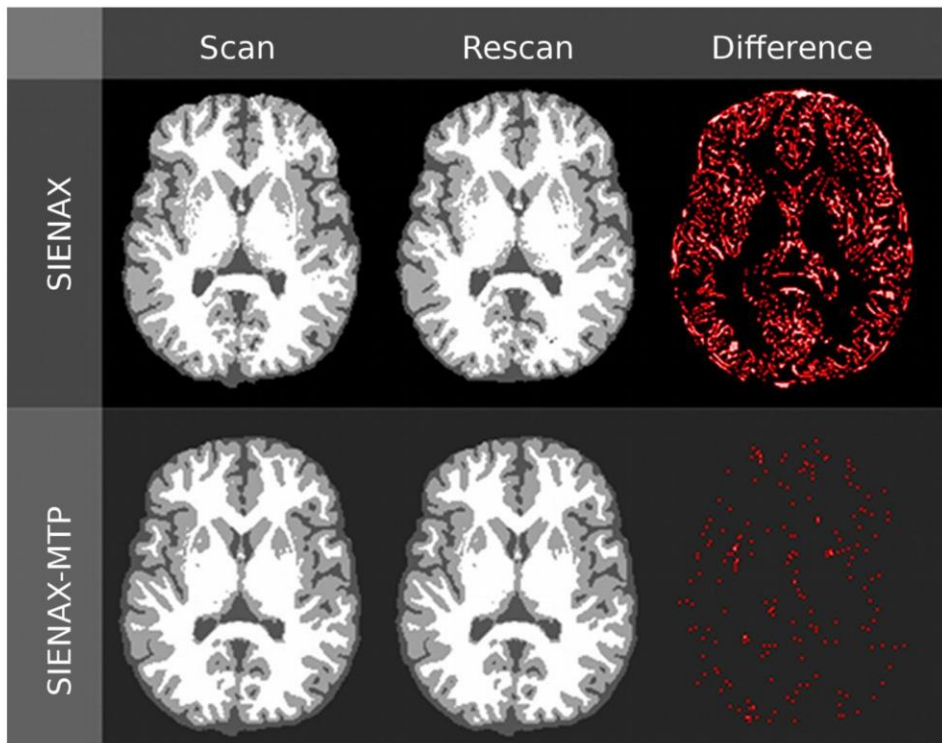
Appendix D – Additional background

This appendix provides more in-depth background and further references for readers interested in more information about MRI, MS, and image processing.

D.1 MRI

Magnetic resonance imaging (MRI) of human subjects was first successfully performed in 1977, with the first commercial whole-body scanner created in 1980 (Ai et al. 2012). In the three decades since, MRI has become an essential medical tool available in more than half of surveyed emergency hospitals in the US and UK (Ginde et al. 2008; Kane and Whiteley 2008).

The unique advantages of MRI that have driven its widespread use, include the fact that it is tomographic, capable of relatively high resolution of the order of a cubic millimetre or better, and that it is almost completely non-invasive. Equally notable, is MRI's ability to acquire a wide variety of data. By controlling the magnetic environment and probing tissue protons with radio frequency (RF) pulses, MRI is able to elucidate many different aspects of the chemical environment of hydrogen atoms.

At its most basic, MRI is made possible by the combination of three key components: a relatively large superconducting magnetic coil, a system for radio frequency (RF) transmission and detection, and a set of magnetic gradient-producing coils (Fig. D1).
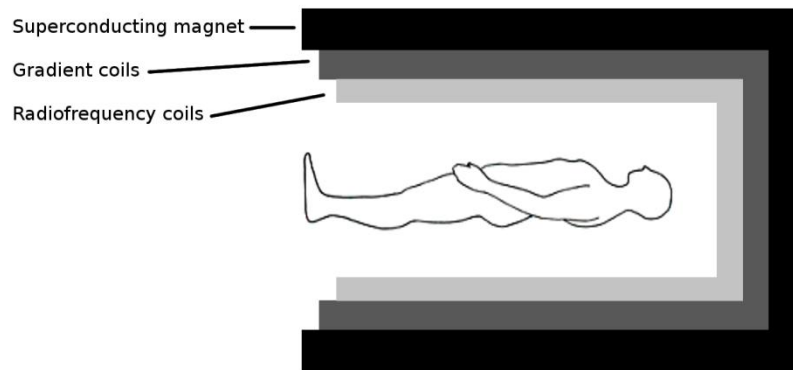
Figure D1: Schematic image of the core components of an MRI machine. The superconducting magnet aligns water protons and provides them with a characteristic frequency. The gradient coils allow for spatial encoding by varying the local magnetic field. The radiofrequency coils interact with the spinning protons to elicit and read electromagnetic signals.

It is impossible to do justice to the complex interplay at the heart of MRI image creation and acquisition in such a short space, but the basic principles are as follows:

1. Hydrogen nuclei (consisting of single protons) have a characteristic physical property called "spin" which can be conceived of as analogous to the angular momentum of a spinning ball or top. This spin also has a definable axis and direction (clockwise vs. counter-clockwise). This can of course be represented mathematically by a single positive or negative vector, but we treat the two separately here for simplicity. For the most part, the spin axes of hydrogen nuclei in vivo are oriented randomly.

2. When a subject is placed in the MRI scanner, these hydrogen nuclei also begin to precess around the main field (B0) axis in much the same way that a spinning top begins to "wobble" around the axis of gravity – the spin axis itself rotates around the field axis. The rate of this precession, called the Larmor frequency, is also a well-defined function on the magnetic field strength (e.g., 42.58 MHz/Tesla for hydrogen nuclei).

3. In addition, although the spin axes were original oriented randomly, they now begin to align more closely with the scanner, as this corresponds to a lower energy state. Note that in reality, only a very small excess of

protons align with instead of against the field, on the order of a few parts per million. This is also a more classical view of the system that ignores the more quantum-dynamical underlying principles. See (Haacke 1999) for a more detailed discussion.

4. When these precessing hydrogen nuclei are exposed to RF energy at or near their Larmor frequency (called an excitation pulse), their precession becomes phased, and their spin axes are also tipped away from the main magnetic field. With many protons spinning in phase and tipped away from the main field, a rotating net magnetic moment is produced in a plane perpendicular to the main field. In turn, this rotating magnetic moment induces a fluctuating electrical field that can be detected and measured with properly positioned RF receive coils and careful amplification.

5. Once "energized" by an RF field, protons do not simple remain at their newly "tipped" angle and stay in phase. Instead, they trade energy with their surroundings and with each other in such a way that they begin to "relax" back into alignment with the main field and also lose phase coherence with each other. The process of tipping back into alignment with the main field is called T1 relaxation, whereas the process of losing phase coherence is called T2 relaxation. Both occur simultaneously, with T2 often much faster than T1, and each dependent of various aspects of the local chemical environment. Because of this, the strength of the RF signal read out at from a particular portion of tissue any given time after the initial excitation pulse is largely governed by three factors: the overall number of hydrogen nuclei in the tissue (proton density, PD), the T1 relaxation rate of the tissue, and the T2 relaxation rate of the tissue. For simplicity, the important distinction between T2* and T2 is overlooked here.

6. Finally, recalling that protons are excitable only by RF pulses at or near their Larmor frequency, and that the Larmor frequency itself is dependent on the magnetic field, it is apparent that by varying the local magnetic field it is possible to ensure that only protons in one specific area will be excited. This is the job of gradient coils, which produce spatial variations in the otherwise homogeneous main magnetic field in order to excite single slices. Within a slice, carefully timed application of additional

gradient fields can code already-excited nuclei in each portion of tissue with a specific identifying frequency and phase. By exploiting the mathematics of the Fourier transform, a great deal of data can be acquired simultaneously in this manner.

7. Through the process of image reconstruction, a 3-dimensional grid (or series of 2-dimensional grids) is built up, composed of individual elements called voxels, for "volume elements" (analogous to pixels, but with depth as well as height and width).

Despite their brevity, the above principles provide some basic intuition for how conventional sequences like T1-weighted, T2-weighted, and PD-weighted images are acquired. By adjusting gradient/RF strength and application times, many variants can be acquired with different emphasis on scanning time, signal to noise ratio, and resolution. However, much more creative uses are possible with specialized RF/gradient techniques. For example, precise application of opposite polarity gradients can have no net effect on stationary nuclei, but dramatically affect moving nuclei, leading to diffusion contrast (Le Bihan 2003). Similarly, a well-tuned frequency offset in the RF field can be used to only excite protons in certain tissues, resulting in the potential for magnetization transfer contrast (Henkelman et al. 2001).

In practice, within a single one-hour scanning session, ten or more full conventional and/or non-conventional 3-dimensional tissue maps can be produced, each emphasizing a different aspect of the local chemical environment and/or sensitive to different pathology. For clinical routine imaging, they are generally read qualitatively by a trained radiologist in a suitable visual format. For quantitative analysis, they are transferred in a standard format called DICOM, which ensures standard recording formats for most relevant acquisition parameters.

A more complete description of the physics of MRI is beyond the scope of this document, but the interested reader is referred to (Hornak 2011) for an informal but informative overview, or to (Bernstein et al. 2004; Haacke 1999) for more comprehensive and rigorous treatment of the subject.

D.2 Multiple sclerosis

Multiple sclerosis (MS) is a chronic autoimmune disease of the central nervous system (CNS), which is characterized by inflammation, demyelination, axonal loss, and gliosis (Fig. D2). It is thought to affect more than 2 million individuals worldwide (Rosati 2001), and is the most common neurological disease amongst young adults. It commonly strikes individuals in their 20's or 30's (Koch-Henriksen and Sørensen 2010), and is usually characterized by acute clinical relapses followed by significant periods of remission. Eventually, though, most cases later enter a more debilitating progressive phase. The effects of MS are almost exclusively confined to the central nervous system, and the disease has as its hallmark focal pathological plaques with extensive demyelination located in otherwise myelinated central WM. Symptoms vary dramatically depending on affected areas, and include sensory deficits, muscle weakness, loss of balance, bladder dysregulation, visual changes, and cognitive impairment (Poser et al. 1982).
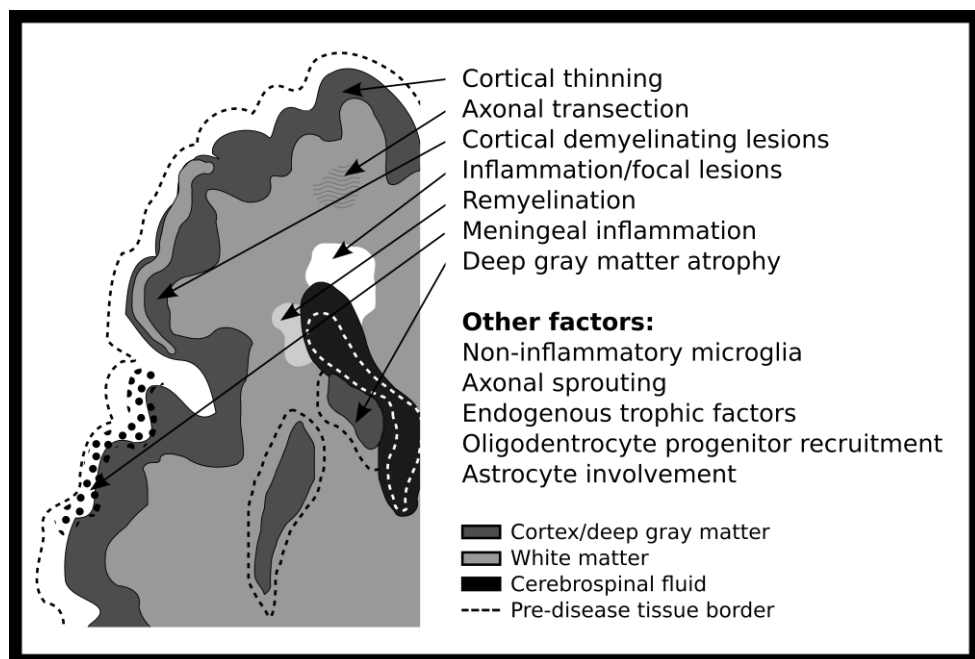
Figure D2: Heterogeneity of MS disease-related changes. In addition to the hallmark focal plaques, many other pathological and reparatory mechanisms occur both simultaneously and in series. The precise distribution, timing, and interaction between many of these processes is still imperfectly understood.

Although first identified by Jean-Martin Charcot in 1866 (Goetz et al. 1995), the aetiology of MS remains unknown. Indeed, over the 150 years since it was first discovered many disparate factors have been implicated. Environmental studies have revealed an association with latitude, with populations at higher latitudes being more susceptible (Simpson et al. 2011). Although there is evidence for an effect of sunlight exposure potentially mediated through vitamin D (Munger et al. 2004), there are also substantial east-west variations that indicate other climate related factors may be at work (Ebers and Sadovnick 1993). Infectious diseases have also been implicated, and there is convincing evidence that Epstein-Barr virus may play a role (Ascherio and Munger 2007). Other factors like degree of outdoor work (Kampman et al. 2007) and smoking (Riise et al. 2003; Zivadinov and Weinstock-Guttman 2009; Manouchehrinia et al. 2013; Salzer et al. 2013) have shown significant correlations with MS risk and outcome. Also, diet has been suggested and often discounted, although it seems that serum vitamin D levels may play a meaningful role (Munger et al. 2004). In addition to these environmental factors, there appears to be an important genetic component. MS co-occurrence in monozygotic twins is significantly more likely than in dizygotic twins or non-twin siblings (Willer et al. 2003). Additional studies with half-siblings have made a compelling case for a genetic/epigenetic basis for MS (Ebers et al. 2004). Beyond this, DNA microarray techniques have made it possible to cast a wide net in searching for likely contributors to MS. A recent large, genome-wide study identified specific risk alleles, including those of interleukin (IL2 and IL7) and human leukocyte antigen (HLA-DR) genes (Hafler et al. 2007), and another pointed to CD6, IRF8 and TNFRSF1A SNP locales (De Jager et al. 2009). However, despite the encouraging insights gained through these various studies, no overall comprehensive model of the disease exists and much about MS remains unknown, perhaps due to the complexity of epigenetic effects (Huynh and Casaccia 2013).

In recent years much of the research into MS has focused on the immunological aspects of the disease. There is overwhelming evidence that MS is associated with inflammatory markers in the CNS (Brück et al. 1995). As seems to be the case with most aspects of MS, the immune responses involved appear to be multifaceted. For some time, CD4+ helper T cells have been implicated, and are clearly involved in the disease (Viglietta and Baecher-Allan 2004). However,

many other cell types and immunomodulatory cascades are also involved in MS, including CD8+ T cells and CD56+ natural killer cells (Skulina and Schmidt 2004; Takahashi et al. 2004). Some populations like CD4+ Th1 T cells seem to have a disease-promoting effect, whereas others like CD25+CD4+ T cells are regulatory and suppress the immune response (Kasper and Shoemaker 2010). Of particular recent interest is the potential role of B cells in MS etiology (Magliozzi et al. 2007), and their potential as a therapeutic target (Hauser and Waubant 2008).

Histopathological work has also shed much light on the precise nature of MS. Perhaps most importantly, it has made increasingly clear the fact that MS involves considerably more tissue damage than the conventionally seen focal T2 lesions, as well as the fact that even within such lesions considerably more activity occurs than simple demyelination (Lassmann et al. 2007). It has also demonstrated extensive alterations in previously little-studied GM, and has informed better understanding of its more subtle lesions (Geurts et al. 2009; Filippi et al. 2013; Fischer et al. 2013). Meningeal inflammation has also been found (Choi et al. 2012), which may more directly influence GM pathology. Mechanisms of neurodegeneration beyond Wallerian degeneration have also been elucidated, including direct axonal transection (Trapp and Peterson 1998). Perhaps one of the most interesting recent findings from histopathology is an evaluation of newly formed MS lesions. Surprisingly, this showed extensive oligodendrocyte apoptosis and microglial activation but minimal lymphocyte or myelin macrophage presence, contrary to current disease models (Barnett and Prineas 2004). Follow-up study with more cases confirmed the absence of T or B cells and the presence of apoptotic oligodendrocytes in early expanding borders of MS lesions (Henderson et al. 2009). Further investigation demonstrated a role for oxidative stress in pre-inflammatory oligodendrocytes (Haider et al. 2011). These findings have led some to hypothesize that MS may not be an autoimmune disease, but rather a degenerative disease resulting in an immune component (Stys et al. 2012; Trapp and Nave 2008). Additional work has pointed to the importance of astrocytes in remyelination (Skripuletz et al. 2013), and has shown other factors that can inhibit (Stoffels et al. 2013) or promote (Yuen et al. 2013) it. Given all of this, it has become clear that MS is a multifactorial disease, and the ability to monitor all these aspects shown via

histopathology with in vivo MRI is an important goal (Filippi et al. 2012). Additionally, the simultaneous combination of histopathology and post-mortem MRI may drive further advancements (Kolasinski et al. 2012).

Arguably one of the greatest recent changes for those afflicted with MS is the fact that clinical care has progressed from passive observation, symptom management, and palliative care to meaningful disease modifying treatments that can significantly slow MS progress. These include large classes of immune-modifying agents such as the beta interferons (Jacobs and Cookfair 1996) and glatiramer acetate (Ford et al. 2006), as well as more modern approaches that limit lymphocyte movement or proliferation such as teriflunomide (O'Connor and Wolinsky 2011), fingolimod (Kappos et al. 2010), and natalizumab (Polman et al. 2006).

D.3 Medical image processing and analysis

For the first 15 years or so of its existence neurological MRI analysis was largely qualitive in nature, with most quantitative elements restricted to relatively simple calculations or low-level reconstruction applications. However in the late 1990s, clinicians and MRI physicists started to turn to the world of signal processing and statistics to develop new methodologies. Since then, neurological image analysis has become a sub-discipline in its own right (Dhawan 2011; Dougherty 2009); albeit an immature discipline that is developing. While medical imaging provides a unique set of challenges, it benefits from algorithms developed in other fields. Application and interpretation of these algorithms in a neurological context, however, remains major challenge. While the systems in other disciplines may be well understood, this is not the case with neurological disease, where much is unknown about the intracranial system. Consequently, techniques originally developed in other disciplines have to be tailored to MRI image analysis and validated against experimental data.

From the earliest stages of tomographic medical imaging (CT), computational power and sophisticated algorithms were required to reconstruct images from raw data, and this is still the case today. Where CT used the inverse Radon transform and iterative reconstruction, MRI now makes extensive use of the fast

Fourier transform (FFT). In the past few decades, though, computing speeds and memory have expanded to allow for considerably more possibilities than mere reconstruction. Because of this, the field of image analysis has grown dramatically, and has made quantitative interpretation and/or transformation of images a viable alternative or complement to qualitative radiological reads.

Arguably some of the most critical advances in the field have come from areas that are traditionally considered "pre-processing". In particular, co-registration – the mutual alignment of scans from different image types, time points, or subjects – has substantially matured to the point where it is highly reliable, due to the work of many independent investigators (Ashburner et al. 1997; Friston et al. 1995; Jenkinson et al. 2002; Woods and Grafton 1998). Co-registration is generally considered to be a linear operation, which although it can perform excellently between a single individual's scans, is not able to precisely align anatomy between different subjects. This concern can however be addressed by non-linear image warping, or normalization, which goes beyond affine transforms to allow detailed anatomy-specific warping/morphing. Such normalization may make use of hundreds or thousands of parameters, and is an extremely difficult mathematical and computational problem. Despite the challenges, though, significant progress was made even more than a decade ago (Woods and Grafton 1998) and improvements are still consistently being made. For example, much work has gone into invertible, diffeomorphic normalization (Klein et al. 2009; Murphy et al. 2011). Taken together, the value of fast, reliable, and reproducible post-acquisition alignment and normalization is difficult to overstate. Without such tools, it would be extremely difficult or impossible to track subtle changes across time, to evaluate the multi-spectral MRI characteristics of single tissue locations, or to compare localized changes across subjects.

With co-registration and normalization available, simultaneous progress has been made on the creation and standardization of neurological maps, such as the famous atlas of neurology produced by Talairach and Tourneaux (Talairach and Tournoux 1988). Although crude by modern standards, at the time this represented a great step forward. However, it was based on a single subject data, and it was not until much later that the ability to warp many subjects into a

common space in a highly detailed manner was possible. Only when this obstacle was overcome was it possible to create much more accurate and representative atlases (Grabner et al. 2006) and ICBM 452 non-linear atlases (Mazziotta et al. 2001). With the ability to create high-quality atlases came the ability to perform group-wise spatial analyses. While early work was based on "collapsing" subject data into single metrics, the ability to normalize subjects into a common space led to the ability to precisely evaluate differences in specific, small anatomical regions all the way down to the voxel level. Some of the key techniques in this area include voxel based morphometry (Ashburner and Friston 2000) and deformation-based mapping (Chung et al. 2001). These techniques have been the basis for many independent investigations into morphometric changes in a variety of diseases and in normal aging (Draganski et al. 2004; Honea and Crow 2005; Prinster et al. 2006; Thompson et al. 2003).

At the same time, improvements in image processing algorithms have led to accurate, reproducible tissue segmentation and object extraction techniques. These allow for classification on a voxel-by-voxel basis into specific tissue compartments or for labeling of voxels as specific anatomical structures. Examples of segmentation techniques include EMS (Van Leemput et al. 2001), FAST (Zhang et al. 2001), and SPM unified segmentation (Ashburner and Friston 2005). Examples of object extraction tools include FIRST (Patenaude et al. 2011) and cortical parcellation in Freesurfer (Fischl 2012). Other advances have occurred in lower-level processing areas, including a variety of techniques for image enhancement and/or restoration. For example, non-linear smoothing has provided a "best of both worlds" ability to diminish measurement noise without paying the classical price of lost spatial resolution (Black and Sapiro 1998; Smith and Brady 1997). Other techniques have focused more on ameliorating MRI-specific artefacts, including techniques for intra-session motion correction (Jenkinson et al. 2002), eddy current correction (Bodammer and Kaufmann 2004), and susceptibility-induced distortion unwarping (Jezzard and Balaban 1995). Finally, a number of more sequence-specific techniques have also developed symbiotically with MRI physics, including DTI-based tractography (Ciccarelli et al. 2008), fMRI ICA (Calhoun et al. 2009) and network analysis (Bullmore and Sporns 2009).

Appendix D references

Ai, T., Morelli, J., Hu, X. and Hao, D. 2012. A Historical Overview of Magnetic Resonance Imaging, Focusing on Technological Innovations. *Investigative radiology* 47(12), pp. 725–41.

Ascherio, A. and Munger, K.L. 2007. Environmental risk factors for multiple sclerosis. Part I: the role of infection. *Annals of neurology* 61(4), pp. 288–99.

Ashburner, J. and Friston, K. 2005. Unified segmentation. *Neuroimage* 26(3), pp. 839–851.

Ashburner, J. and Friston, K. 2000. Voxel-based morphometry—the methods. *Neuroimage* 11(6), pp. 805–821.

Ashburner, J., Neelin, P., Collins, D., Evans, A. and Friston, K. 1997. Incorporating prior knowledge into image registration. *Neuroimage* 6(4), pp. 344–352.

Barnett, M.H. and Prineas, J.W. 2004. Relapsing and remitting multiple sclerosis: pathology of the newly forming lesion. *Annals of neurology* 55(4), pp. 458–68.

Bernstein, M.A., King, K.F. and Zhou, X.J. 2004. *Handbook of MRI Pulse Sequences*. Elsevier.

Le Bihan, D. 2003. Looking into the functional architecture of the brain with diffusion MRI. *Nature reviews. Neuroscience* 4(6), pp. 469–80.

Black, M. and Sapiro, G. 1998. Robust anisotropic diffusion. *Image Processing, IEEE Transactions on* 7(3), pp. 421–432.

Bodammer, N. and Kaufmann, J. 2004. Eddy current correction in diffusion-weighted imaging using pairs of images acquired with opposite diffusion gradient polarity. *Magnetic Resonance in Medicine* 51(1), pp. 188–193.

Brück, W., Porada, P. and Poser, S. 1995. Monocyte/macrophage differentiation in early multiple sclerosis lesions. *Annals of neurology* 38(5), pp. 788–796.

Bullmore, E. and Sporns, O. 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* 10(3), pp. 186–198.

Calhoun, V., Liu, J. and Adalı, T. 2009. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *Neuroimage* 45(1), pp. S163–S172.

Choi, S.R., Howell, O.W., Carassiti, D., Magliozzi, R., Gveric, D., Muraro, P. a, Nicholas, R., Roncaroli, F. and Reynolds, R. 2012. Meningeal inflammation plays a role in the pathology of primary progressive multiple sclerosis. *Brain* 135(Pt 10), pp. 2925–37.

Chung, M., Worsley, K., Paus, T. and Cherif, C. 2001. A unified statistical approach to deformation-based morphometry. *NeuroImage* 14(3), pp. 595–606.

Ciccarelli, O., Catani, M., Johansen-Berg, H., Clark, C. and Thompson, A. 2008. Diffusion-based tractography in neurological disorders: concepts, applications, and future developments. *Lancet neurology* 7(8), pp. 715–27.

Dhawan, A.P. 2011. *Medical Image Analysis*. John Wiley & Sons.

Dougherty, G. 2009. *Digital Image Processing for Medical Applications*. Cambridge University Press.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U. and May, A. 2004. Neuroplasticity: changes in grey matter induced by training. *Nature* 427(6972), pp. 311–2.

Ebers, G., Sadovnick, A. and Dyment, D. 2004. Parent-of-origin effect in multiple sclerosis: observations in half-siblings. *The Lancet* 363(9423), pp. 1773–1774.

Ebers, G.C. and Sadovnick, A.D. 1993. The Geographic Distribution of Multiple Sclerosis: A Review. *Neuroepidemiology* 12(1), pp. 1–5.

Filippi, M., Preziosa, P., Pagani, E., Copetti, M., Mesaros, S., Colombo, B., Horsfield, M. A., Falini, A., Comi, G., Lassmann, H. and Rocca, M. A. 2013. Microstructural magnetic resonance imaging of cortical lesions in multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England)* 19(4), pp. 418–26.

Filippi, M., Rocca, M. A., Barkhof, F., Brück, W., Chen, J.T., Comi, G., DeLuca, G., De Stefano, N., Erickson, B.J., Evangelou, N., Fazekas, F., Geurts, J.J.G., Lucchinetti, C., Miller, D.H., Pelletier, D., Popescu, B.F.G. and Lassmann, H. 2012. Association between pathological and MRI findings in multiple sclerosis. *Lancet neurology* 11(4), pp. 349–60.

Fischer, M.T., Wimmer, I., Höftberger, R., Gerlach, S., Haider, L., Zrzavy, T., Hametner, S., Mahad, D., Binder, C.J., Krumbholz, M., Bauer, J., Bradl, M. and Lassmann, H. 2013. Disease-specific molecular events in cortical multiple sclerosis lesions. *Brain : a journal of neurology* 136(Pt 6), pp. 1799–815.

Fischl, B. 2012. FreeSurfer. *Neuroimage* 62(2), pp. 774–781.

Ford, C., Johnson, K. and Lisak, R. 2006. A prospective open-label study of glatiramer acetate: over a decade of continuous use in multiple sclerosis patients. *Multiple sclerosis* 12(3), pp. 309–320.

Friston, K., Ashburner, J. and Frith, C. 1995. Spatial registration and normalization of images. *Human brain mapping* 3(3), pp. 165–189.

Geurts, J.J.G., Stys, P.K., Minagar, A., Amor, S. and Zivadinov, R. 2009. Gray matter pathology in (chronic) MS: modern views on an early observation. *Journal of the neurological sciences* 282(1-2), pp. 12–20.

Ginde, A.A., Foianini, A., Renner, D.M., Valley, M. and Camargo, C.A. 2008. Availability and quality of computed tomography and magnetic resonance imaging equipment in U.S. emergency departments. *Academic emergency medicine* 15(8), pp. 780–3.

Goetz, C.G., Bonduelle, M. and Gelfand, T. 1995. *Charcot: Constructing Neurology*. Oxford University Press.

Grabner, G., Janke, A. and Budge, M. 2006. Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults. In: *Medical Image Computing and Computer-Assisted Intervention*. Berlin, pp. 58–66.

Haacke, E.M. 1999. *Magnetic Resonance Imaging: Physical Principles and Sequence Design*. Wiley.

Hafler, D.A., Compston, A., Sawcer, S., Lander, E.S., Daly, M.J., De Jager, P.L. and De Bakker, P.I. 2007. Risk Alleles for Multiple Sclerosis Identified by a Genomewide Study. *The New England journal of medicine* 357(9), p. 851.

Haider, L., Fischer, M.T., Frischer, J.M., Bauer, J., Höftberger, R., Botond, G., Esterbauer, H., Binder, C.J., Witztum, J.L. and Lassmann, H. 2011. Oxidative damage in multiple sclerosis lesions. *Brain* 134(Pt 7), pp. 1914–24.

Hauser, S. and Waubant, E. 2008. B-cell depletion with rituximab in relapsing–remitting multiple sclerosis. *New England Journal of Medicine* 358(7), pp. 676–688.

Henderson, A.P.D., Barnett, M.H., Parratt, J.D.E. and Prineas, J.W. 2009. Multiple sclerosis: distribution of inflammatory cells in newly forming lesions. *Annals of neurology* 66(6), pp. 739–53.

Henkelman, R., Stanisz, G. and Graham, S. 2001. Magnetization transfer in MRI: a review. *NMR in Biomedicine* 14(2), pp. 57–64.

Honea, R. and Crow, T. 2005. Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *American Journal of Psychiatry* 162(12), pp. 2233–2245.

Hornak, J.P. 2011. The Basics of MRI. http://www.cis.rit.edu/htbooks/mri/ [Online].

Huynh, J.L. and Casaccia, P. 2013. Epigenetic mechanisms in multiple sclerosis: implications for pathogenesis and treatment. *Lancet neurology* 12(2), pp. 195–206.

Jacobs, L. and Cookfair, D. 1996. Intramuscular interferon beta-1a for disease progression in relapsing multiple sclerosis. *Annals of neurology* 39(3), pp. 285–294.

De Jager, P.L., Jia, X., Wang, J., de Bakker, P.I.W., Ottoboni, L., Aggarwal, N.T., Piccio, L., Raychaudhuri, S., Tran, D., Aubin, C., Briskin, R., Romano, S., Baranzini, S.E., McCauley, J.L., Pericak-Vance, M.A., Haines, J.L., Gibson, R.A., Naeglin, Y., Uitdehaag, B., Matthews, P.M., Kappos, L., Polman, C., McArdle, W.L., Strachan, D.P., Evans, D., Cross, A.H., Daly, M.J., Compston, A., Sawcer, S.J., Weiner, H.L., Hauser, S.L., Hafler, D.A. and Oksenberg, J.R. 2009. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nature genetics* 41(7), pp. 776–82.

Jenkinson, M., Bannister, P., Brady, M. and Smith, S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17(2), pp. 825–841.

Jezzard, P. and Balaban, R. 1995. Correction for geometric distortion in echo planar images from B0 field variations. *Magnetic resonance in medicine* 34(1), pp. 65–73.

Kampman, M.T., Wilsgaard, T. and Mellgren, S.I. 2007. Outdoor activities and diet in childhood and adolescence relate to MS risk above the Arctic Circle. *Journal of neurology* 254(4), pp. 471–7.

Kane, I. and Whiteley, W. 2008. Availability of CT and MR for assessing patients with acute stroke. *Cerebrovascular Diseases* 25(4), pp. 375–377.

Kappos, L., Radue, E. W., O'Connor, P., Polman, C., Hohlfeld, R., Calabresi, P. and Selmaj, K. 2010. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. *New England Journal of Medicine* 362(5), pp. 387–401.

Kasper, L.H. and Shoemaker, J. 2010. Multiple sclerosis immunology: The healthy immune system vs the MS immune system. *Neurology* 74(1 Supplement 1), pp. S2–S8.

Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M., Christensen, G.E., Collins, D.L., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J. and Parsey, R. V 2009. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage* 46(3), pp. 786–802.

Koch-Henriksen, N. and Sørensen, P.S. 2010. The changing demographic pattern of multiple sclerosis epidemiology. *Lancet neurology* 9(5), pp. 520–32.

Kolasinski, J., Stagg, C.J., Chance, S. A., Deluca, G.C., Esiri, M.M., Chang, E., Palace, J. a, McNab, J. a, Jenkinson, M., Miller, K.L. and Johansen-Berg, H. 2012. A combined post-mortem magnetic resonance imaging and quantitative histological study of multiple sclerosis pathology. *Brain : a journal of neurology* 135(Pt 10), pp. 2938–51.

Lassmann, H., Brück, W. and Lucchinetti, C. 2007. The immunopathology of multiple sclerosis: an overview. *Brain pathology* 17(2), pp. 210–218.

Van Leemput, K., Maes, F., Vandermeulen, D., Colchester, A. and Suetens, P. 2001. Automated segmentation of multiple sclerosis lesions by model outlier detection. *Medical Imaging, IEEE Transactions on* 20(8), pp. 677–688.

Magliozzi, R., Howell, O., Vora, A., Serafini, B., Nicholas, R., Puopolo, M., Reynolds, R. and Aloisi, F. 2007. Meningeal B-cell follicles in secondary progressive multiple sclerosis associate with early onset of disease and severe cortical pathology. *Brain* 130(Pt 4), pp. 1089–104.

Manouchehrinia, A., Tench, C.R., Maxted, J., Bibani, R.H., Britton, J. and Constantinescu, C.S. 2013. Tobacco smoking and disability progression in multiple sclerosis: United Kingdom cohort study. *Brain : a journal of neurology* 136(Pt 7), pp. 2298–304.

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K. and Woods, R. 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 356(1412), pp. 1293–1322.

Munger, K.L., Zhang, S.M., O'Reilly, E., Hernan, M.A., Olek, M.J., Willett, W.C. and Ascherio, A. 2004. Vitamin D intake and incidence of multiple sclerosis. *Neurology* 62(1), pp. 60–65.

Murphy, K., van Ginneken, B., Reinhardt, J.M., Kabus, S., Ding, K., et al. 2011. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. *IEEE transactions on medical imaging* 30(11), pp. 1901–20.

O'Connor, P. and Wolinsky, J. 2011. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. *New England Journal of Medicine* 365(14), pp. 1293–1303.

Patenaude, B., Smith, S.S.M., Kennedy, D.N.D. and Jenkinson, M. 2011. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 56(3), pp. 907–22.

Polman, C.H., O'Connor, P.W., Havrdova, E., Hutchinson, M., Kappos, L., Miller, D.H. and Phillips, J.T. 2006. A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *New England Journal of Medicine* 354(9), pp. 899–910.

Poser, S., Raun, N. and Poser, W. 1982. Age at onset, initial symptomatology and the course of multiple sclerosis. *Acta Neurologica Scandinavica* 66(3), pp. 355–362.

Prinster, A., Quarantelli, M. and Orefice, G. 2006. Grey matter loss in relapsing–remitting multiple sclerosis: a voxel-based morphometry study. *Neuroimage* 29(3), pp. 859–867.

Riise, T., Nortvedt, M. and Ascherio, A. 2003. Smoking is a risk factor for multiple sclerosis. *Neurology* 61(8), pp. 1122–1124.

Rosati, G. 2001. The prevalence of multiple sclerosis in the world: an update. *Neurological sciences* 22(2), pp. 117–139.

Salzer, J., Hallmans, G., Nyström, M., Stenlund, H., Wadell, G. and Sundström, P. 2013. Smoking as a risk factor for multiple sclerosis. *Multiple sclerosis (Houndmills, Basingstoke, England)* 19(8), pp. 1022–7.

Simpson, S., Blizzard, L., Otahal, P., Van der Mei, I. and Taylor, B. 2011. Latitude is significantly associated with the prevalence of multiple sclerosis: a meta-analysis. *Journal of neurology, neurosurgery, and psychiatry* 82(10), pp. 1132–41.

Skripuletz, T., Hackstette, D., Bauer, K., Gudi, V., Pul, R., Voss, E., Berger, K., Kipp, M., Baumgärtner, W. and Stangel, M. 2013. Astrocytes regulate myelin clearance through recruitment of microglia during cuprizone-induced demyelination. *Brain : a journal of neurology* 136(Pt 1), pp. 147–67.

Skulina, C. and Schmidt, S. 2004. Multiple sclerosis: brain-infiltrating CD8+ T cells persist as clonal expansions in the cerebrospinal fluid and blood. *Proceedings of the National Academy of Sciences of the United States of America* 101(8), pp. 2428–2433.

Smith, S. and Brady, J. 1997. SUSAN—A new approach to low level image processing. *International journal of computer vision* 23(1), pp. 45–78.

Stoffels, J.M.J., de Jonge, J.C., Stancic, M., Nomden, A., van Strien, M.E., Ma, D., Sisková, Z., Maier, O., Ffrench-Constant, C., Franklin, R.J.M., Hoekstra, D., Zhao, C. and Baron, W. 2013. Fibronectin aggregation in multiple sclerosis lesions impairs remyelination. *Brain : a journal of neurology* 136(Pt 1), pp. 116–31.

Stys, P.K., Zamponi, G.W., Minnen, J. van and Geurts, J.J. 2012. Will the real multiple sclerosis please stand up? *Nature Reviews Neuroscience* 13(7), pp. 507–514.

Takahashi, K., Aranami, T. and Endoh, M. 2004. The regulatory role of natural killer cells in multiple sclerosis. *Brain* 127(9), pp. 1917–1927.

Talairach, J. and Tournoux, P. 1988. *Co-planar Stereotaxic Atlas of the Human Brain: 3-dimensional Proportional System*. Thieme Medical Pub.

Thompson, P.M., Hayashi, K.M., de Zubicaray, G., Janke, A.L., Rose, S.E., Semple, J., Herman, D., Hong, M.S., Dittmer, S.S., Doddrell, D.M. and Toga, A.W. 2003. Dynamics of Gray Matter Loss in Alzheimer's Disease. *J. Neurosci.* 23(3), pp. 994–1005.

Trapp, B. and Nave, K. 2008. Multiple sclerosis: an immune or neurodegenerative disorder? *Annu. Rev. Neurosci.* 31, pp. 247–269.

Trapp, B. and Peterson, J. 1998. Axonal transection in the lesions of multiple sclerosis. *New England Journal of Medicine* 338(5), pp. 278–285.

Viglietta, V. and Baecher-Allan, C. 2004. Loss of functional suppression by CD4+ CD25+ regulatory T cells in patients with multiple sclerosis. *The Journal of experimental medicine* 199(7), pp. 971–979.

Willer, C.J., Dyment, D.A., Risch, N.J., Sadovnick, A.D. and Ebers, G.C. 2003. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proceedings of the National Academy of Sciences of the United States of America* 100(22), pp. 12877–82.

Woods, R. and Grafton, S. 1998. Automated image registration: II. Intersubject validation of linear and nonlinear models. *ournal of computer assisted tomography* 22(1), pp. 153–165.

Yuen, T.J., Johnson, K.R., Miron, V.E., Zhao, C., Quandt, J., Harrisingh, M.C., Swire, M., Williams, A., McFarland, H.F., Franklin, R.J.M. and Ffrench-Constant, C. 2013. Identification of endothelin 2 as an inflammatory factor that promotes central nervous system remyelination. *Brain : a journal of neurology* 136(Pt 4), pp. 1035–47.

Zhang, Y., Brady, M. and Smith, S. 2001. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging* 20(1), pp. 45–57.

Zivadinov, R. and Weinstock-Guttman, B. 2009. Smoking is associated with increased lesion volumes and brain atrophy in multiple sclerosis. *Neurology* 73(7), pp. 504–510.